Cognition 109 (2008) 295–315

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

# Can an agent's false belief be corrected by an appropriate communication? Psychological reasoning in 18-month-old infants

Hyun-joo Song [a,*], Kristine H. Onishi [b], Renée Baillargeon [c], Cynthia Fisher [c]

[a] Department of Psychology, Yonsei University, 134 Shinchon-dong Seodaemun-gu, Seoul 120-749, South Korea
[b] Department of Psychology, McGill University, 1205 Dr Penfield, Montreal, Quebec, H3A 1B1 Canada
[c] Department of Psychology, University of Illinois, 603 East Daniel Street, Champaign, IL 61820, USA

A R T I C L E I N F O

*Article history:*
Received 22 August 2006
Revised 2 August 2008
Accepted 22 August 2008

*Keywords:*
Infant cognition
Cognitive development
Theory of mind
Belief
Action comprehension
Psychological reasoning

A B S T R A C T

Do 18-month-olds understand that an agent's false belief can be corrected by an appropriate, though not an inappropriate, communication? In Experiment 1, infants watched a series of events involving two agents, a ball, and two containers: a box and a cup. To start, agent1 played with the ball and then hid it in the box, while agent2 looked on. Next, in agent1's absence, agent2 moved the ball from the box to the cup. When agent1 returned, agent2 told her "The ball is in the cup!" (informative-intervention condition) or "I like the cup!" (uninformative-intervention condition). During test, agent1 reached for either the box (box event) or the cup (cup event). In the informative-intervention condition, infants who saw the box event looked reliably longer than those who saw the cup event; in the uninformative-intervention condition, the reverse pattern was found. These results suggest that infants expected agent1's false belief about the ball's location to be corrected when she was told "The ball is in the cup!", but not "I like the cup!". In Experiment 2, agent2 simply pointed to the ball's new location, and infants again expected agent1's false belief to be corrected. These and control results provide additional evidence that infants in the second year of life can attribute false beliefs to agents. In addition, the results suggest that by 18 months of age infants expect agents' false beliefs to be corrected by relevant communications involving words or gestures.



## 1. Introduction

Our ability to make sense of agents' actions rests in large part on our ability to understand the mental states that underlie their actions. Critical to this understanding is the recognition that mental states sometimes conflict with reality: for example, an agent acting in a setting may hold a false belief about the location or contents of an object in the setting, or she may perceive a deceptive object as one thing when it is in fact another.

Developmental psychologists have long been interested in determining at what age children become able to attribute false beliefs and false perceptions to agents. Initial investigations suggested that this ability did not emerge until about 4 years of age (e.g., Flavell, 1988; Gopnik & Wellman, 1994; Perner, 1991, 1995; Wellman, Cross, & Watson, 2001). This evidence came primarily from tasks in which children are asked direct questions about an agent's beliefs or perceptions (e.g., Baron-Cohen, Leslie, & Frith, 1985; Gopnik & Astington, 1988; Moore, Pure, & Furrow, 1990; Perner, Leekam, & Wimmer, 1987; Wimmer & Perner, 1983). For example, in a standard false-belief task (e.g., Baron-Cohen et al., 1985), children listen to a story enacted with props: a first character hides a toy in locationA and leaves; while she is gone, a second character moves the toy to locationB. When asked where the first character will look for her toy upon her return, most 4-year-olds correctly point to locationA; in contrast, most 3-year-olds point to locationB, suggesting that they do not understand that the first character will hold a false belief about the toy's location. Similarly, in a standard

* Corresponding author. Tel.: +82 2 2123 2449.
*E-mail address:* hsong@yonsei.ac.kr (H. Song).

0010-0277/$ - see front matter 
doi:10.1016/j.cognition.2008.08.008

false-perception task (e.g., Gopnik & Astington, 1988), children explore a fake object such as a sponge that looks like a rock. When asked how the object will appear to a naive agent, most 3-year-olds demonstrate no understanding that the agent will be misled by the object's appearance.

However, recent investigations conducted with novel tasks designed to tap children's *spontaneous* (as opposed to *elicited*) responses suggest that the ability to attribute false beliefs and false perceptions may emerge long before 4 years of age. To date, two kinds of spontaneous tasks have yielded positive results with children aged 2 years and younger: anticipatory-looking and violation-of-expectation (VOE) tasks. In the next sections, we review this evidence and then introduce the present research, which built on these findings.

### *1.1. Anticipatory-looking tasks*

Clements and Perner (1994) were the first to report that children younger than 4 years of age demonstrate some understanding of false belief in anticipatory-looking tasks. Children aged 29–53 months heard a story similar to that in a standard false-belief task, enacted with props. However, instead of being asked directly where the first character would look for her toy upon her return, children heard the experimenter say to himself "I wonder where she's going to look?". Immediately following this anticipation prompt, children's looking behavior was examined to determine whether they looked at locationA, where the first character's false belief should lead her to search, or at locationB, where the toy was currently hidden. Beginning at 35 months, most children looked at locationA, suggesting that some understanding of false belief is present by 3 years of age. Garnham (née Clements) and Ruffman (2001) subsequently obtained similar results with children aged 25–49 months, though Ruffman and Perner (2005) later expressed skepticism about these results.

Southgate, Senju, and Csibra (2007) recently tested 25-month-old children in a novel nonverbal anticipatory-looking task using an eye-tracker. The children first watched two videotaped familiarization events in which a bear puppet hid a toy in one of two boxes while a female agent looked on. The agent wore a visor and sat behind a panel between two small, closed doors, one above each box; only the agent's head was visible above the panel. After the bear hid the toy, the two doors lit up, and then the agent opened the correct door to retrieve the toy. The bear hid the toy in the left box in the first familiarization trial and in the right box in the second trial; only children who correctly anticipated which door the agent would use in the second trial (i.e., who looked at the right door before the agent opened it) were retained in the experiment. During the test trial, the agent saw the bear hide the toy in the left box. At that point, a phone rang behind the agent, who turned toward the sound; while the agent was facing away, the bear removed the toy from the left box, briefly hid it in the right box, and then left with it. The phone then stopped ringing, the agent turned back toward the boxes, and the doors lit up. Most children correctly anticipated the agent's behavior and looked at the door above the left box, where she falsely believed the toy to be hidden. Similar results were obtained in another condition where the agent falsely believed the toy to be hidden in the right box. Together, these results suggest that, by 2 years of age, children already possess some understanding of false belief.

### *1.2. Violation-of-expectation tasks*

Evidence that children younger than 2 years also possess some understanding of false belief comes from VOE reports by Onishi and Baillargeon (2005) and Surian, Caldi, and Sperber (2007); two other reports, by Song and Baillargeon (in press) and Scott and Baillargeon (in press), are described in later sections.

In Onishi and Baillargeon (2005), 15-month-old infants first received three familiarization trials involving live events. In the first trial, a toy stood on an apparatus floor between a green and a yellow box; the boxes' openings faced each other and were covered with fringe. A female agent opened doors in the back wall of the apparatus, played with the toy for a few seconds, and then hid it inside the green box; the agent then paused, with her hand inside the box, until the trial ended. During the second and third familiarization trials, the agent reached inside the green box (as though to grasp the toy she had previously hidden there) and then paused. Next, the infants received a belief-induction trial; there were four versions of this trial, designed to yield two true-belief (TB) and two false-belief (FB) conditions: the agent could believe, truly or falsely, that the toy was hidden in the green or in the yellow box. For example, in the FB-green condition, the toy moved from the green to the yellow box in the agent's absence. In the FB-yellow condition, the agent watched the toy move to the yellow box and then left; in her absence, the toy returned to the green box. During the test trial, the agent reached inside the green (green-box event) or the yellow (yellow-box event) box and then paused. In each condition, infants expected the agent to reach where she believed—truly or falsely—the toy to be hidden, and they looked reliably longer when she reached to the other location. Thus, in the FB-green condition the infants who saw the yellow-box event looked reliably longer than those who saw the green-box event, whereas in the FB-yellow condition the reverse looking pattern was found.

In Surian et al. (2007), 13-month-old infants first received four familiarization trials involving computer-animated events. In each trial, a caterpillar moved into view and watched an experimenter's hand hide one object (e.g., an apple) behind one screen and a different object (e.g., a piece of cheese) behind another screen; the caterpillar always approached the same screen to chew on the same, preferred object. During the test trial, the hand hid the two objects in the reverse locations, either after (TB condition) or before (FB condition) the caterpillar arrived on the scene. The infants in the TB condition expected the caterpillar to approach the preferred object in its new location, and they looked reliably longer when it approached the original location instead. In contrast, the infants in the FB condition expected the caterpillar to approach the original location—since it had not seen the hand hide the objects and must falsely assume that they had been hidden in the same locations as before—and thus they looked reliably longer when it approached the new location.

Together, these experiments suggest that, by 13–15 months of age, infants can already attribute false beliefs to agents, whether human or non-human. As we discuss more fully in the general discussion (Section 4), we assume that infants' understanding of false belief operates largely without explicit or conscious awareness; all claims in this article about how infants might understand or reason about agents' false beliefs should therefore be interpreted in that light.

### 1.3. Challenges

The anticipatory-looking and VOE results summarized in the previous sections suggest two conclusions: first, children in the second year of life (i.e., between about 13 and 25 months) already possess some understanding of false belief; and second, children are more likely to reveal this understanding in tasks that tap spontaneous, as opposed to elicited, responses. Here we focus on the first of these conclusions (for possible causes of children's difficulties with elicited tasks, see e.g., Bloom & German, 2000; Carlson & Moses, 2001; Carlson, Moses, & Breton, 2002; Fodor, 1992; Hala, Hug, & Henderson, 2003; Kovács, in press; Leslie, German, & Polizzi, 2005; Leslie & Polizzi, 1998; Russell, 1996; Russell, Mauthner, Sharpe, & Tidswell 1991; Scott & Baillargeon, in press).

Is the first conclusion above correct? Could weaker or more conservative interpretations be offered for the results presented in the previous sections that do not endow infants with such sophisticated psychological-reasoning abilities? To date, three such interpretations have been offered.

One alternative interpretation (Perner & Ruffman, 2005; Ruffman & Perner, 2005) is that, when shown an event in which an agent watches a goal object being hidden in a location, infants form a three-way association between the agent, the object, and its hiding location; this association then guides infants' anticipatory responses, as well as their looking behavior when events deviate from the association. If we assume that infants can form such a three-way association in a single trial, and that any new association trumps previous associations (e.g., when the agent sees the toy move to the yellow box in the belief-induction trial of the FB-yellow condition in Onishi and Baillargeon (2005), infants revise the association involving the green box formed in the three previous familiarization trials), then we can explain all of the data presented in the previous sections.

Another alternative interpretation (Onishi & Baillargeon, 2005; Perner & Ruffman, 2005; Ruffman & Perner, 2005) is that infants may bring to the laboratory a behavioral rule (learned in the course of everyday life) that agents who are searching for an object typically search for it where they last saw it. If we assume that infants not only can form such behavioral rules but also can readily generalize them to novel, non-human agents, then we can again account for all of the results described above.

A third alternative interpretation (Southgate et al., 2007; see also Garnham & Ruffman, 2001) is that infants may bring to the laboratory a general expectation that ignorance leads to error: thus, they expect an agent who is ignorant about an object's location to search for it in the wrong location. This explanation does not apply to the results of Southgate et al., since their infants correctly anticipated which box the agent would search even though the bear puppet had left with the toy so that both boxes were in fact wrong locations. However, it does apply to the results of Onishi and Baillargeon (2005) and Surian et al. (2007); perhaps the infants in the FB conditions of these experiments looked reliably longer when the agent searched for the object in its current location, not because they reasoned that the agent's false belief should have led her to search the other location, but because they reasoned that the agent was ignorant about the object's current location and thus was likely to search for it in the wrong location.

As might be expected, questions have in turn been raised about the plausibility of the alternative interpretations listed above (e.g., Csibra & Southgate, 2006; Leslie, 2005; Surian et al., 2007). For example, what evidence is there that infants can form three-way associations between agents, objects, and locations (e.g., Csibra & Southgate, 2006)? What evidence is there that infants can form behavioral rules, or can generalize such rules to non-human agents (e.g., Leslie, 2005; Southgate et al., 2007; Surian et al., 2007)? Finally, what evidence is there that infants expect ignorance to consistently lead to error (e.g., He & Baillargeon, 2007; Scott & Baillargeon, in press; Scott, Song, Baillargeon, & Leslie, 2007)?

The debate over whether infants in the second year of life are already able to attribute false beliefs to agents is likely to be heated and protracted. Our own approach to the debate has been twofold. One objective has been to test the alternative interpretations described above empirically. For example, we have asked in a series of VOE experiments whether infants respond differently when an agent is ignorant, as opposed to mistaken, about a toy's location (e.g., He & Baillargeon, 2007; Scott & Baillargeon, in press; Scott et al., 2007). Results (discussed in Section 4) indicate that infants do not in fact expect ignorant agents to err. When an agent does not know whether a toy is hidden in locationA or locationB, and infants know it is in locationA, they do not expect the agent to search locationB; rather, they have no expectation as to which location the agent will search.

Our other objective has been to explore infants' ability to reason about false beliefs more broadly, so as to better understand the nature, range, and development of this ability. The present research fits within this second objective, which is discussed more fully below.

