Cognitive Psychology 82 (2015) 32–56

Contents lists available at ScienceDirect

Cognitive Psychology

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

Infants understand deceptive intentions to implant false beliefs about identity: New evidence for early mentalistic reasoning

Rose M. Scott[a,*], Joshua C. Richman[b], Renée Baillargeon[b]

[a] Psychological Sciences, University of California Merced, 5200 North Lake Road, Merced, CA 95343, United States
[b] Psychology, University of Illinois, 603 E. Daniel St., Champaign, IL 61820, United States

ARTICLE INFO

Article history:
Accepted 17 August 2015

Keywords:
Deception
False-belief understanding
Psychological reasoning
Social cognition
Infant development
Cognitive development

ABSTRACT

Are infants capable of representing false beliefs, as the mentalistic account of early psychological reasoning suggests, or are they incapable of doing so, as the minimalist account suggests? The present research sought to shed light on this debate by testing the minimalist claim that a signature limit of early psychological reasoning is a specific inability to understand false beliefs about identity: because of their limited representational capabilities, infants should be unable to make sense of situations where an agent mistakes one object for another, visually identical object. To evaluate this claim, three experiments examined whether 17-month-olds could reason about the actions of a deceptive agent who sought to implant in another agent a false belief about the identity of an object. In each experiment, a thief attempted to secretly steal a desirable rattling toy during its owner's absence by substituting a less desirable silent toy. Infants realized that this substitution could be effective only if the silent toy was visually identical to the rattling toy (Experiment 1) and the owner did not routinely shake her toy when she returned (Experiment 2). When these conditions were met, infants expected the owner to be deceived and to mistake the silent toy for the rattling toy she had left behind (Experiment 3). Together, these results cast doubt on the minimalist claim that infants cannot represent false beliefs about identity. More generally, these results indicate that infants in the 2nd year of life can reason not only about the actions of agents who *hold* false beliefs, but also about the actions of agents who seek to *implant* false beliefs, thus providing new support for the

* Corresponding author.
E-mail address: rscott@ucmerced.edu (R.M. Scott).

http://dx.doi.org/10.1016/j.cogpsych.2015.08.003

mentalistic claim that an abstract capacity to reason about false beliefs emerges early in human development.

1. Introduction

Adults routinely interpret others' actions in terms of underlying mental states, and developmental researchers have long been interested in determining when and how this ability develops. Over the past two decades, numerous reports have presented evidence that infants can attribute to agents motivational states (e.g., goals and dispositions), epistemic states (e.g., knowledge and ignorance), and counterfactual states (e.g., false beliefs and pretense) (for reviews, see Baillargeon, Scott, & Bian, in press; Baillargeon et al., 2015). These findings have led many investigators to adopt a *mentalistic* account, which assumes that infants are equipped with a psychological-reasoning system that provides them with a skeletal causal framework for representing and learning about agents' mental states (e.g., Barrett et al., 2013; Buttelmann, Carpenter, & Tomasello, 2009; Carruthers, 2013; Kovács, Téglás, & Endress, 2010; Luo, 2011; Onishi & Baillargeon, 2005; Scott, Baillargeon, Song, & Leslie, 2010; Scott, Roby, & Smith, in press; Southgate, Senju, & Csibra, 2007; Surian, Caldi, & Sperber, 2007).

Recently, however, a number of researchers have offered an alternative, *minimalist* account of these prior findings, which particularly affects claims concerning infants' ability to attribute counterfactual states (e.g., Apperly & Butterfill, 2009; Butterfill & Apperly, 2013; Low, Drummond, Walmsley, & Wang, 2014; Low & Watts, 2013). This account assumes that two distinct systems underlie human psychological reasoning. The (conscious, nonautomatic, slow, flexible) *late-developing* system emerges around age 4 as a result of linguistic, executive-function, and metarepresentational advances; this advanced system is capable of representing false beliefs and other counterfactual states, and it enables correct responses in traditional false-belief tasks (e.g., Baron-Cohen, Leslie, & Frith, 1985; Wellman & Bartsch, 1988; Wellman, Cross, & Watson, 2001; Wimmer & Perner, 1983). The (unconscious, automatic, fast, inflexible) *early-developing* system is already present in infancy; although it cannot represent false beliefs and other counterfactual states, it can track simpler, belief-like states that are sufficient to allow infants to succeed at non-traditional false-belief tasks (e.g., Buttelmann et al., 2009; Onishi & Baillargeon, 2005; Scott et al., 2010; Southgate et al., 2007).