### 1.4. A variety of belief-inducing situations

In their commentary on Onishi and Baillargeon (2005), Perner and Ruffman (2005) wrote: "The conclusions from the [standard] false-belief task are warranted only because understanding of false-belief around 4 years can be demonstrated in a variety of belief-inducing situations" (p. 216). In the same spirit, we have begun to examine whether infants can attribute to agents not only false beliefs about location, as shown in the experiments described in the previous sections, but also false perceptions as well as false beliefs about identity, contents, properties, and number (e.g., He & Baillargeon, 2007, 2008; Scott & Baillargeon, in press; Scott et al., 2007; Song & Baillargeon, in press).

To illustrate, a recent experiment (Song & Baillargeon, in press) examined 14.5-month-olds' ability to attribute false perceptions to others. The infants first watched live familiarization events in which a female agent faced two toys, a doll with blue pigtails and a stuffed skunk. Across trials, the doll and skunk were presented on placemats or inside shallow containers; the agent consistently reached for the doll, suggesting that she preferred it over the skunk. During the test trial, in the agent's absence (false-perception condition), the doll and skunk were hidden in two boxes with lids: the doll in a plain box and the skunk in a box with a tuft of blue hair protruding from under its lid. The agent then returned, reached for either the plain box (plain-box event) or the hair box (hair box event), and then paused. The infants who saw the plain-box event looked reliably longer than those who saw the hair-box event, suggesting that they expected the agent to falsely perceive the tuft of hair as belonging to the doll and hence to falsely believe that the doll was hidden in the hair box and the skunk in the plain box. In another condition (true-perception condition), the agent witnessed the hiding of the doll and skunk. The infants who saw the hair-box event now looked reliably longer than those who saw the plain-box event, suggesting that they expected the agent to search for her doll where she had seen it being hidden. These conclusions were supported by the results of another experiment in which the agent consistently reached for the skunk instead of for the doll. The infants in the false-perception condition now expected the agent to reach for the plain box, whereas those in the true-perception condition expected her to reach for the hair box.

The present research also sprang from our objective to explore infants' psychological reasoning in a variety of situations. However, rather than asking what other false beliefs infants might attribute to agents, here we examined what communications infants might view as sufficient to *correct* agents' false beliefs. We speculated that positive results would shed light on another facet of infants' ability to reason about false beliefs and as such would help constrain the interpretation of infants' performance in false-belief tasks more generally.

### 1.5. The present research

As adults, our ability to reason about others' beliefs rests on several key assumptions, which include the following. First, we recognize that beliefs are not direct reflections of reality, which must always be accurate, but representations, which may or may not be accurate. Second, we realize that, in simple situations at least, beliefs are neither arbitrary nor capricious, but are determined by causally relevant factors and hence can be inferred from contextual information. For example, if Colette hides an object in locationA, she will hold a true belief that the object is in locationA; however, if in her absence the object is transferred to locationB, she will now hold a false belief that the object is still in locationA. Finally, we understand that beliefs may be induced in others, and that false beliefs may be corrected, through appropriate interventions. To return to our example, we would not attempt to correct Colette's false belief that the object is still in locationA by clapping our hands or playing the tuba. Only appropriate interventions, such as showing Colette that the object is now in locationB, or communicating this information through words or gestures, would result in her holding a correct belief about the object's location.

The evidence reviewed in the previous sections suggests that infants in the second year of life already possess something akin to the first two assumptions above: they realize that beliefs may be true or false, and that whether an agent holds a true or a false belief about a scene depends on causally relevant factors. For example, infants expect an agent who hides a toy in locationA and then witnesses its displacement to locationB to hold a true belief that the toy is now in locationB; in contrast, infants expect an agent who does not witness this displacement to hold a false belief that the toy is still in locationA. Infants thus possess some understanding of the conditions that may lead an agent to hold a true or a false belief about a toy's location (Onishi & Baillargeon, 2005; Song & Baillargeon, in press; Southgate et al., 2007; Surian et al., 2007).

If infants realize that beliefs are determined by causally relevant factors, they may expect false beliefs to be corrected by appropriate, though not inappropriate, communications. The present experiments examined this issue and asked whether 18-month-old infants would expect an agent's false belief about a toy's location to be corrected when another agent communicated information about the toy's new location through words (Experiment 1) or pointing gestures (Experiment 2).

## 2. Experiment 1

The infants in Experiment 1 were assigned to an informative- or an uninformative-intervention condition. The infants in both conditions first received *familiarization* trials in which one agent (agent1) faced two lidded containers: a blue box on the left and a red cup on the right (from the infants' perspective). While agent2 looked on, agent1 played with a ball and then hid it in the box (see Fig. 1). The infants next received a *belief-induction* trial in which agent1 was absent and agent2 moved the ball to the cup (see Fig. 2). At this point, agent1 could be said to possess a false belief that the ball was still in the box.

Next, the infants received an *intervention* trial that differed in the two conditions (see Fig. 3). In the informative-intervention condition, the infants witnessed a communication that could correct agent1's false belief about the toy's location: specifically, agent2 told agent1 "The ball is in the cup!" twice. In the uninformative-intervention condition, the infants witnessed a communication that should not correct agent1's false belief: agent2 simply told agent1 "I like the cup!" twice. Thus, whereas agent2's utterance in the informative-intervention condition conveyed information about the ball's current location (the

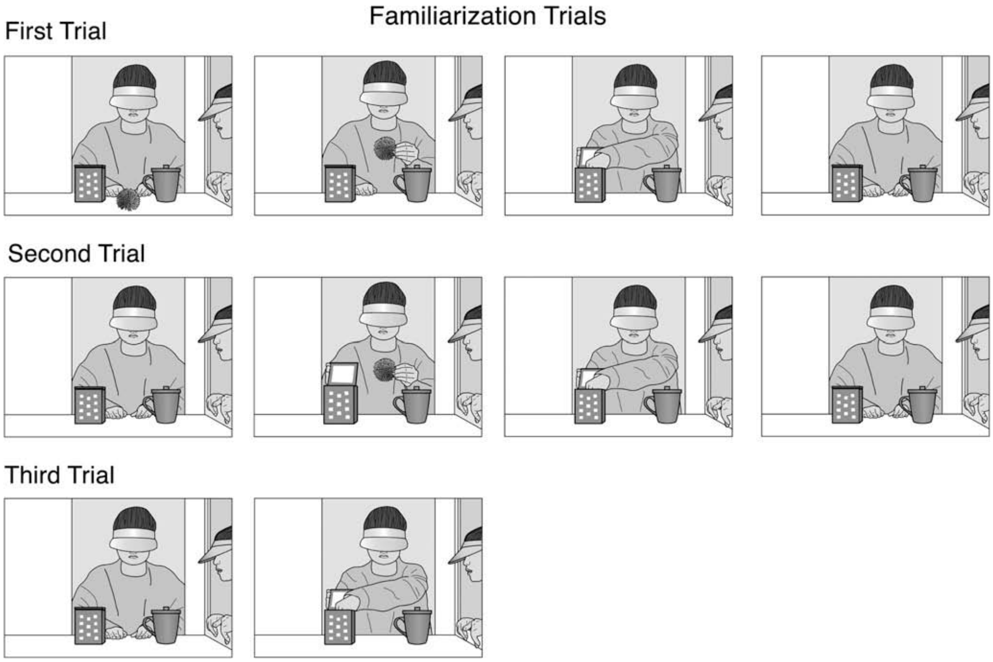


**Fig. 1.** Schematic drawing of the events shown during the three familiarization trials in Experiment 1.

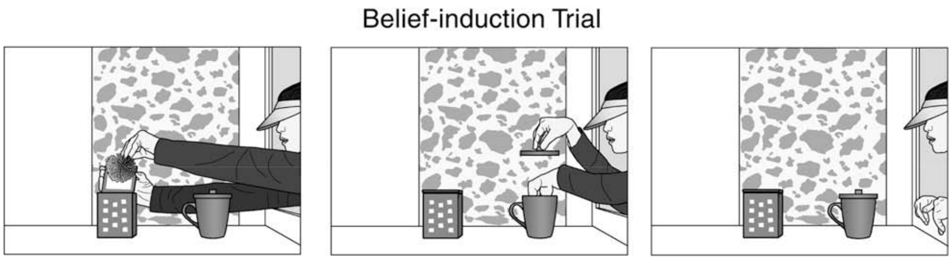


**Fig. 2.** Schematic drawing of the event shown during the belief-induction trial in Experiment 1.

cup), agent2's utterance in the uninformative-intervention condition provided no such information. Previous research suggested that 18-month-old infants would understand and distinguish between these two simple communications (e.g., Hirsh-Pasek & Golinkoff, 1996; Meints, Plunkett, Harris, & Dimmock, 2002).

Finally, the infants in both conditions received a single *test* trial (see Fig. 4) in which agent2 was absent and agent1 reached either for the box (box event) or the cup (cup event).

Our predictions were as follows. If the infants in the informative-intervention condition (1) realized that agent1 did not witness the ball's displacement during the belief-induction trial and hence held the false belief that it was still in the box, and (2) understood that this false belief *could* be corrected by agent2's relevant communication during the intervention trial ("The ball is in the cup!"), then they should expect agent1 to reach for the cup and they should be surprised when she reached for the box instead. The infants who saw the box event should thus look reliably longer than those who saw the cup event.

Conversely, if the infants in the uninformative-intervention condition (1) again realized that agent1 did not witness the ball's displacement during the belief-induction trial and hence held the false belief that it was still in the box, and (2) understood that this false belief could *not* be corrected by agent2's irrelevant communication during the intervention trial ("I like the cup!"), then they should expect agent1 to retain her false belief and to reach for the box, and they should be surprised when she reached for the cup instead. The infants who saw the cup event should thus look reliably longer than those who saw the box event.

Opposite looking patterns were thus predicted for the infants in the informative- and uninformative-intervention conditions. Of course, other patterns were also possible. For example, the infants in the two conditions might realize, following the belief-induction trial, that agent1 now falsely believed that the ball was still hidden in the box, but be unable to reason about how this belief might be corrected. In such a case, the infants in both conditions should expect agent1 to reach for the box, where she falsely

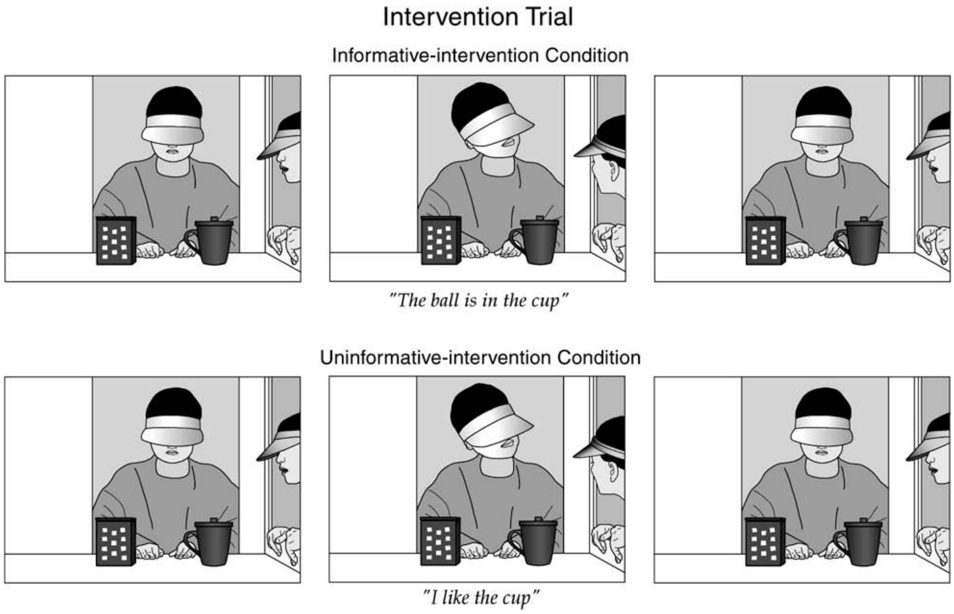


**Fig. 3.** Schematic drawing of the events shown during the intervention trial in the informative- and uninformative-intervention conditions of Experiment 1.

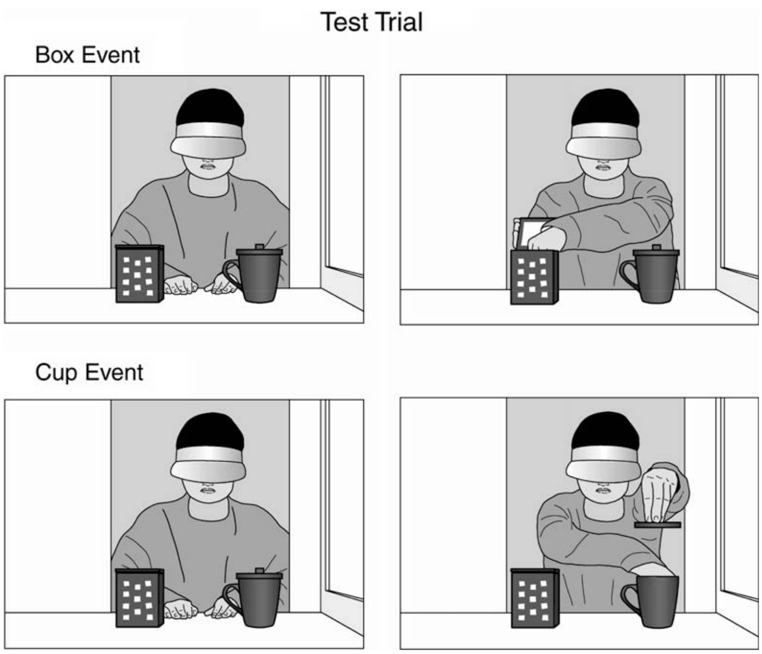


**Fig. 4.** Schematic drawing of the box and cup test events shown in Experiment 1.

believed the ball to be hidden. Thus, in both conditions, the infants who saw the cup event should look reliably longer than those who saw the box event. Alternatively, the infants in the two conditions might understand that agent1's false belief could be corrected but have difficulty distinguishing between agent2's informative and uninformative communications. For example, upon hearing the words "the cup" in either communication, the infants might conclude that agent1 now knew the ball's correct location. In such a case, the infants in the two conditions should expect agent1 to reach for the cup. Thus, in both conditions, the infants who saw the box event should look reliably longer than those who saw the cup event.