Is the mentalistic or the minimalist account of early false-belief understanding correct? In the following sections, we begin by discussing two key aspects of the minimalist account. First, we explain how the early-developing system, which is incapable of representing false beliefs, can nevertheless correctly predict the actions of an agent who holds a false belief about the *location* or *properties* of an object. Next, we describe some of the signature limits of the early-developing system, which include an inability to handle situations where an agent holds a false beliefs about the *identity* of an object. According to the minimalist account, "mistakes about the identities of objects can be used to distinguish minimal from full-blown theory-of-mind cognition" (Butterfill & Apperly, 2013, p. 622); only the late-developing system has the representational capability to correctly predict the actions of an agent who mistakes one object for another. We then review previous evidence that infants can reason about false beliefs about identity, which proponents of the minimalist account argue is open to an alternative interpretation that implicates only the early-developing system. Finally, we introduce the present research, which sought to provide a new test of the minimalist account of early false-belief understanding. Instead of examining whether infants could reason about the actions of an agent who *held* a false belief about the identity of an object, here we asked whether infants could reason about the deceptive actions of an agent who *sought to implant in another agent* a false belief about the identity of an object.

We reasoned that positive results in this new deception task would cast doubt on the claim that infants are equipped only with a minimal, early-developing system that is incapable of representing false beliefs about identity (or indeed any other false beliefs). In addition, the demonstration that infants not only can reason about the actions of an agent who mistakes one object for another, but also can understand a deceptive agent's efforts to lure another agent into making such a mistake, would provide new evidence for the mentalistic claim that false-belief understanding emerges early in human development (e.g., Baillargeon, Scott, & He, 2010; Baillargeon et al., 2015, in press; Barrett et al., 2013; Scott et al., in press).

2. The minimalist account

2.1. How does the early-developing system enable success at non-traditional false-belief tasks?

According to the minimalist account, the early-developing system does not represent beliefs as such, but instead tracks simpler, belief-like "registrations" (Butterfill & Apperly, 2013). A registration is a relation between an agent and a specific object: upon encountering an object, an agent registers its location and properties. Registrations can be used to predict future actions: for example, the early-developing system will expect an agent who is searching for an object to look for it in the location where the agent last registered it.

After an agent registers an object's location and properties, this registration can become outdated if in the agent's absence the object is moved to another location or its properties are altered (Butterfill & Apperly, 2013). This poses no difficulties for the early-developing system, which can use registrations to predict actions whether these registrations are true or false. For example, consider a task in which an agent hides an object in one location and then leaves; in the agent's absence, the object is transferred to another location (e.g., Onishi & Baillargeon, 2005) or removed from the scene (e.g., Southgate et al., 2007). By tracking where the agent last registered the object, the early-developing system can predict that the agent, upon returning to the scene, will search for the object in its original (as opposed to current) location. As another example, consider a false-belief task in which an agent watches an experimenter demonstrate that a green object rattles when shaken, whereas a red object does not (Scott et al., 2010). Next, in the agent's absence, the experimenter alters the two objects (i.e., transfers the contents of the green object to the red object), so that the red object now rattles when shaken but the green object no longer does. By tracking what information the agent registered about each object's properties, the early-developing system can predict that the agent, upon returning to the scene, will select the (now silent) green object when asked to produce a rattling noise.

In sum, because the early-developing system predicts agents' actions by considering whatever true or false information is available to them about objects' locations and properties (including contents), it is sufficient to explain infants' success at nearly all non-traditional false-belief tasks published to date (e.g., Buttelmann, Over, Carpenter, & Tomasello, 2014; Knudsen & Liszkowski, 2012; Senju, Southgate, Snape, Leonard, & Csibra, 2011; Song, Onishi, Baillargeon, & Fisher, 2008; Surian et al., 2007; Träuble, Marinović, & Pauen, 2010). We return to possible exceptions in Section 3, after we discuss some of the signature limits that are thought to characterize the early-developing system.

2.2. What are some of the signature limits of the early-developing system?

2.2.1. Understanding false beliefs about identity

Because the early-developing system tracks registrations instead of representing beliefs, one of its signature limits concerns false beliefs that involve "the particular way in which an agent sees an object" (Low & Watts, 2013, p. 308), such as false beliefs about identity. In principle, genuine belief representations can capture any propositional content that agents can entertain, including false beliefs about the locations, properties, or identities of objects in a scene. In contrast, registrations can only capture relations between agents and specific objects—they do not "allow for a distinction between what is represented and how it is represented" (Apperly & Butterfill, 2009, p. 963). Thus, when an agent and an infant both view the same object but hold different beliefs about what the object *is*, the early-developing system is unable to correctly predict the agent's actions.

To illustrate, consider a scene (described by Butterfill & Apperly, 2013) in which an infant sits opposite an agent with a screen between them; two identical balls rest on the infant's side of the screen, occluded from the agent's view. One ball emerges to the left of the screen and returns behind it, and then the second ball emerges to the right of the screen and leaves the scene. Adults would expect the agent to hold a false belief about the identity of the second ball: the late-developing system would appreciate that the agent is likely to falsely represent the second ball as the first ball. In contrast, infants should expect the agent to treat the two balls as distinct objects: because the early-developing system cannot take into account *how* the agent might represent the second ball, it will simply track the agent's registration of each specific ball as it comes into view. Thus, after the second ball leaves the scene, adults should view it as unexpected if the agent searched behind the screen for the first ball, but infants should not.

To restate this first signature limit in more general terms, when an agent encounters a specific object *x*, the early-developing system can track the agent's registration of the location and properties of *x*, and it can use this registration to predict the agent's subsequent actions, even if its contents become false through events that occur in the agent's absence. If the agent next encountered another object *y*, the early-developing system could again track the agent's registration of *y*—but it would have no way of representing a situation where the agent mistook *y* for *x*. Because a registration relates to a specific object, it is not possible for the registration of *y* to be about *x*: the registration of *y* must be about *y*, just as the registration of *x* must be about *x*. Only the late-developing system, which is capable of representing false beliefs and other counterfactual states, could understand that the agent held a false belief about the identity of *y* and saw it as *x* even though it was really *y*.

2.2.2. Understanding complex goals

A second signature limit of the early-developing system is that, just as it tracks registrations rather than represents beliefs, it tracks goals in simple functional terms, as outcomes brought about by bodily movements (Butterfill & Apperly, 2013). In this respect, the minimalist account is similar to the non-mentalistic teleological account proposed by Csibra, Gergely, and their colleagues, which assumes that early psychological reasoning deals exclusively with physical variables: a teleological explanation specifies only the layout of a scene (e.g., the presence and location of obstacles), the agent's actions in the scene, and the physical end-state brought about by these actions (e.g., Csibra, Gergely, Bíró, Koós, & Brockbank, 1999; Gergely & Csibra, 2003; Gergely, Nádasdy, Csibra, & Bíró, 1995). From a minimalist perspective, infants should be able to track a variety of object-directed goals (e.g., carrying, grasping, shaking, storing, throwing, or stealing objects), but should be unable to understand more complex goals, such as goals that reference others' mental states. In particular, it should be difficult for the early-developing system to understand acts of strategic deception aimed at implanting false beliefs in others. Attributing goals that involve anticipating and manipulating the contents of others' mental states should be well beyond the purview of a system that "has only a minimal grasp of goal-directed action" and tracks goals as physical end-states brought about by bodily movements (Butterfill & Apperly, 2013, p. 614).