### 2.1. Method

#### 2.1.1. Participants

Participants were 28 healthy term infants, 14 male and 14 female, from English-speaking families ($M$ = 18 months, 18 days; range: 18 months, 3 days to 19 months, 9 days).

Another 9 infants were tested but not included in the analyses because they were inattentive (3), drowsy (1), or fussy (1), because they looked more than 3 SD from the mean of their condition in the test trial (2) or looked for the maximum amount of time allowed on the belief-induction trial (1), or because of parental interference (1). Half the infants were randomly assigned to the informative-intervention condition, and half to the uninformative-intervention condition; within each condition, half the infants saw the box event, and half saw the cup event.

The infants' names in this and in the following experiment were obtained primarily from purchased mailing lists and from birth announcements in the local newspaper. Parents were offered reimbursement for their transportation expenses but were not otherwise compensated for their participation.

*2.1.2. Apparatus*

The apparatus consisted of a wooden display booth 126 cm high, 102 cm wide, and 58 cm deep, mounted 76 cm above the room floor. The infant faced an opening 50 cm high and 95 cm wide in the front of the apparatus; between trials, a curtain consisting of a muslin-covered frame 60 cm high and 101 cm wide was lowered in front of this opening. The side walls of the apparatus were painted white, the back wall was made of white foam board, and the floor was covered with beige cardboard.

Agent1 sat on a wooden chair centered behind a window in the back wall of the apparatus. This back window was 63.5 cm high and 41 cm wide; it extended from the apparatus floor and was located 7 cm from the right wall. During the belief-induction trial, the back window was filled with a panel covered with a salient black and white pattern; it was hoped that this pattern would help infants notice that agent1 was not present in the trial. Agent1 wore a blue shirt and an ivory visor that hid her eyes from the infants; a muslin curtain behind agent1 hid the test room.

Agent2 knelt on pillows behind a window in the right wall of the apparatus. This side window was 51 cm high and 38 cm wide, and was located 4 cm above the apparatus floor and 7.5 cm from the back wall. During the test trial, the side window was filled with a muslin curtain. Agent2 wore a dark green shirt and a light green visor.

The box and cup stood on the apparatus floor 15 cm in front of the back window, 4.5 cm from the window's left and right edge, respectively; the box and cup were 11 cm apart at their closest point. The box was made of foam board, covered with blue contact paper, and decorated with small white squares; it was 10.5 cm tall, 8.5 cm wide, and 8.5 cm deep. The box's lid was also made of foam board and covered with blue contact paper, and was 8.5 cm square and 0.5 cm thick; it was hinged to the back of the box with blue tape. The cup was made of red porcelain and had a large handle oriented to the left; it was 10.5 cm tall, 9 cm in diameter at the top, and 12.5 cm wide at its widest point (with its handle included). The cup's lid was made of foam board, was covered with red contact paper, and was 1 cm tall and 9 cm in diameter; attached to the center of the lid was a small rectangular red knob 1 cm high, 1.5 cm wide, and 1.5 cm deep. Centered between and 3 cm in front of the box and cup was a green plastic koosh ball 6 cm in diameter.

The infants were tested in a brightly lit room. Three 20-W fluorescent light bulbs attached to the front and back walls of the apparatus provided additional light. Two frames, each 184 cm high and 76 cm wide and covered with blue cloth, stood at an angle on either side of the apparatus; these frames served to isolate the infants from the test room.

*2.1.3. Trials*

Each trial consisted of an *initial* and a *final* phase; looking times during the two phases were computed separately (for reasons explained below, the intervention trial had only an initial phase). During the initial phase of a trial, the agents performed the scripted actions appropriate for the trial and then paused; during the final phase, the infant watched this paused scene until the trial ended. The duration of the initial phase was fixed and depended on the specific actions performed. The duration of the final phase was infant-controlled (see below for the specific criteria used to end trials). When a trial ended, a supervisor lowered the curtain at the front of the apparatus, and stimuli were readied for the next trial. Inter-trial intervals lasted about 10 s, and each new trial began with the raising of the curtain.

In the following descriptions, the numbers in parentheses indicate the number of seconds taken to perform the actions described. To help the agents adhere to the events' scripts, a metronome beat softly once per second. A camera mounted behind and next to the infant projected an image of the events onto a TV screen located behind the apparatus in the test room; the supervisor monitored the events to confirm that they followed the prescribed scripts.

*2.1.3.1. Informative-intervention condition.*

*First familiarization trial.* At the start of the first familiarization trial, agent1 sat at the back window, facing forward, with her bare hands resting on the floor 9 cm behind and centered between the box and cup. Agent2 knelt at the side window, facing forward, with her left profile to the infants; her hands loosely grasped the lower edge of the window. Both agents looked at a neutral point between the two containers. During the trial, agent1 looked at the ball and box as she acted on them; agent2 looked at the actions performed by agent1. Because the agents wore visors, their eyes were generally not visible to the infants; nevertheless, the infants could gather some information about the direction of each agent's gaze from her head orientation.

The initial phase of the first familiarization trial lasted 15 s. After a pause (1 s), agent1's left hand grasped the ball (1 s), moved it to a position about 20 cm above the apparatus floor, between the two containers (1 s), and rotated it to the left and right twice, changing orientation once per second (4 s). Agent1's right hand then grasped the box's lid (1 s) and opened it (1 s). Next, agent1's left hand placed the ball inside the box (2 s), and then withdrew to its starting position on the apparatus floor (1 s). Finally, agent1's right hand closed the box's lid (2 s) and withdrew to the apparatus floor (1 s). The two agents then paused in their respective positions. During the final phase, the infants watched this paused scene until the trial ended.

*Second familiarization trial.* The set-up at the start of the second familiarization trial was similar to that of the first trial, except that the ball was inside the box. The initial phase again lasted 15 s. After a pause (1 s), agent1's right hand grasped the box's lid (1 s) and opened it (1 s). Agent1's left hand then reached into the box (1 s), pulled out the ball and moved it to the position above and between the two containers (1 s), and rotated it to the left and right twice (4 s). Next, agent1's left hand lowered the ball back inside the box (2 s) and withdrew to the apparatus floor (1 s). Finally, agent1's right hand closed the box's lid (2 s), and withdrew to the apparatus floor (1 s). The two agents then paused, and the infants watched this paused scene until the trial ended.

*Third familiarization trial.* The set-up at the start of the third familiarization trial was identical to that of the second trial. The initial phase lasted 3 s. To start, agent1's right hand grasped the box's lid (1 s) and opened it (1 s). Next, agent1's left hand reached into the box (1 s). The two agents then paused, and the infants watched this paused scene until the trial ended.

*Belief-induction trial.* The set-up at the start of the belief-induction trial was similar to that of the third familiarization trial, with two exceptions: agent1 was absent and the back window was filled with the patterned panel. The initial phase lasted 12 s. Agent2's right hand grasped the box's lid (1 s) and opened it (1 s). Agent2's left hand then reached inside the box (1 s), pulled out the ball, and moved it toward herself, to a position about 5 cm above the apparatus floor, 9 cm from the cup, and 3 cm from the side window (1 s). Agent2's right hand then closed the box (1 s), grasped the cup's lid (1 s), and lifted it about 9 cm above the cup (1 s). Next, agent2's left hand placed the ball inside the cup (2 s), and withdrew to its starting position at the lower edge of the side window (1 s). Finally, agent2's right hand replaced the cup's lid (1 s) and withdrew to the lower edge of the side window (1 s). Agent2 then paused, and the infants watched this paused scene until the trial ended.

*Intervention trial.* The set-up at the start of the intervention trial was similar to that of the second and third familiarization trials. The intervention trial only had an initial phase, which lasted 9 s. After a pause (1 s), the two agents looked at each other (1 s). Agent2 then said "The ball is in the cup!" (2 s) twice, with a pause (1 s) after each utterance. Agent2 was a female native speaker of English, and spoke in an infant-directed manner at a listening level of about 68 dB (measured with a sound-level meter placed at the infants' location). Next, both agents turned their heads back to look at the neutral position between the two containers (1 s), and then the trial ended.

No final phase was included in the intervention trial because we were concerned that the infants might become puzzled or confused as time passed and agent1 failed to reach for one of the containers, suggesting perhaps that she had lost interest in the ball or had failed to understand agent2's communication.

*Test trial.* The set-up at the start of the test trial was similar to that of the intervention trial except that only agent1 was present; agent2's window was filled with a muslin curtain. During the trial, the infants saw either a box or a cup event; the initial phase of either event lasted 3 s. In the *box* event, agent1's right hand grasped the box's lid (1 s) and opened it (1 s); her left hand then reached inside the box (1 s). Agent1 then paused, with her right hand holding the lid and her left hand inside the box, and the infants watched this paused scene until the trial ended. In the *cup* event, agent1's left hand grasped the cup's lid (1 s) and lifted it about 9 cm (1 s); her right hand then reached inside the cup (1 s). Agent1 then paused, as before, until the trial ended.

*2.1.3.2. Uninformative-intervention condition.* The infants in the uninformative-intervention condition received the same trials as those in the informative-intervention condition, with one exception: during the intervention trial, agent2 said "I like the cup!" twice.

#### *2.1.4. Procedure*

During the experiment, the infant sat on a parent's lap in front of the apparatus; the infant's head was approximately 50 cm from the curtain. Parents were instructed to close their eyes and to remain silent and neutral during the entire experiment.

The infant's looking behavior was monitored by two observers who viewed the infant through peepholes in the cloth-covered frames on either side of the apparatus. Each observer held a button linked to a computer and depressed the button when the infant looked at the event. Looking times during the initial and final phases of each trial were computed separately. During the familiarization and test trials, the looking times recorded by the primary observer were used to determine when the final phase of a trial had ended (see below). During the belief-induction and intervention trials, the primary observer typically left the room and the secondary observer served as the sole observer. This procedural safeguard ensured that the primary observer was naive during the test trial as to which condition and experiment the infant was assigned to.

All infants first received the three *familiarization* trials described above. Examination of the infants' looking times during the initial phase of each trial revealed that they were highly attentive: they looked on average for 14.9/15 s during the first trial, 14.5/15 s during the second trial, and 2.5/3 s during the third trial. The final phase of each trial ended when the infant (1) looked away for 2 consecutive seconds after having looked for at least 2 cumulative seconds, or (2) looked for 60 cumulative seconds without looking away for 2 consecutive seconds.

Next, all infants received the *belief-induction* trial described above. The infants were very attentive during the initial phase and looked for 11.9/12 s on average. The final phase ended when the infant (1) looked away for 2 consecutive seconds after having looked for at least 3 cumulative seconds, or (2) looked for 40 cumulative seconds without looking away for 2 consecutive seconds.

Following the belief-induction trial, the infants received the *intervention* trial appropriate for their condition. As before, the infants were highly attentive during the initial phase and looked for 8.6/9 s in both conditions (recall that this trial had no final phase).

Finally, each infant received a single *test* trial; half the infants in each intervention condition saw the box event, and half saw the cup event. The infants were attentive during the initial phase and looked for 2.8/3 s on average. The final phase ended when the infant (1) looked away for 2 consecutive seconds after having looked for at least 5 cumulative seconds, or (2) looked for 60 cumulative seconds without looking away for 2 consecutive seconds.

To assess interobserver agreement during the familiarization and test trials (recall that the primary observer was typically absent during the belief-induction and intervention trials), the final phase of each trial was divided into 100-ms intervals, and the computer determined in each interval whether the two observers agreed that the infant was or was not looking at the event. Percent agreement was calculated for each trial by dividing the number of intervals in which the observers agreed by the total number of intervals in the trial. Agreement was measured for 24 infants (only one observer was present for the other 4 infants) and averaged 94% per trial per infant.

Preliminary analyses of the test data revealed no significant main effect of sex and no significant interactions involving sex, all $Fs(1, 20) < 1.35$, $ps > .26$; the data were therefore collapsed across sex in subsequent analyses.