2.2.3. Reasoning about complex interactions among mental states

Finally, a third signature limit of the early-developing system is that it cannot deal with cognitively demanding situations in which predicting an agent's actions requires reasoning about a complex, interlocking set of mental states that interact causally (Low et al., 2014). According to the minimalist account, such a complex causal structure "places demands on working memory, attention, and executive function that are incompatible with automaticity" (Butterfill & Apperly, 2013, p. 629). The early-developing system is largely encapsulated from other cognitive processes, which makes it automatic and fast—but also inflexible and limited in the number of interlocking mental states it can successfully consider and integrate.

3. Previous tests of the minimalist account: Can infants attribute false beliefs about identity?

3.1. Initial tests

Is it the case that infants can track *whether* an agent sees an object, but not *how* an agent sees an object, so that they are unable to correctly predict the actions of an agent who holds a false belief about the identity of an object? Two initial reports provided suggestive evidence that infants in the 2nd year of life can attribute false beliefs about identity to agents (Scott & Baillargeon, 2009; Song & Baillargeon, 2008).

In Song and Baillargeon (2008), 14-month-olds first received familiarization trials in which a female agent sat centered behind two toys: a doll with blue pigtails and a stuffed skunk with a pink bow. Across trials, an experimenter's gloved hands placed the toys on placemats or in shallow containers; which toy was on the left and which toy was on the right varied across trials. In each trial, the agent reached for the skunk, suggesting that she preferred it over the doll. In the next, box-orientation trial, the agent was absent; two large boxes with lids rested on the apparatus floor, and the gloved hands demonstrated that the right box's lid had a tuft of blue hair (similar to one of the doll's pigtails) attached to it. At the start of the test trial, the agent was again absent; the experimenter hid the doll in the plain box and the skunk in the hair box. The agent then returned, reached for either the plain or the hair box, and then paused. The infants expected the agent to reach for the plain box and looked reliably longer when she reached for the hair box instead (this looking pattern reversed if the agent witnessed the gloved hands' actions). These results suggested that the infants expected the agent (1) to mistake the tuft of hair for one of the doll's pigtails and hence (2) to falsely conclude that the doll was hidden in the hair box and the skunk in the plain box (because both toys were always present in the familiarization trials).

In Scott and Baillargeon (2009), 18-month-olds first received familiarization trials in which a female agent sat centered behind a 1-piece penguin that did not come apart and a disassembled 2-piece penguin. Across trials, gloved hands placed the 1-piece penguin and the two pieces of the disassembled 2-piece penguin on platforms or in shallow containers; which toy was on the left and which toy was on the right varied across trials. In each trial, the agent hid a small key in the bottom piece of the 2-piece penguin and then assembled it; once assembled, the 2-piece penguin was identical to the 1-piece penguin. In the test trials, the agent was initially absent; the gloved hands assembled the 2-piece penguin, placed it under a transparent cover, and then placed the 1-piece penguin under an opaque cover. The agent then returned with her key, reached for either the transparent or the opaque cover, and then paused. The infants expected the agent to reach for the opaque cover and looked reliably longer when she reached for the transparent cover instead (this looking pattern reversed if the agent witnessed the gloved hands' actions). These results suggested that the infants expected the agent (a) to mistake the penguin visible under the transparent cover for the 1-piece penguin (because the 2-piece penguin had always been disassembled at the start of the familiarization trials) and hence (b) to falsely conclude that the disassembled 2-piece penguin was hidden under the opaque cover (because both penguins were always present in the familiarization trials).