### *2.2. Results*

The infants' looking times during the final phases of the three *familiarization* trials were averaged and analyzed by means of a 2 × 2 analysis of variance (ANOVA) with condition (informative- or uninformative-intervention) and test event (box or cup) as between-subjects factors. No effect was significant, all $Fs(1, 24) < 1.04$, $ps > .32$, suggesting that the infants in the four experimental groups tended to look equally during these trials (informative-intervention/box-event, $M = 23.0$, $SD = 13.4$; informative-intervention/cup-event, $M = 29.4$, $SD = 16.4$; uninformative-intervention/box-event, $M = 28.5$, $SD = 12.4$; uninformative-intervention/cup-event, $M = 32.8$, $SD = 13.2$). A similar analysis of the infants' looking times during the final phase of the *belief-induction* trial also produced no significant effects, all $Fs(1, 24) < 1.08$, $ps > .31$, suggesting that the infants looked about equally during this trial (informative-intervention/box-event, $M = 17.5$, $SD = 7.7$; informative-intervention/cup-event, $M = 13.9$, $SD = 11.7$; uninformative-intervention/box-event, $M = 16.7$, $SD = 11.1$; uninformative-intervention/cup-event, $M = 12.5$, $SD = 9.1$).

The infants' looking times during the final phase of the *test* trial (see Fig. 5) were analyzed as above. The main effect of condition was marginally significant, $F(1, 24) = 4.15$, $p < .06$, and the main effect of test event was not significant, $F(1, 24) < 1$. However, the interaction between condition and test event was significant, $F(1, 24) = 15.57$, $p < .001$. Planned comparisons indicated that (1) in the informative-intervention condition, the infants who saw the box event ($M = 22.1$, $SD = 4.8$) looked reliably longer than those who saw the cup event ($M = 13.7$, $SD = 5.8$), $F(1, 24) = 5.12$, $p < .05$, and (2) in the uninformative-intervention condition, the infants who saw the cup event

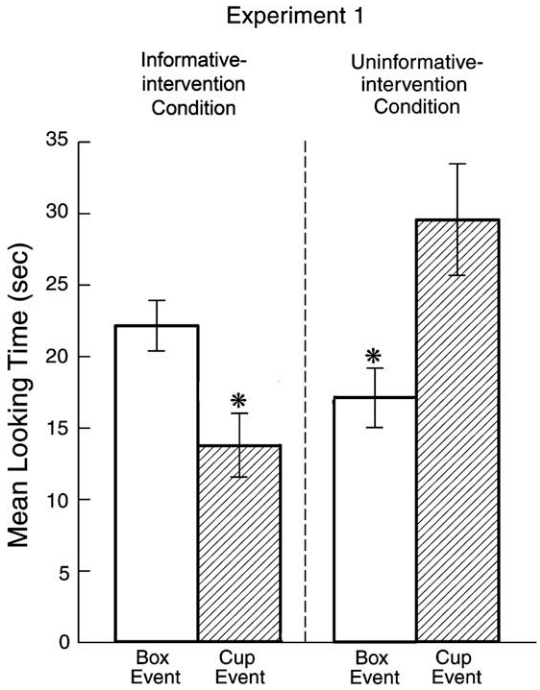


**Fig. 5.** Mean looking times to the box and cup test events for the infants in the informative- and uninformative-intervention conditions of Experiment 1. Error bars represent standard errors for each condition.

($M = 29.5$, $SD = 10.2$) looked reliably longer than those who saw the box event ($M = 17.1$, $SD = 5.6$), $F(1, 24) = 11.00$, $p < .005$.[1]

Non-parametric Wilcoxon rank-sum tests confirmed the results of the informative- ($W_S = 34$, $p < .025$) and uninformative-intervention ($W_S = 34$, $p < .025$) conditions.

### *2.3. Discussion*

In the informative-intervention condition, the infants who saw the box event looked reliably longer than those who saw the cup event; in the uninformative-intervention condition, the reverse pattern was found. These results suggest that the infants in the informative-intervention condition (1) realized following the belief-induction trial that agent1 now held the false belief that the ball was still in the box;[2] (2) assumed that agent2's communication in

[1] The test data were also subjected to an analysis of covariance (ANCOVA), using as covariates the infants' average looking times during the familiarization trials and their looking times during the belief-induction trial. The results of the ANCOVA replicated those of the ANOVA: the interaction between condition and test event was again significant, $F(1, 22) = 15.13$, $p < .001$. Planned comparisons confirmed that in the informative-intervention condition, the infants who saw the box event looked reliably longer than those who saw the cup event, $F(1, 22) = 5.38$, $p < .05$, whereas in the uninformative-intervention condition, those who saw the cup event looked reliably longer than those who saw the box event, $F(1, 22) = 10.16$, $p < .005$.

[2] The present data are insufficient to determine whether the infants (1) realized during the belief-induction trial that agent1 now held a false belief about the ball's location, or (2) came to this realization only when they saw agent1 again at the start of the intervention trial. The issue of whether infants keep track of the beliefs of absent agents is interesting in its own right and worthy of investigation, but does not bear on the main conclusions drawn here.

the intervention trial ("The ball is in the cup!") was sufficient to correct this false belief; and hence (3) expected agent1 to reach for the cup in the test trial and were surprised when she reached for the box instead.

Conversely, the infants in the uninformative-intervention condition (1) again realized following the belief-induction trial that agent1 now held the false belief that the ball was still in the box; (2) assumed that agent2's communication during the intervention trial ("I like the cup!") was not sufficient to correct this false belief; and hence (3) expected agent1 to reach for the box in the test trial and were surprised when she reached for the cup instead.

The present results thus add to the recent evidence from VOE experiments (Onishi & Baillargeon, 2005; Song & Baillargeon, in press; Surian et al., 2007) that infants in the second year of life can attribute false beliefs to agents. The present results also extend these findings to older, 18-month-old infants, and to situations more similar to those used in standard and anticipatory-looking false-belief tasks, in which one agent hides an object in one location and, after the agent leaves, another agent moves it to a different location (e.g., Baron-Cohen et al., 1985; Clements & Perner, 1994; Garnham & Ruffman, 2001; Wimmer & Perner, 1983). Furthermore, the present results suggest that, by 18 months of age, infants expect false beliefs to be corrected by appropriate interventions, and they show some ability to distinguish between informative and uninformative communications.

In future research, we plan to build on the present findings to ascertain just how precise are infants' expectations about what might constitute informative communications. The results of the uninformative-intervention condition suggest that the infants deemed agent2's mention of the cup ("I like the cup!") insufficient to inform agent1 about the ball's current location. Would infants take any mention of the ball and the cup in the same utterance (e.g., "The ball and the cup!", "I like the ball and the cup!", or "The ball is acorp the cup!") to be sufficient to amend agent1's false belief about the ball's location? Or would infants expect agent1's false belief to be corrected only by communications that make explicit (or appear to make explicit) the ball's new location (in the cup)? Would age effects emerge in infants' responses to these various communications? Future experiments will explore these and related questions.[3]

[3] Although our discussion has focused on what infants might construe as informative communications, questions could also be raised about uninformative communications. For example, we have assumed that the infants in the uninformative-intervention condition understood agent2's utterance ("I like the cup!") and deemed it insufficient to correct agent1's false belief about the ball's location. However, there are other possibilities: for example, it might be suggested that the infants interpreted agent2's utterance less literally, either as a signal that she planned to act on the cup or as warning to agent1 to stay away from the cup. These interpretations leave open the possibility that infants might take a reference to the cup in a more neutral utterance as sufficient to amend agent1's false belief about the ball's location. A test of this possibility would be to run a modified uninformative-intervention condition in which agent2 utters a more neutral statement such as "Here is the cup!" or "Look at the cup!". Evidence that infants again expect agent1 to reach for the box would support our claim that the infants in the uninformative-intervention condition of Experiment 1 took agent2's utterance at face value (i.e., as a statement that she liked the cup), and realized that this utterance was insufficient to inform agent1 about the ball's current location.

## 3. Experiment 2

In Experiment 2, we sought further evidence that 18-month-old infants understand that an agent's false belief about an object's location can be corrected. The infants were assigned to an intervention or a no-intervention condition, and the procedure used in the intervention condition was similar to that in Experiment 1, with two exceptions: agent2 used a pointing gesture, rather than words, to convey information about the ball's new location, and the red cup was replaced with a red can.

There were two main reasons for using this pointing intervention. One was to determine whether infants would interpret agent2's pointing gesture as providing agent1 with information about the ball's location. In Experiment 1, agent2's utterance "The ball is in the cup" conveyed clear and explicit information about the ball's new location. In Experiment 2, agent2's gesture of pointing at the can was less specific, in that there were other possible reasons for agent2 to make this gesture: for example, to indicate that she liked the can and wanted to draw attention to it (e.g., Woodward & Guajardo, 2002). Thus, we wanted to find out how 18-month-old infants, after seeing agent2 move the ball from the box to the can in agent1's absence, would interpret agent2's pointing at the can upon agent1's return: as an informative communication that could correct agent1's false belief about the ball's location, or as an uninformative communication insufficient to amend this false belief.

Another reason for using a pointing intervention in Experiment 2 concerned future extensions of the present procedure to younger infants. Previous research indicates that infants aged 12 months and older not only produce points, but also understand their communicative function (e.g., Butterworth & Grover, 1988; Carpenter, Nagell, & Tomasello, 1998; Franco & Butterworth, 1996; Lempers, 1979; Leung & Rheingold, 1981; Liszkowski, Carpenter, Henning, Striano, & Tomasello, 2004; Liszkowski, Carpenter, Striano, & Tomasello, 2006; Woodward & Guajardo, 2002). For instance, Liszkowski et al. (2004) found that 12-month-olds were more likely to repeat their pointing gesture when an adult was not fixating the event of infants' interest, presumably to draw the adult's attention to the event and establish joint attention. Thus, part of the motivation for Experiment 2 was the consideration that a pointing manipulation, if successful, could perhaps be adapted in future experiments to test infants younger than 18 months, who might not understand sentences such as "The ball is in the cup!", but might still understand points.

### 3.1. Method

The infants in the intervention and no-intervention conditions first received three *familiarization* trials (see Fig. 6) similar to those in Experiment 1: while agent2 looked on, agent1 played with the ball and hid it in the box. Next, the infants received one *belief-induction* trial (see Fig. 7) similar to that in Experiment 1: while agent1 was absent, agent2 moved the ball from the box to the can. At this point, agent1 could be said to hold a false belief that the ball was still in the box.

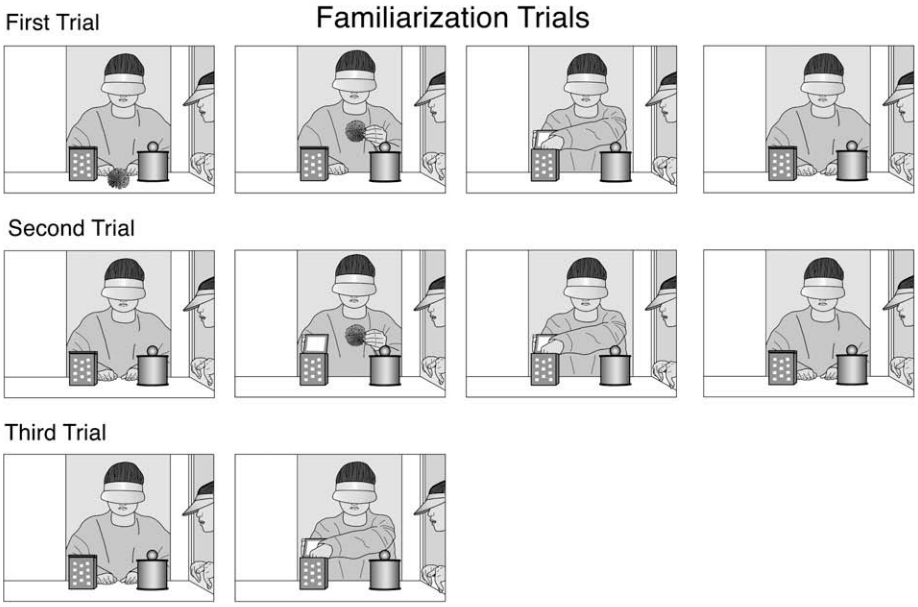


**Fig. 6.** Schematic drawing of the events shown during the three familiarization trials in Experiment 2.

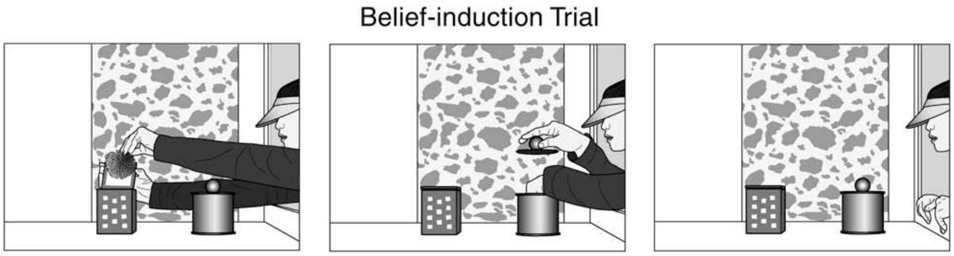


**Fig. 7.** Schematic drawing of the event shown during the belief-induction trial in Experiment 2.

Next, the infants in the intervention condition received one *intervention* trial (see Fig. 8) during which agent2 pointed to the can twice. A possible interpretation of agent2's points was that they communicated information to agent1 about the ball's new location, since agent1 was absent when agent2 moved the ball to the can during the previous, belief-induction trial. The infants in the no-intervention condition did not receive an intervention trial.

Finally, as in Experiment 1, all of the infants received a single *test* trial (see Fig. 9) in which agent2 was absent and agent1 reached for the box (box event) or can (can event).

We predicted that the infants in the intervention and no-intervention conditions, like those in Experiment 1, would realize following the belief-induction trial that agent1 now held the false belief that the ball was still in the box. In the no-intervention condition, the infants should expect agent1 to act on this belief, since no intervention came to correct it: they should thus expect agent1 to reach into the box, and they should be surprised when she reached into the can instead. The infants who saw the can event should thus look reliably longer than those who saw the box event.

As for the infants in the intervention condition, at least two alternatives were possible. On one hand, if the infants determined that agent2's pointing gesture was sufficient to inform agent1 about the ball's new location, then they should conclude that agent1 now held the true belief that the ball was in the can. As a result, they should expect agent1 to reach into the can, and they should be surprised when she reached into the box instead. The infants who saw the box event should thus look reliably longer than those who saw the can event. On the other hand, if the infants perceived agent2's pointing gesture to be ambiguous and hence uninformative, then they should expect agent1 to retain her false belief that the ball was in the box, and they should be surprised when she reached into the can. Like the infants in the no-intervention condition, the infants who saw the can event should thus look reliably longer than those who saw the box event. Different looking patterns were thus predicted in the intervention condition depending on whether the infants viewed agent2's pointing as informative or as uninformative.

#### 3.1.1. Participants

Participants were 32 healthy term infants, 16 male and 16 female ($M$ = 18 months, 17 days; range: 18 months, 7 days to 19 months, 6 days). Another 13 infants were tested but not included in the analyses because they were


**Fig. 8.** Schematic drawing of the event shown during the intervention trial in the intervention condition of Experiment 2.

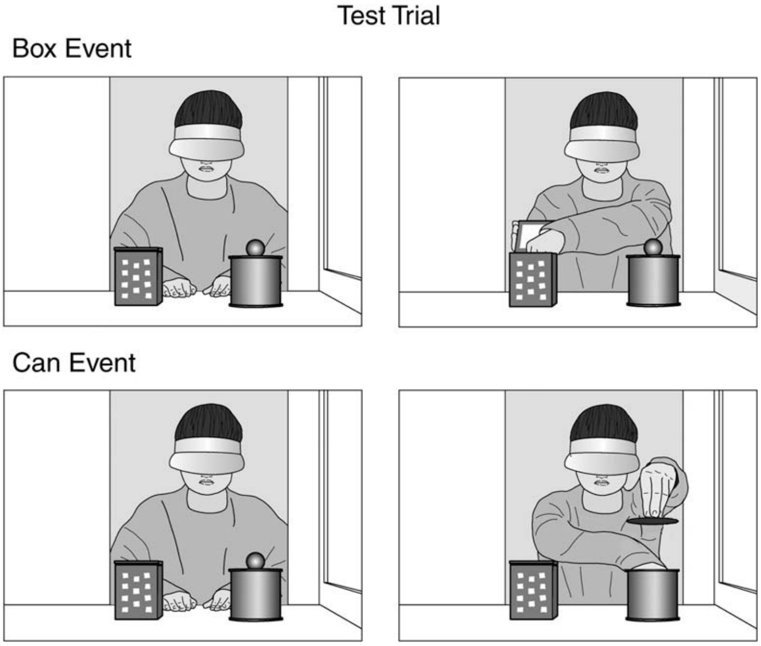


**Fig. 9.** Schematic drawing of the box and can test events shown in Experiment 2.

inattentive (3), fussy (1), or distracted (1), looked for the maximum amount of time allowed in the belief-induction trial (2) or test trial (3), or looked more than 3 SD from the mean of their condition in the test trial (3). Half the infants were randomly assigned to the intervention condition, and half to the no-intervention condition; within each condition, half the infants saw the box event, and half saw the can event.

### *3.1.2. Apparatus*

The apparatus and stimuli used in Experiment 2 were identical to those in Experiment 1, except that the red cup was replaced with a red can. This can was 9.5 cm high, 10 cm in diameter, and covered with red contact paper; its lid was 0.25 cm thick, 10 cm in diameter, and also covered with red contact paper. Attached to the center of the lid was a round knob 3.5 cm in diameter, made of rubber, and decorated with multi-colored stripes.

### *3.1.3. Trials*

The infants in Experiment 2 received trials similar to those in Experiment 1, with a few exceptions. First, the cup was replaced with a can. In Experiment 1, using a cup was helpful because most 18-month-olds know the word "cup", so the infants could understand statements such as "The ball is in the cup!". However, we were concerned that pointing to the cup in Experiment 2 might suggest irrelevant, cup-related interpretations (e.g., that agent2 wanted to drink from the cup). To avoid such interpretations, we replaced the cup with another generic container, a can. Second, the infants in no-intervention condition did not receive an intervention trial. Third, the infants in the intervention condition received a different intervention trial than in Experiment 1. The set-up at the start of the trial was the same as in Experiment 1, and the trial again consisted of a 9-s initial phase. After a pause (1 s), the two agents looked at each other (1 s). Next, agent2 turned to the can and pointed to it with her right index finger, moving her finger up and down three times (1 s); she then paused (1 s), pointed at the can again (1 s), and withdrew her hand to its original position at the bottom of the window (1 s). Finally, both agents turned their heads back to look at the neutral position between the two containers (1 s), and paused (2 s). The trial then ended. As in Experiment 1, no final phase was included because we were concerned that the infants might become puzzled as time passed and agent1 failed to reach for one of the containers.

### 3.1.4. Procedure

As in Experiment 1, the infants in Experiment 2 received three familiarization trials, one belief-induction trial, one intervention trial (intervention condition only), and one test trial, in which they saw the box or can event. Examination of the infants' looking times during the initial phase of each trial revealed that they were highly attentive: the infants looked for 14.5/15 s during the first familiarization trial, 14.7/15 s during the second familiarization trial, 2.8/3 s during the third familiarization trial, 11.9/12 s during the belief-induction trial, 8.9/9 s during the intervention trial (intervention condition only), and 3/3 s during the test trial. The criteria used to end the trials were the same as in Experiment 1. As before, the primary observer was typically absent during the belief-induction and intervention trials. Interobserver agreement during the final phases of the familiarization and test trials was measured for 31 of the 32 infants and averaged 95% per trial per infant.

Preliminary analyses of the test data revealed no significant main effect of sex and no significant interactions involving sex, all $Fs(1, 24) < 1.60$, $ps > .21$; the data were therefore collapsed across sex in subsequent analyses.

### 3.2. Results

The infants' looking times during the final phases of the three *familiarization* trials were averaged and analyzed by means of a $2 \times 2$ ANOVA with condition (intervention or no-intervention) and test event (box or can) as between-subjects factors. No effect was significant, all $Fs(1, 28) < 1.27$, $ps > .27$, suggesting that the infants in the four experimental groups tended to look equally during the familiarization trials (intervention/box-event, $M = 27.1$, $SD = 15.8$; intervention/can-event, $M = 25.6$, $SD = 8.7$; no-intervention/box-event, $M = 21.1$, $SD = 12.1$; no-intervention/can-event: $M = 22.5$, $SD = 7.0$). A similar analysis of the infants' looking times during the final phase of the *belief-induction* trial again revealed no significant effects, all $Fs(1, 28) < 1.03$, $ps > .32$, suggesting that the infants looked about equally during the belief-induction trial (intervention/box-event, $M = 14.9$, $SD = 10.9$; intervention/can-event, $M = 18.4$, $SD = 10.8$; no-intervention/box-event, $M = 12.3$, $SD = 8.0$; no-intervention/can-event, $M = 15.9$, $SD = 9.0$).

The infants' looking times during the final phase of the *test* trial (see Fig. 10) were analyzed as above. Neither the main effect of condition nor that of test event was significant, both $Fs(1, 28) < 1$. However, the interaction between condition and test event was significant, $F(1, 28) = 18.13$, $p < .00025$. Planned comparisons indicated that (1) in the intervention condition, the infants who saw the box event ($M = 26.7$, $SD = 7.0$) looked reliably longer than those who saw the can event ($M = 16.9$, $SD = 5.7$), $F(1, 28) = 5.86$, $p < .025$, and (2) in the no-intervention condition, the infants who saw the can event ($M = 29.1$, $SD = 12.7$) looked reliably longer than those who saw the box event ($M = 14.6$, $SD = 4.0$), $F(1, 28) = 12.97$, $p < .0025$.[4]

Non-parametric Wilcoxon rank-sum tests confirmed the results of the intervention ($W_S = 41$, $p < .01$) and no-intervention ($W_S = 44$, $p < .025$) conditions.

### 3.3. Further results

The results of the intervention condition in Experiment 2 suggested that the infants viewed agent2's pointing at the can as an intervention sufficient to correct agent1's false belief about the ball's location: it informed her that the ball was now in the can. However, another, less interesting interpretation of these results was that the infants simply expected agent1 to reach for whichever container agent2 pointed to, without considering agent1's belief about the ball's location. The infants might have assumed, for example, that the agents were now playing a game in which agent2 pointed to the container agent1 should reach for next.

To examine this alternative interpretation, another group of 18-month-old infants was tested in a control condition identical to the intervention condition, except that the belief-induction trial was different (see Fig. 11): while agent1 looked on, agent2 took the ball out of the box, played with it, and then replaced it inside the box. Thus, following the belief-induction trial, both agent1 and agent2 held the true belief that the ball was in the box.

We reasoned that if the infants in the intervention condition expected agent1 to reach for the can in the test trial simply because agent2 had pointed to it in the preceding, intervention trial, then the infants in the control condition, who saw the same intervention trial, should also expect agent1 to reach for the can, and they should be surprised when she reached for the box instead. Thus, the infants who saw the box event should again look reliably longer than those who saw the can event. On the other hand, if the infants in the intervention condition expected agent1 to reach for the can in the test trial because (1) agent1 was not present during the belief-induction trial when agent2 moved the ball from the box to the can, and (2) agent2's pointing gesture served to inform agent1 about the ball's new location, then the infants in the control condition should respond differently. In this condition, agent1 and agent2 were both present when agent2 hid the ball in the box. Thus, the infants should reason that, whatever purpose lay behind agent2's pointing at the can, this gesture did not convey information about the ball's location and hence was irrelevant: agent1 sought the ball, she believed it was hidden in the box, and hence she should reach for the box. The infants who saw the can event should thus look reliably longer than those who saw the box event.

[4] As in Experiment 1, the test data were subjected to an ANCOVA using as covariates the infants' average looking times during the familiarization trials and their looking times during the belief-induction trial. The ANCOVA again revealed a significant interaction between condition and test event, $F(1, 26) = 17.03$, $p < .0005$. Planned comparisons confirmed that in the intervention condition, the infants who saw the box event looked reliably longer than those who saw the can event, $F(1, 26) = 5.56$, $p < .05$, whereas in the no-intervention condition, those who saw the can event looked reliably longer than those who saw the box event, $F(1, 26) = 12.19$, $p < .0025$.

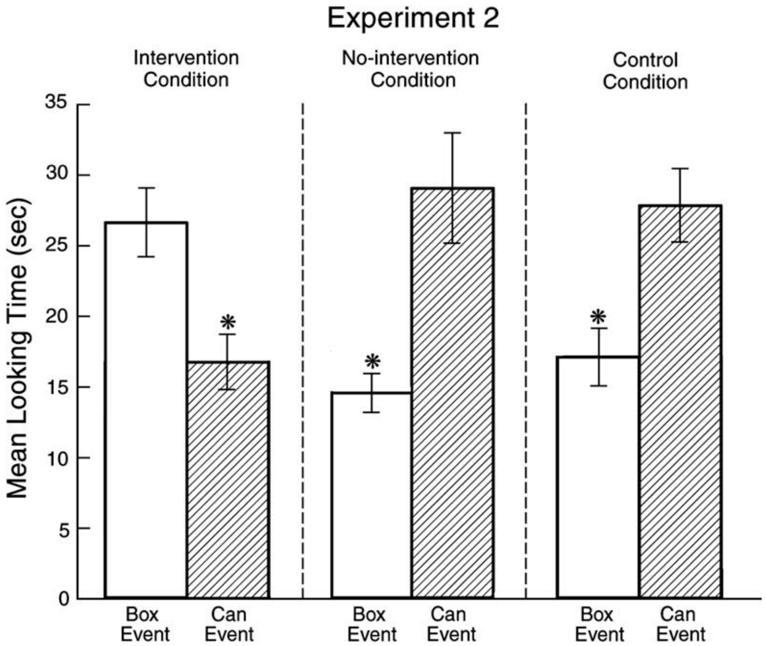


**Fig. 10.** Mean looking times to the box and can test events for the infants in the intervention, no-intervention, and control conditions of Experiment 2. Error bars represent standard errors for each condition.