3.2. The object-type interpretation

The results from these two experiments would seem to indicate that contrary to the minimalist account, infants can take into account how agents construe objects and understand that agents may hold false beliefs about identity. Butterfill and Apperly (2013) and Low and Watts (2013) have questioned this conclusion, however, on the grounds that in each experiment infants' reasoning could have involved expectations about *object types* as opposed to object identities (see also Low et al., 2014; Zawidzki, 2011). Specifically, the infants in the experiment of Song and Baillargeon (2008) might have reasoned as follows: at the start of each familiarization trial, the agent registered the presence of two types of objects, a doll with blue pigtails and a toy skunk; when the agent entered the scene in the test trial, she expected these two types of objects to again be present; therefore, upon registering the blue tuft attached to the hair box, she expected to find the skunk in the plain box. Likewise, the infants in the experiment of Scott and Baillargeon (2009) might have reasoned that when the agent entered the scene in each test trial, she expected two types of objects to again be present, an assembled penguin and a disassembled penguin; therefore, upon registering the assembled penguin under the transparent cover, she expected to find the disassembled penguin under the opaque cover.

Thus, because in both experiments infants' reasoning could have focused simply on the types of objects the agent expected to be present, neither experiment unequivocally contradicts the minimalist account of early false-belief understanding and more specifically the claim that infants are equipped only with an early-developing system that is incapable of handling false beliefs about identity. Instead, what these two experiments indicate is that the early-developing system can "predict actions on the basis of how things appear to observers who are ignorant of their true nature" (Butterfill & Apperly, 2013, p. 624).

This object-type interpretation is puzzling. The claim that the early-developing system is capable of handling false beliefs about object types would seem to blur the critical line drawn by the minimalist account between registrations and representations. If a registration is a relation to a specific object, its location, and properties, then how could an agent who encounters an object register what type of object it *appears to be*, as opposed to what type of object it really *is*? If the registration of *x* must be about *x*, and the registration of *y* must be about *y*, then how could an agent who encounters a novel tuft of hair mistake it for a (previously registered) doll's pigtail? Or how could an agent who encounters an assembled 2-piece penguin mistake it for a (previously registered) 1-piece penguin?

3.3. A further test

Despite the fact that the object-type interpretation seems inconsistent with the purported nature of the early-developing system (for similar arguments, see Carruthers, in press; Christensen & Michael, in press; Thompson, 2014), this interpretation is often used to argue that the findings of Song and Baillargeon (2008) and Scott and Baillargeon (2009) can be reconciled with the minimalist account (e.g., Butterfill & Apperly, 2013; Low, in press; Low & Watts, 2013; Low et al., 2014; Zawidzki, 2011). In an attempt to circumvent these arguments, Buttelmann, Suhrke, and Buttelmann (2015) recently devised a novel task: they tested infants' ability to attribute to an agent a false belief about the identity of a *single* object that could be represented in two different ways.

In each of four trials, 18-month-olds and an agent encountered a deceptive object, such as an object that appeared to be a toy duck. The agent then left the room, and in her absence the infants learned the object's true identity (e.g., the duck was in fact a brush). The deceptive object was then placed on a high shelf. When the agent returned and reached vainly for the deceptive object, the infants were shown two test objects, one that matched the deceptive object's appearance (e.g., a toy duck) and one that matched its true identity (e.g., a brush), and they were asked to give the agent what she wanted. The infants tended to choose the test object that matched the deceptive object's appearance rather than the test object that matched its identity (this pattern reversed if the agent was present when the object's true identity was revealed). Buttelmann et al. concluded that the infants understood that that the agent held a false belief about the identity of the deceptive object (e.g., she thought it was a toy duck) and used this belief to decide which test object to retrieve for her.

Unfortunately, these results are also open to a possible object-type interpretation. During the first encounter with each deceptive object, both the agent *and* the infants were ignorant of its true nature. The infants' early-developing system would therefore have tracked that the agent registered the presence of a particular type of object (e.g., a toy duck). Because the agent was absent when the object's true identity was subsequently revealed (e.g., a brush), her registration of the object was not updated. Consequently, when the agent returned and reached for the deceptive object on the shelf, the infants could simply consult her non-updated or incomplete registration of the object to choose an appropriate test object for her.