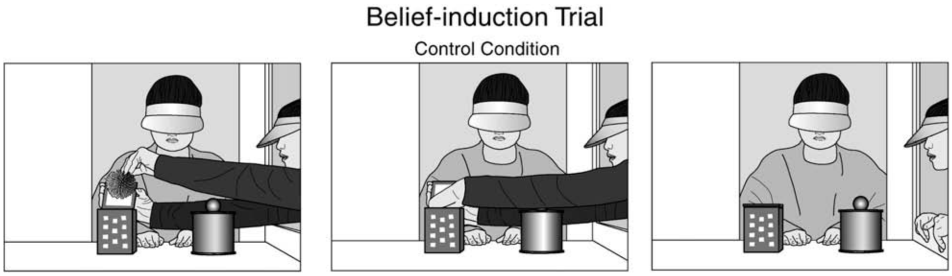


**Fig. 11.** Schematic drawing of the event shown during the belief-induction trial in the control condition of Experiment 2.

Participants were 16 healthy term infants, 9 male and 7 female (range: 18 months, 7 days to 19 months, 0 days; *M* = 18 months, 16 days). Another 6 infants were tested but not included in the analyses, because they were distracted (1), fussy (1), or difficult to observe (1), looked for the maximum amount of time allowed in the test trial (2), or looked more than 3 SD from the mean of their condition in the test trial (1). Half the infants saw the box event and half saw the can event.

As in the intervention condition, the belief-induction trial consisted of a 12-s initial phase followed by a final phase. The set-up at the beginning of the trial was the same as in the second and third familiarization trials, with both agents present. To start, agent2's right hand grasped the box's lid (1 s) and opened it (1 s); next, agent2's left hand reached into the box (1 s), removed the ball (1 s), rotated it to the left and right twice (4 s), returned it inside the box (1 s), and then withdrew to its initial position on the lower edge of the window (1 s). Finally, agent2's right hand closed the box's lid (1 s) and withdrew to the window (1 s). During the final phase, the agents paused in their same positions until the trial ended.

Examination of the initial phase of each trial revealed that the infants were highly attentive: they looked for 14.8/15 s, 14.5/15 s, and 2.5/3 s, respectively, during the first three familiarization trials, for 11.8/12 s during the belief-induction trial, for 8.5/9 s during the intervention trial, and for 2.9/3 s during the test trial.

Interobserver agreement during the final phases of the familiarization and test trials was measured for all 16 infants and averaged 95% per trial per infant.

The infants' looking times during the final phases of the *familiarization* trials were averaged and compared to those of the infants in the intervention condition by means of $2 \times 2$ ANOVA with condition (intervention or control) and test event (box or can) as between-subjects factors. The analysis yielded no significant effects, all $Fs(1, 28) < 1$, suggesting that the infants tended to look equally during the familiarization trials (intervention/box-event, *M* = 27.1, SD = 15.8; intervention/can-event, *M* = 25.6, SD = 8.7; control/box-event, $M = 23.5$, SD = 8.4; control/can-event: $M = 24.6$, SD = 8.9). Analysis of the infants' looking times during the *belief-induction* trial again produced no significant effects, all $Fs(1, 28) < 1.66$, $p > .20$, suggesting that the infants also looked about equally during the belief-induction trial (intervention/box-event, $M = 14.9$, SD = 10.9; intervention/can-event, $M = 18.4$, SD = 10.8; control/box-event, $M = 12.3$, SD = 8.8; control/can-event, $M = 18.7$, SD = 12.5).

The infants' looking times during the final phase of the *test* trial (see Fig. 10) were analyzed as above. Neither the main effect of condition nor that of test event was significant, both $Fs(1, 28) < 1$. However, the interaction between condition and test event was significant, $F(1, 28) = 19.33$, $p < .00025$. Planned comparisons indicated that (1) in the intervention condition, the infants who saw the box event ($M = 26.7$, SD = 7.0) looked reliably longer than those who saw the can event ($M = 16.9$, SD = 5.7), $F(1, 28) = 8.85$, $p < .01$, and (2) in the control condition, the infants who saw the can event ($M = 27.9$, SD = 7.4) looked reliably longer than those who saw the box event ($M = 17.2$, SD = 5.9), $F(1, 28) = 10.52$, $p < .005$.[5]

Non-parametric Wilcoxon rank-sum tests confirmed the results of the intervention ($W_S = 41$, $p < .01$) and control ($W_S = 42$, $p < .01$) conditions.

### 3.4. Discussion

In the intervention condition, the infants who saw the box event looked reliably longer than those who saw the can event; in the no-intervention condition, the infants showed the reverse looking pattern. These results suggest that the infants in the intervention condition (1) realized following the belief-induction trial that agent1 now held the false belief that the ball was still in the box; (2) assumed that agent2's pointing at the can during the intervention trial was sufficient to correct this false belief; and hence (3) expected agent1 to reach into the can and were surprised when she reached into the box instead.

Like the infants in the intervention condition, those in the no-intervention condition realized following the belief-induction trial that agent1 now held a false belief about the ball's location. Because no intervention came to correct this false belief, the infants expected agent1 to act on it. Thus, they expected agent1 to look for the ball in the box, where she falsely believed it to be hidden, and they were surprised when she reached for the can instead.

The preceding interpretations were supported by the result of the control condition. As expected, the infants in this condition looked reliably longer at the can than at the box event. This result suggests that the infants (1) realized following the belief-induction trial that agent1 still held a true belief about the ball's location (since agent2 had replaced the ball into the box, in plain view of agent1); (2) assumed that agent2's pointing at the can in the intervention trial was not meant to convey information about the ball's new location; and hence (3) expected agent1 to search for the ball in the box in the test trial and were surprised when she reached into the can instead. The result of the control condition makes clear that the infants in the intervention condition did not expect agent1 to reach into the can in the test trial simply because agent2 had pointed to it in the preceding trial; the infants in the control condition received exactly the same intervention trial and yet they expected agent1 to reach for the box, where she believed the ball was hidden, as opposed to the can, where agent2 had pointed.

The results of Experiment 2 extend previous results in several ways. First, they add to the recent evidence—reviewed in the Introduction—that infants in the second year of life are able to attribute false beliefs to others (Onishi & Baillargeon, 2005; Song & Baillargeon, in press; Southgate et al., 2007; Surian et al., 2007). Second, the results support those of Experiment 1 and provide additional evidence that 18-month-old infants understand that false beliefs can be corrected by appropriate interventions. In the present research, agent1 was not present when agent2 transferred the ball from the box, where agent1 had hidden it, to a new hiding place; as a result, agent1 held a false belief about the ball's location. Infants expected this false belief to be corrected when, upon agent1's return, agent2 used either words (Experiment 1) or gestures (Experiment 2) to communicate that the ball was now in the new hiding place.

Finally, the results of Experiment 2 extend O'Neill's (1996) findings about young children's use of pointing gestures to inform others (see also Liszkowski et al., 2006). In O'Neill's experiments, 2-year-old children watched an experimenter hide an object in one of two containers located out of reach, and then they could ask their parent to retrieve the object for them. The children were significantly more likely to gesture to the correct container when the parent had not witnessed the hiding event than when she had. The children thus recognized that their parent was ignorant about the object's hiding place, and they acted to provide the necessary information. The results of Experiment 2 extend these results by showing that infants as young as 18 months understand the use of a pointing gesture to communicate information to an agent about an object's hiding place, even when (1) the agent is an unfamiliar adult, rather than a parent; (2) the agent is mistaken, as opposed to simply ignorant, about the object's hiding place; (3) the gesture is performed by the individual who hid the object, rather than by the infants themselves; and (4) the infants are attempting to make sense of the agent's subsequent actions (i.e., determine where the agent should look for the object), rather than trying to achieve some goal of their own (i.e., obtain the object).

#### 3.4.1. Updating or overwriting?

The results of the informative-intervention condition in Experiment 1 and of the intervention condition in Experiment 2 suggest that the infants expected agent1's false belief about the ball's location to be corrected when she

[5] The test data were also subjected to an ANCOVA using as covariates the infants' average looking times during the familiarization trials and their looking times during the belief-induction trial. The ANCOVA again revealed a significant interaction between condition and test event, $F(1, 26) = 17.71$, $p < .0005$. Planned comparisons confirmed that in the intervention condition, the infants who saw the box event looked reliably longer than those who saw the can event, $F(1, 26) = 8.77$, $p < .01$, whereas in the control condition, those who saw the can event looked reliably longer than those who saw the box event, $F(1, 26) = 9.07$, $p < .01$.

was informed—through words or pointing gestures—about the ball's new location. But how did this correction process actually work? There are at least two possibilities. One is that, as trials progressed, the infants successively *updated* their representation of agent1's belief, from her initial true belief to her subsequent false belief and her final corrected belief, keeping track in some manner of the history of her beliefs. The other possibility is that the infants simply *overwrote* their representation of agent1's belief with a new representation whenever events required such a change, without preserving a memory of her successive beliefs.

Did the infants in the informative-intervention condition of Experiment 1 and in the intervention condition of Experiment 2 update or overwrite their representation of agent1's belief across trials? The present results are insufficient to answer this question: they suggest only that the infants expected agent2's communications to correct agent1's false belief. Whether the infants also kept track of the history of agent1's beliefs will have to be established through further research.

The preceding analysis presupposes that the infants in the informative-intervention condition of Experiment 1 and in the intervention condition of Experiment 2 attributed to agent1 a false belief about the ball's location, but is this assumption correct? It might be suggested that the infants considered agent1's belief about the ball's location only upon observing her search. Because agent1 did not search in the belief-induction or in the intervention trial, the infants would not have considered her belief about the ball's location until the test trial. Since she had just been informed by agent2 of the ball's new location, the infants in these conditions might never have attributed to agent1 a false belief about the ball's location—only a true belief based on agent2's communication (The infants in the uninformative-intervention condition of Experiment 1 and in the no-intervention condition of Experiment 2 had to attribute to agent1 a false belief, to explain why she searched for the ball in the wrong location).

The interpretation just proposed differs radically from the one we have put forth here: rather than expecting agent2's communication to *correct* agent1's false belief about the ball's location, the infants would simply have expected this communication to *inform* agent1 about where to search. However, we believe that this alternative interpretation is unlikely, for two reasons. First, keeping track of agent1's belief about the ball's location as events unfolded (whether through updating or overwriting) would have been much easier for infants, in terms of information-processing load, than attempting in the test trial to recall what had happened in previous trials so as to determine where agent1 might search. To see why, consider in particular the results of the uninformative-intervention condition of Experiment 1: to form the correct expectation that the agent would reach for the box, the infants would have had to remember what had happened in the familiarization, belief-induction, and intervention trials. It seems to us more plausible that the infants recognized early on that the agent wanted the ball and kept track in each trial of her belief about its location.

A second reason for preferring the present interpretation has to do with the results of Southgate et al. (2007) discussed in the Introduction. These results make clear that, in situations where an agent repeatedly searches for an object in one of two locations, 25-month-olds actively anticipate or predict where the agent will search, beginning in the familiarization trials. Recent findings suggest that 18-month-old infants also anticipate where an agent will search for an object (Neumann, Thoermer, & Sodian, 2008). These data support the notion that the infants in Experiments 1 and 2 expected agent1 to search for her ball whenever she was present, and that they therefore consistently attended to her belief about its location.

## 4. General discussion

The infants in the present experiments expected agent1 to act in accordance with her belief about the ball's location, whether this belief was true or false. When agent1 held the *true* belief that the ball was in the box (control condition of Experiment 2), the infants expected her to search in the box and were surprised when she reached for the other container instead. Similarly, when agent1 held the *false* belief that the ball was in the box, because she was absent when it was moved to the other container (no-intervention condition of Experiment 2), the infants again expected her to search in the box and were surprised when she did not.

The infants also understood that agent1's false belief about the ball's location could be *corrected* by a communication from agent2. If, upon agent1's return, agent2 either stated that the ball was in the other container (informative-intervention condition of Experiment 1), or pointed to the other container (intervention condition of Experiment 2), the infants now expected agent1 to reach for the other container and were surprised when she searched in the box instead.

Finally, not only did the infants understand that agent1's false belief about the ball's location could be corrected by a communication from agent2, they also showed some ability to distinguish between an *appropriate* and an *inappropriate* communication. In the uninformative-intervention condition of Experiment 1, instead of telling agent1 "The ball is in the cup!", agent2 simply said "I like the cup!". Although agent2 mentioned the cup, the infants viewed this communication as insufficient to inform agent1 about the ball's new location. As a result, they expected agent1 to remain mistaken about the ball's location, and they were surprised when she reached for the cup.

The present findings thus provide additional evidence (Onishi & Baillargeon, 2005; Song & Baillargeon, in press; Southgate et al., 2007; Surian et al., 2007) that infants in the second year of life attribute false beliefs to agents. The present findings also extend prior results in showing that, by 18 months of age, infants understand that agents' false beliefs can be corrected by appropriate interventions involving words or pointing gestures.

### 4.1. Alternative interpretations

As was mentioned in the Introduction, a number of alternative interpretations have been offered for recent findings from anticipatory-looking and VOE tasks that some understanding of false belief emerges long before 3 or 4 years of age (Onishi & Baillargeon, 2005; Song & Baillargeon, in press; Southgate et al., 2007; Surian et al., 2007). To illustrate, three alternative interpretations have been offered for the results of Onishi and Baillargeon (2005). One (*association-based*) interpretation is that the infants formed a three-way association between the agent, the toy, and the location where the toy became hidden, and they looked reliably longer when the test trial deviated from this association (Perner & Ruffman, 2005; Ruffman & Perner, 2005). Another (*rule-based*) interpretation is that the infants brought to the task a behavioral rule that agents who are searching for a hidden object will search for it where they last saw it disappear (Perner & Ruffman, 2005; Ruffman & Perner, 2005). Yet another *(ignorance-based*) interpretation is that the infants possessed a general expectation that ignorance leads to error: thus, agents who are ignorant about a toy's location will search for it in the wrong location (Southgate et al., 2007). We next consider each interpretation in turn, in light of the present and other findings.

#### 4.1.1. Association-based interpretation

Let us accept, for the sake of argument, that infants who watch an agent (agent1) hide a toy in one of two locations (locationA) form an agent1-toy-locationA association (e.g., Perner & Ruffman, 2005; but see Csibra & Southgate, 2006). In order to account for the present findings as well as those of Onishi and Baillargeon (2005), Southgate et al. (2007), and Surian et al. (2007), we *also* need to assume that this association, even when reinforced for several trials, is immediately replaced by a new agent1-toy-locationB association (1) if agent1 simply watches the toy disappear in locationB; (2) if agent1, after being absent from the situation, is told by agent2 that the toy is in locationB; or (3) if agent1, after being absent from the situation, sees agent2 point to locationB. On the other hand, the agent1-toy-locationA association is *not* revised (1) if agent1, after being absent from the situation, is told some irrelevant fact about locationB, or (2) if agent1 was present during the trial preceding the pointing to locationB and saw agent2 replace the toy in locationA. The difficulty here, of course, is that it is highly implausible that a low-level association could be readily modulated in this manner, by various high-level factors having to do with conditions under which agent1 might or might not acquire relevant knowledge about the toy's new location.

Another difficulty with the association-based interpretation is that it cannot easily explain the data of Song and Baillargeon (in press). To illustrate, let us assume that the infants in the false-perception condition of the skunk experiment formed a three-way association during the last two familiarization trials between the agent, the skunk, and the shallow container in which the skunk was placed. This association could not have been revised during test, because the agent did not enter the apparatus until after the doll was hidden in the plain box and the skunk in the hair box. Thus, the fact that the infants who saw the hair-box event looked reliably longer than those who saw the plain-box event could not be explained by saying that the hair-box event deviated from the three-way association formed in previous trials: new hiding locations were used in the test trial that were not present during the familiarization trials.

#### 4.1.2. Rule-based interpretation

Let us accept for the moment that infants can extract behavioral rules as they observe agents act in various situations (e.g., Perner & Ruffman, 2005). In order to account for the present findings as well as those reviewed in the Introduction (Onishi & Baillargeon, 2005; Song & Baillargeon, in press; Southgate et al., 2007; Surian et al., 2007), we would need to grant infants the following rules: (1) agents searching for an object will search for it where they saw it disappear; (2) agents searching for an object will search for it where they saw it disappear unless they receive a relevant communication (as construed by the infants) that the object is now in a new location, in which case they will search the new location; (3) agents who see two objects being placed in the same two locations across trials (e.g., apple behind left screen, cheese behind right screen) will search for their preferred object in the location where it was placed in the past; and (4) agents who see two objects being placed in different locations across trials (e.g., doll on left placemat and skunk on right placemat; doll in right shallow container and skunk in left shallow container), and who are finally faced with two lidded boxes, one plain and the other with a part protruding from under its lid that resembles a part belonging to one of the objects, will search the box with the part if they happen to prefer the object with the similar part, and will search the plain box if they happen to prefer the other object. As may already be obvious from this litany of rules, the difficulty with the rule-based interpretation is that, as more experiments provide evidence suggestive of early false-belief understanding, more rules must be invoked to account for infants' responses. Not only is it increasingly far-fetched that infants could have formed all of these rules prior to coming to the laboratory, but parsimony alone supports the simpler explanation that infants can attribute both true and false beliefs to agents (see also Onishi & Baillargeon, 2005; Southgate et al., 2007; Surian et al., 2007).

#### 4.1.3. Ignorance-based interpretation

All of the VOE experiments reviewed in the Introduction (Onishi & Baillargeon, 2005; Song & Baillargeon, in press; Surian et al., 2007), as well as the present experiments, are open to the ignorance-based interpretation raised by Southgate et al. (2007): infants might have attributed to the agent simply a state of ignorance about her toy's location, and they expected this state to lead her to search in the wrong location. Because Southgate et al. used an anticipatory-looking task in which both search locations were wrong (the toy was removed from the scene and thus both locations were empty), their results are not open to this alternative interpretation: their 25-month-olds correctly anticipated where the agent would search when she held a false belief about the toy's location.

The ignorance-based interpretation differs from the association- and rule-based interpretations in one key respect: it assumes that infants attribute internal states to agents to make sense of their actions. At stake here are critical questions about the range of internal states infants can attribute to agents, and about the age of onset of the computational systems involved in processing these internal states. To explain why this is the case, we first describe one possible account of these computational systems; we then return to the ignorance-based interpretation and consider its implications.

### 4.2. An account of early psychological reasoning

According to one account of early psychological reasoning (adapted from Leslie, 1987, 1994a, 1994b, 1995, 2000; see also Johnson, 2000, 2003; Luo & Baillargeon, 2007; Premack, 1990; Premack & Premack, 1995; Scott & Baillargeon, in press; Song & Baillargeon, in press), infants are born with a psychological-reasoning system that provides them with a shallow causal framework for interpreting the intentional actions of agents in terms of *internal states*. The operation of the psychological-reasoning system is assumed to be largely unconscious: infants are not aware of the causal framework they use when reasoning about agents, any more than young children are aware of the grammar of their language as they begin to understand and produce sentences.

The psychological-reasoning system is thought to be composed of at least two subsystems, Subsystem1 (SS1) and Subsystem2 (SS2). In both subsystems, intentional action is assumed to be constrained by a core principle of rationality: when pursuing a goal, agents typically select actions that are not only causally appropriate but also reasonably efficient (e.g., Csibra, 2008; Csibra, Bíró, Koós, & Gergely, 2003; Csibra, Gergely, Bíró, Koós, & Brockbank, 1999; Gergely, Bekkering, & Király, 2002; Gergely & Csibra, 2003; Gergely, Nádasdy, Csibra, & Bíró, 1995).

When infants watch an agent act on objects in a setting, SS1 allows them to attribute two kinds of internal states to the agent: *motivational* and *reality-congruent informational* states. Motivational states specify the agent's motivation in the scene and include goals and dispositions; reality-congruent informational states specify what accurate information the agent possesses (through perception, memory, or inference) or lacks about the setting. SS1 thus allows infants to attribute states of knowledge or ignorance to the agent.

SS2 extends SS1 in that it allows infants to also attribute *reality-incongruent informational* states to the agent. These states concern information about the setting that deviates from reality, in that it is either false or pretend (e.g., Leslie, 1987; Onishi, Baillargeon, & Leslie, 2007). To reason correctly about such states, children must be able to hold in mind two distinct versions of the setting: one that corresponds to reality (as they construe it), and one that is decoupled from reality and corresponds to the agent's representation (e.g., Leslie, 1994a).

Current evidence suggests that SS1 becomes operational in the first months of life and is well in place by the end of the first year. Thus, experiments on *motivational* states have revealed that even young infants can attribute dispositions and goals to agents (e.g., Bíró & Leslie, 2007; Csibra, 2008; Csibra et al., 1999; Hamlin, Wynn, & Bloom, 2007; Johnson, Ok, & Luo, 2007; Kamerawi, Kato, Kanda, Ishiguro, & Hiraki, 2005; Király, Jovanovic, Prinz, Aschersleben, & Gergely, 2003; Luo & Baillargeon, 2005; Sommerville, Woodward, & Needham, 2005; Song & Baillargeon, 2007; Woodward, 1998, 1999). For example, after watching an agent repeatedly reach for objectA as opposed to objectB in a setting, 5-month-olds attribute to the agent a particular disposition, a preference for objectA over objectB. When the objects' positions are reversed, infants expect the agent to approach objectA in its new position, and they look reliably longer if she approaches objectB instead (e.g., Luo & Baillargeon, 2005; Woodward, 1998, 1999).

Similarly, experiments on *reality-congruent informational* states suggest that, by the end of the first year, infants keep track of what objects an agent can or cannot see, and has or has not seen, in a setting (e.g., Brooks & Meltzoff, 2002, 2005; Caron, Kiel, Dayton, & Butler, 2002; Luo & Baillargeon, 2007; Moll & Tomasello, 2004; Surian et al., 2007; Thoermer & Sodian, 2005; Tomasello & Haberl, 2003). Thus, 12.5-month-olds who watch an agent repeatedly reach for objectA over objectB do *not* attribute to the agent a preference for objectA if objectB is hidden from her by a screen; however, they *do* attribute such a preference if the agent is aware of objectB's presence behind the screen, because she saw it there earlier (Luo & Baillargeon, 2007). Recent experiments suggest that infants as young as 6 months consider what objects an agent can or cannot see when interpreting her actions (e.g., Luo & Johnson, in press).

At what point does SS2 become operational? The answer to this question depends on how we interpret the results of the VOE experiments reported here as well as those presented in the Introduction (Onishi & Baillargeon, 2005; Song & Baillargeon, in press; Surian et al., 2007). If Southgate et al. (2007) are correct that the infants attributed only *ignorance* to the agent, then these experiments would simply reflect the operation of SS1, which is sufficient to handle knowledge and ignorance states. The earliest evidence that SS2 is operational (outside of pretense; see Onishi et al., 2007, for discussion) would then come from the findings of Southgate et al. (2007) and Neumann et al. (2008) showing that infants aged 18 months and older anticipate where an agent who holds a false belief about an object's location will search for it. On the other hand, if the infants in the VOE experiments attributed a false belief rather than ignorance to the agent, then the earliest evidence that SS2 is operational would come from the experiment of Surian et al. (2007), at 13 months of age (see also He & Baillargeon, 2007, for preliminary evidence with 11-month-old infants).

### 4.3. Tests of the ignorance-based interpretation

The ignorance-based interpretation proposed by Southgate et al. (2007) rests on a crucial assumption, that infants expect ignorance to lead to error: an agent who does not know that a toy is in locationA, not locationB, will search for it in locationB. Is this assumption correct? In everyday life, ignorance does not systematically lead to error: rather, it leads to random behavior. When one looks for a water glass in an unfamiliar kitchen, one may find the cupboard in which glasses are kept on the first try or only after opening several other cupboards. Do infants expect an ignorant agent to err, or do they hold no expectation as to where she will search? We have addressed this question in four VOE projects with infants aged 11–30 months (e.g., He & Baillargeon, 2007; He, Bolz, & Baillargeon, 2007; Scott & Baillargeon, in press; Scott et al., 2007). Results in all four projects indicated that infants did not, in fact, expect ignorance to lead to error: instead, and quite reasonably, infants held no expectation as to how the ignorant agent would act.

To illustrate, one project focused on 18-month-olds' ability to attribute to an agent a false belief about the identity of an object (Scott & Baillargeon, in press). The infants were assigned to a false-belief, a true-belief, or an ignorance condition. In the *false-belief* condition, the infants first watched live familiarization events in which a female agent faced two toy penguins, one that did not come apart (joined penguin) and one that did (unjoined penguin). Across trials, the two penguins were placed on placemats or inside shallow containers by an experimenter's gloved hands. The unjoined penguin was always presented in two pieces; the agent hid a key in the bottom piece, stacked the top piece on top of the bottom piece (the unjoined penguin then looked identical to the joined penguin), and held on to the penguin until the trial ended. During the test trials, in the agent's absence, the hands stacked the unjoined penguin and placed it under a large transparent cover; the hands then placed the joined penguin under a large opaque cover (either cover was wide enough to hold the two pieces of the unjoined penguin side by side). The agent entered the apparatus at this point, reached for the transparent or opaque cover, and paused. If the infants reasoned that (1) the agent would falsely assume that the penguin under the transparent cover was the joined penguin, since the unjoined penguin had always been presented in two pieces in previous trials, and (2) the agent would falsely conclude that the unjoined penguin was under the opaque cover, then they should expect the agent to reach for the opaque cover and be surprised when she reached for the transparent cover instead. The *true-belief* condition was similar except that the agent watched all of the gloved hands' actions and thus knew that the unjoined penguin was under the transparent cover. Finally, the *ignorance* condition was again similar to the false-belief condition except that both covers were either opaque or transparent, so that the agent could not know, when she entered the apparatus, under which cover the unjoined penguin had been placed. As expected, the infants in the false-belief condition looked reliably longer when the agent reached for the transparent as opposed to the opaque cover; the infants in the true-belief condition showed the reverse looking pattern; and the infants in the ignorance condition looked about equally at the two events (control results indicated that the infants had not merely forgotten where the unjoined penguin was located).

These results suggest that infants respond differently when an agent is mistaken, as opposed to ignorant, about which of two locations her preferred toy has been placed in; they expect the mistaken agent to search the location where she falsely believes her toy to be hidden, but they hold no expectation about which location the ignorant agent will search. These results (and similar results obtained with younger, 11-month-old infants; He & Baillargeon, 2007) help lessen concerns that false-belief VOE findings are open to an ignorance-based interpretation.