Thus, if one accepts the (controversial) claim that the early-developing system can track what type of object a deceptive object will appear to be to a naive agent, neither the findings of Buttelmann et al. (2015) nor those of Scott and Baillargeon (2009) and Song and Baillargeon (2008) provide conclusive evidence against the minimalist account.

4. The present research

Are infants capable of representing false beliefs, as the mentalistic account suggests, or are they limited to tracking registrations, as the minimalist account suggests? The present research was designed to shed light on this debate, by seeking new evidence that infants can attribute false beliefs about identity. As we saw in the last section, arguments have been raised by proponents of the minimalist account that the existing evidence that infants understand false beliefs about identity could simply "reflect reasoning based on a restricted set of facts regarding what the actor has registered" (Low et al., 2014, p. 1522). To avoid these interpretive difficulties, we designed a new violation-of-expectation task involving a deception situation: we asked whether infants could reason about the deceptive actions of an agent who sought to implant in another agent a false belief about the identity of an object.

Specifically, 17-month-old infants watched a situation in which a thief attempted to secretly steal a desirable object during its owner's absence by substituting a less desirable object. To evade detection, the thief had to select a substitute object that the returning owner could mistake for the desirable object she had left behind; in other words, the thief could elude detection only by implanting in the owner a false belief about the identity of the substitute object. Across experiments, infants had to determine which substitute object could deceive the owner, and under what conditions this substitution was likely to prove effective.

We reasoned that positive results indicating that infants understood that the thief wanted the owner to mistake the substitute object for the object she had left there would cast doubt on the minimalist claim that one critical signature limit of early psychological reasoning is an inability to reason about false beliefs about identity. In addition, because of the challenging task used here, positive results would also bear on the other two signature limits discussed earlier: it would be difficult to explain how a minimal system incapable of tracking complex goals or of processing multiple, interlocking mental states could understand an act of strategic deception aimed at implanting a false belief in another agent.

5. Experiment 1

5.1. Design

The infants in Experiment 1 were assigned to a deception or a silent-control condition. We first describe these conditions and then outline the predictions from the mentalistic and minimalist accounts.

The *deception* condition examined whether 17-month-olds could distinguish between an effective and an ineffective act of deception. The infants watched live events involving two female agents, the thief (T) and the owner (O). During O's absence, T stole O's rattling toy and replaced it with a silent toy that was either visually identical (effective deception) or visually distinct (ineffective deception).

The infants received six familiarization trials, which included three rattling-toy trials and three silent-toy trials (Fig. 1). A different toy was used in each trial; the six toys differed only in color and pattern. All trials had an initial phase and a final phase. At the start of the (36-s) initial phase of each *rattling-toy* trial, T sat at the back of a puppet-stage apparatus, and O knelt (out of view) behind a curtained window in the right wall. O knocked twice, opened the curtain, and brought in a toy on a tray. O then shook the toy, causing it to rattle, until a bell rang; O said, "I'll be back!", returned the toy to the tray, and left, closing the curtain. Next, T grasped the toy and shook it, causing it to rattle, until O knocked again, signaling her return; T then quickly put the toy back on the tray. O opened the curtain, picked up the toy, and stored it inside a lidded box next to her window (the toy did not rattle when moved, only when shaken). Both experimenters then paused. During the final phase of the trial, the infants watched this paused scene until the trial ended. The *silent-toy* trials were identical except that the toy made no noise when O shook it, T did not play with the toy during O's absence, and upon her return O threw the toy into a trashcan located across the apparatus, near the left wall (to muffle noises, the trashcan was filled with fabric and discarded toys were removed after each trial).

Fig. 1. Schematic depiction of the events presented in the familiarization trials in Experiment 1. Infants saw three rattling-toy trials and three silent-toy trials; a different toy was used in each trial, and the six toys differed in color and pattern.

Next, the infants received either a matching or a non-matching test trial (Fig. 2). During the (27-s) initial phase of the *matching* trial, while T watched, O brought in a rattling test toy that was visually identical to a silent toy she had previously discarded in the trashcan. O shook the test toy, causing it to rattle, until the bell rang; she then said, "I'll be back!", returned the test toy to the tray, and left. T picked up the test toy, peered into the trashcan, selected the matching silent toy, and placed it on the tray. Next, T hid the test toy in a kangaroo pocket on the front of her shirt and then paused (the toy fell to the bottom of T's pocket and was not visible above the apparatus floor). During the final phase, the infants watched this paused scene until the trial ended (O did not return in the test trial: because our focus was on infants' responses to T's deceptive actions, the test scene paused after these actions). The *non-matching* trial was identical except that the silent toy T retrieved from the trashcan and placed on the tray differed in color from the rattling test toy. For half the infants, the rattling test

Fig. 2. Schematic depiction of the events presented in the test trials of the deception condition of Experiment 1. Infants in the *silent-control* condition saw identical test trials except that the test toy O brought in was silent.

toy was green, the matching silent toy was green, and the non-matching silent toy was yellow; for the other infants, the rattling test toy was yellow, and the matching and non-matching silent toys were reversed.

The *silent-control* condition was identical to the deception condition except that in the test trial O brought in a *silent* test toy.

5.2. Predictions

5.2.1. Mentalistic account

According to the mentalistic account, the infants in the deception condition (a) should realize that only the substitution of the matching silent toy was consistent with T's deceptive goal of stealing the rattling test toy without O's notice and hence (b) should look reliably longer if given the non-matching as opposed to the matching trial. Although these trials were complex, they combined elements that, according to prior research, infants in the 2nd year of life are already able to interpret.

First, the familiarization trials provided information that T preferred the rattling toys over the silent toys: across trials, T consistently played with the rattling toys but ignored the silent toys. Prior research indicates that when an agent selectively acts on one type of object as opposed to another (e.g., toy ducks as opposed to toy frogs; red objects as opposed to objects of other colors), infants in the 2nd year of life take this consistent choice information to reveal an underlying preference (e.g., Kushnir, Xu, & Wellman, 2010; Luo & Beck, 2010; Woodward, 1999). Thus, it seemed likely that the infants in the deception condition would attribute to T a preference for the rattling toys.

Second, the familiarization trials also conveyed that T wanted to keep O ignorant about her (T's) interest in the rattling toys: in each rattling-toy trial, T picked up the toy only after O left, and she quickly returned it to the tray when O knocked to announce her return. Prior research indicates that infants in the 2nd year of life are adept at tracking which agents are knowledgeable or ignorant about events in a scene (e.g., Liszkowski, Carpenter, & Tomasello, 2008; Scott et al., 2010; Song et al., 2008; Tomasello & Haberl, 2003). Thus, the infants in the deception condition should realize that T consistently played with the rattling toys only during O's absence and hence without her knowledge.

Third, in the test trial, and for the first time in the testing session, O introduced a rattling toy that was visually identical to a silent toy she had previously discarded. After O left, T stole this rattling toy by hiding it in her pocket. Prior research indicates that infants in the 2nd year of life already understand stealing—or taking away the toy someone has been playing with—as a negative, antisocial action (e.g., Hamlin, Mahajan, Liberman, & Wynn, 2013; Hamlin, Wynn, Bloom, & Mahajan, 2011). The infants in the deception condition should therefore recognize that T meant to steal the rattling test toy when she hid it in her pocket.

Fourth, T did not merely steal the rattling test toy: she also placed one of the discarded silent toys on the tray, suggesting that she wanted her theft to go unnoticed by O (this was consistent with T's secretive behavior during the familiarization trials). By replacing the rattling test toy with the matching silent toy, T could accomplish her deceptive goal: when O returned, she would mistake the matching silent toy for the rattling toy she had left behind. As discussed earlier, prior research suggests that 14.5- to 18-month-olds may be able to attribute to an agent a false belief about the identity of an object (Buttelmann et al., 2015; Scott & Baillargeon, 2009; Song & Baillargeon, 2008). If 17-month-olds can appreciate not only the perspective of an agent who *holds* such a false belief, but also the perspective of an agent who *seeks to implant* such a false belief, then the infants in the deception condition should recognize that by substituting the matching silent toy, T wanted O to believe it was the rattling toy she had left behind.