### 4.4. Concluding remarks

If infants in the second year of life possess some understanding of false belief, as suggested by recent anticipatory-looking and VOE tasks (Neumann et al., 2008; Onishi & Baillargeon, 2005; Scott & Baillargeon, in press; Song & Baillargeon, in press; Southgate et al., 2007; Surian et al., 2007), then we should be able to find evidence of this understanding in a wide variety of tasks. In line with this analysis, we found in the present experiments that 18-month-old infants not only attributed a false belief to an agent who did not witness the transfer of her toy to a new location, but also expected this false belief to be corrected when the agent received a relevant, but not an irrelevant, communication about the toy's new location. These results add to the evidence that SS2—the computational subsystem assumed to be involved in the processing of reality-incongruent informational states—is already functional in the second year of life.

## Acknowledgements

This research was supported by grants from Yonsei University to Hyun-joo Song (2006-1-0029), from Fonds Québécois de la Recherche sur la Société et la Culture (FQRSC) and the Natural Sciences and Engineering Research Council of Canada (NSERC) to Kristine H. Onishi, and from NICHD to Renée Baillargeon (HD-021104) and Cynthia Fisher (HD-044458). We thank Gergely Csibra and Luca Surian for helpful comments; the staff of the University of Illinois Infant Cognition Laboratory for their help with the data collection; and the parents and infants who participated in the research.

## References

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition, 21*, 37–46.

Bíró, S., & Leslie, A. M. (2007). Infants' perception of goal-directed actions: Development through cue-based bootstrapping. *Developmental Science, 10*, 379–398.

Bloom, P., & German, T. P. (2000). Two reasons to abandon the false belief task as a test of theory of mind. *Cognition, 77*, B25–B31.

Brooks, R., & Meltzoff, A. N. (2002). The importance of eyes: How infants interpret adult looking behavior. *Developmental Psychology, 38*, 958–966.

Brooks, R., & Meltzoff, A. N. (2005). The development of gaze following and its relation to language. *Developmental Science, 8*, 535–543.

Butterworth, G., & Grover, L. (1988). The origins of referential communication in human infancy. In L. Weiskrantz (Ed.), *Thought without language* (pp. 5–24). Oxford: Clarendon.

Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development, 72*, 1032–1053.

Carlson, S. M., Moses, L. J., & Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contributions of inhibitory control and working memory. *Infant and Child Development, 11*, 73–92.

Caron, A. J., Kiel, E. J., Dayton, M., & Butler, S. (2002). Comprehension of the referential intent of looking and pointing between 12 and 15 months. *Journal of Cognition and Development, 3*, 445–464.

Carpenter, M., Nagell, K., & Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of

age. *Monographs of the Society for Research in Child Development, 255*, 1–143.
Clements, W. A., & Perner, J. (1994). Implicit understanding of belief. *Cognitive Development, 9*, 377–395.
Csibra, G. (2008). Goal attribution to inanimate agents by 6.5-month-old infants. *Cognition, 107*, 705–717.
Csibra, G., Bíró, S., Koós, O., & Gergely, G. (2003). One-year-old infants use teleological representations of actions productively. *Cognitive Science, 27*, 111–133.
Csibra, G., Gergely, G., Bíró, S., Koós, O., & Brockbank, M. (1999). Goal attribution without agency cues: The perception of 'pure reason' in infancy. *Cognition, 72*, 237–267.
Csibra, G., & Southgate, V. (2006). Evidence for infants' understanding of false beliefs should not be dismissed. Response to Ruffman and Perner. *Trends in Cognitive Sciences, 10*, 4–5.
Flavell, J. H. (1988). The development of children's knowledge about the mind: From cognitive connections to mental representations. In J. W. Astington, P. L. Harris, & D. R. Olson (Eds.), *Developing theories of mind* (pp. 244–267). New York: Cambridge University Press.
Fodor, J. A. (1992). A theory of the child's theory of mind. *Cognition, 44*, 283–296.
Franco, F., & Butterworth, G. (1996). Pointing and social awareness: Declaring and requesting in the second year. *Journal of Child Language, 23*, 307–336.
Garnham, W. A., & Ruffman, T. (2001). Doesn't see, doesn't know: Is anticipatory looking really related to understanding of belief? *Developmental Science, 4*, 94–100.
Gergely, G., Bekkering, H., & Király, I. (2002). Rational imitation in preverbal infants. *Nature, 415*, 6873.
Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naïve theory of rational action. *Trends in Cognitive Sciences, 7*, 287–292.
Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition, 56*, 165–193.
Gopnik, A., & 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 Development, 59*, 26–37.
Gopnik, A., & Wellman, A. (1994). The "theory theory". In L. Hirschfield & S. A. Gelman (Eds.), *Domain specificity in culture and cognition* (pp. 257–293). New York: Cambridge University Press.
Hala, S., Hug & Henderson, H. (2003). Executive functioning and false belief understanding in preschool children: Two tasks are harder than one. *Journal of Cognition and Development, 4*, 275–298.
Hamlin, J. K., Wynn, K., & Bloom, P. (2007). Social evaluation by preverbal infants. *Nature, 450*, 557–559.
He, Z., & Baillargeon, R. (2007). *Understanding of false belief in 11-month-old infants.* Paper presented at the Biennial meeting of the society for research in child development, Boston, MA.
He, Z., & Baillargeon, R. (2008). *False-belief understanding about number in 19-month-old infants*. Paper to be presented at the Biennial international conference on infant studies, Vancouver, Canada.
He, Z., Bolz, M., & Baillargeon, R. (2007). *Understanding of false belief in 2.5-year-olds in a violation-of-expectation task*. Paper presented at the Biennial meeting of the society for research in child development, Boston, MA.
Hirsh-Pasek, K., & Golinkoff, R. M. (1996). *The origins of grammar*. Cambridge, MA: MIT Press.
Johnson, S. C. (2000). The recognition of mentalistic agents in infancy. *Trends in Cognitive Sciences, 4*, 22–28.
Johnson, S. C. (2003). Detecting agents. *Philosophical Transactions of the Royal Society of London, Series B, 358*, 549–599.
Johnson, S. C., Ok, S., & Luo, Y. (2007). The attribution of attention: 9-month-olds' interpretation of gaze as goal-directed action. *Developmental Science, 10*, 530–537.
Kamerawi, K., Kato, M., Kanda, T., Ishiguro, H., & Hiraki, K. (2005). Six-and-a-half-month-old children positively attribute goals to human action and to humanoid-robot motion. *Cognitive Development, 20*, 303–320.
Király, I., Jovanovic, B., Prinz, W., Aschersleben, G., & Gergely, G. (2003). The early origins of goal attribution in infancy. *Consciousness and Cognition, 12*, 752–769.
Kovács, Á. M. (in press). Early bilingualism enhances mechanisms of false-belief reasoning. *Developmental Science*.
Lempers, J. D. (1979). Young children's production and comprehension of nonverbal deictic behaviors. *Journal of Genetic Psychology, 135*, 93–102.
Leslie, A. M. (1987). Pretense and representation: The origins of "theory of mind". *Psychological Review, 94*, 412–426.
Leslie, A. M. (1994a). Pretending and believing: Issues in the theory of ToMM. *Cognition, 50*, 211–238.
Leslie, A. M. (1994b). ToMM, ToBy, and Agency: Core architecture and domain specificity. In L. Hirschfeld & S. A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture* (pp. 119–148). New York: Cambridge University Press.
Leslie, A. M. (1995). A theory of agency. In D. Sperber, D. Premack, & A. J. Premack (Eds.), *Causal cognition: A multidisciplinary debate* (pp. 121–141). Oxford, United Kingdom: Clarendon Press.
Leslie, A. M. (2000). How to acquire a 'representational theory of mind'. In D. Sperber (Ed.), *Metarepresentations: A multidisciplinary perspective* (pp. 197–223). Oxford: Oxford University Press.
Leslie, A. M. (2005). Developmental parallels in understanding minds and bodies. *Trends in Cognitive Sciences, 9*, 459–462.
Leslie, A. M., German, T. P., & Polizzi, P. (2005). Belief-desire reasoning as a process of selection. *Cognitive Psychology, 50*, 45–85.
Leslie, A. M., & Polizzi, P. (1998). Inhibitory processing in the false belief task: Two conjectures. *Developmental Science, 1*, 247–253.
Leung, E. H. L., & Rheingold, H. L. (1981). Development of pointing as a social gesture. *Developmental Psychology, 17*, 215–220.
Liszkowski, U., Carpenter, M., Henning, A., Striano, T., & Tomasello, M. (2004). 12-month-olds point to share attention and interest. *Developmental Science, 7*, 297–307.
Liszkowski, U., Carpenter, M., Striano, T., & Tomasello, M. (2006). 12- and 18-month-olds point to provide information for others. *Journal of Cognition and Development, 7*, 173–187.
Luo, Y., & Baillargeon, R. (2005). Can a self-propelled box have a goal? Psychological reasoning in 5-month-old infants. *Psychological Science, 16*, 601–608.
Luo, Y., & Baillargeon, R. (2007). Do 12.5-month-old infants consider what objects others can see when interpreting their actions? *Cognition, 105*, 489–512.
Luo, Y., & Johnson, S. C. (in press). Recognizing the role of perception in action at 6 months. *Developmental Science*.
Meints, K., Plunkett, K., Harris, P. L., & Dimmock, D. (2002). What is 'on' and 'under' for 15-, 18- and 24- month-olds? Typicality effects in early comprehension of spatial prepositions. *British Journal of Developmental Psychology, 20*, 113–130.
Moll, H., & Tomasello, M. (2004). 12- and 18-month-olds follow gaze behind barriers. *Developmental Science, 7*, F1–F9.
Moore, C., Pure, K., & Furrow, D. (1990). Children's understanding of the modal expression of speaker certainty and uncertainty and its relation to the development of a representational theory of mind. *Child Development, 61*, 722–730.
Neumann, A., Thoermer, C., & Sodian, B. (2008). *False belief understanding in 18-month-olds' anticipatory looking behavior: An eye-tracking study.* Paper presented at the XXIX international congress of psychology, Berlin, Germany.
O'Neill, D. K. (1996). Two-year-old children's sensitivity to a parent's knowledge state when making requests. *Child Development, 67*, 659–677.
Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science, 308*, 255–258.
Onishi, K. H., Baillargeon, R., & Leslie, A. M. (2007). 15-month-old infants detect violations in pretend scenarios. *Acta Psychologica, 124*, 106–128.
Perner, J. (1991). *Understanding the representational mind*. Cambridge, MA: MIT Press.
Perner, J. (1995). The many faces of belief: Reflections on Fodor's and the child's theory of mind. *Cognition, 57*, 241–269.
Perner, J., Leekam, S. R., & Wimmer, H. (1987). Three-year-olds' difficulty with false belief: The case for a conceptual deficit. *British Journal of Developmental Psychology, 5*, 125–137.
Perner, J., & Ruffman, T. (2005). Infants' insight into the mind: How deep? *Science, 308*, 214–216.
Premack, D. (1990). The infant's theory of self-propelled objects. *Cognition, 36*, 1–16.
Premack, D., & Premack, A. J. (1995). Origins of human social competence. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 205–218). Cambridge, MA: MIT Press.
Ruffman, T., & Perner, J. (2005). Do infants really understand false belief? Response to Leslie. *Trends in Cognitive Sciences, 9*, 462–463.
Russell, J. (1996). *Agency: Its role in mental development*. Hove, England: Erlbaum.
Russell, J., Mauthner, N., Sharpe, S., & Tidswell, T. (1991). The 'windows task' as a measure of strategic deception in preschoolers and autistic subjects. *British Journal of Developmental Psychology, 9*, 331–349.
Scott, R., & Baillargeon, R. (in press). Which penguin is this? Attributing false beliefs about object identity at 18 months. *Child Development*.

Scott, R., Song, H., Baillargeon, R., & Leslie, A. M. (2007). *18.5-month-old infants attribute to others false beliefs about objects' internal properties*. Paper presented at the Biennial meeting of the society for research in child development, Boston, MA.
Sommerville, J. A., Woodward, A. L., & Needham, A. (2005). Action experience alters 3-month-old infants' perception of others' actions. *Cognition, 96*, B1–B11.
Song, H., & Baillargeon, R. (2007). Can 9.5-month-old infants attribute to an actor a disposition to perform a particular action on objects? *Acta Psychologica, 124*, 79–105.
Song. H., & Baillargeon, R. (in press). Infants' reasoning about others' false perceptions. *Developmental Psychology*.
Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief by two-year-olds. *Psychological Science, 18*, 587–592.
Surian, L., Caldi, S., & Sperber, D. (2007). Attribution of beliefs to 13-month-old infants. *Psychological Science, 18*, 580–586.
Thoermer, C., & Sodian, B. (2005). *One-year-olds' level-1 visual perspective-taking abilities in a false-labeling context*. Paper presented at the Biennial meeting of the society for research in child development, Atlanta, GA.
Tomasello, M., & Haberl, K. (2003). Understanding attention: 12- and 18-month-olds know what's new for other persons. *Developmental Psychology, 39*, 906–912.
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.
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.
Woodward, A. L. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition, 69*, 1–34.
Woodward, A. L. (1999). Infants' ability to distinguish between purposeful and nonpurposeful behavior. *Infant Behavior and Development, 22*, 145–160.
Woodward, A. L., & Guajardo, J. J. (2002). Infants' understanding of the point gesture as an object-directed action. *Cognitive Development, 17*, 1061–1084.