To summarize, the mentalistic account predicted that the infants in the deception condition would build a causally coherent interpretation of T's actions that involved multiple, interlocking mental states: (a) T had a *preference* for the rattling toys; (b) when O introduced the rattling test toy, which was visually identical to a previously discarded silent toy, T formed the *goal* of secretly stealing the rattling test toy; (c) substituting the matching silent toy was consistent with T's deceptive goal, because O would hold a *false belief about the identity* of the substitute object; and (d) substituting the non-matching silent toy was inconsistent with T's deceptive goal, because O would *know* which toy it was as soon as she saw it.

Finally, the mentalistic account predicted that the infants in the silent-control condition would be unable to build a causally coherent interpretation of T's actions in either trial and hence would look about equally whether they received the non-matching or the matching trial. Prior research indicates that when infants are unable to generate an explanation for an agent's initial actions, they hold no expectation for the agent's subsequent actions (e.g., Csibra et al., 1999; Gergely et al., 1995; Woodward, 1999; Woodward & Sommerville, 2000). Because T had never expressed interest in the silent toys, her motivation for stealing the silent test toy was unclear; after all, T could have taken silent toys from the trashcan at any time in the familiarization trials. The infants should thus look equally whether T substituted the matching or the non-matching silent toy for the rattling test toy. Negative results in this condition would also rule out low-level interpretations of positive results in the deception condition (e.g., the infants merely attended to the color of the toy on the tray in the test trial and looked longer when it changed from green to yellow or vice versa; Heyes, 2014).

5.2.2. Minimalist account

According to the minimalist account, the infants in the deception condition should be unable to reason about T's deceptive actions and hence should look about equally whether they received the non-matching or the matching trial. From a minimalist perspective, the present task posed at least two difficulties for the early-developing system.

First, because the task focused on the actions of T (the thief) rather than those of O (the owner), and T was present throughout all trials and witnessed all events that occurred, the infants could not succeed simply by tracking what information T had or had not registered about the scene. Instead, the infants needed to take into account *T's reasoning about O's future registration* of the substitute toy. Since the early-developing system is unable to (a) track complex goals, such as deceptive goals that involve anticipating and manipulating others' mental states, or (b) process interactions among multiple, causally interlocking mental states, it seemed unlikely that the infants would be able to understand T's deceptive goal of implanting a false belief in O.

Second, even assuming such understanding were somehow possible, there remained the difficulty that T had to anticipate *how* O would perceive the substitute toy. Because the early-developing system cannot handle false beliefs about identity, in the matching trial it should expect O to register the substitute toy as the silent matching toy it really was, even though it was visually identical to the rattling test toy. O could not register *y* (the silent matching toy on the tray) as *x* (the rattling test toy she had left there), any more than the agent in the hypothetical two-ball scene described by Butterfill and Apperly (2013) could register *y* (the second, visually identical ball to emerge from the screen) as *x* (the first ball to emerge into view). Since neither the substitution in the matching trial nor that in the non-matching trial could deceive O, it did not matter which silent toy T placed on the tray, and the infants should look equally at either substitution.

Could the early-developing system predict that T would expect O to mistake the silent matching toy for the rattling test toy by considering what type of object the toy on the tray would *appear* to be to O? By design, an object-type interpretation similar to the one offered for the findings of Song and Baillargeon (2008) and Scott and Baillargeon (2009) was not possible in the present experiment, because the two toys were visually identical so that there were no visual cues that could lead O to register the toy on the tray as a rattling type of toy. Moreover, an alternative object-type interpretation similar to the one offered for the findings of Buttelmann et al. (2015) was also not possible, because the rattling test toy did not change in O's absence: it remained a rattling type of toy when hidden in T's pocket.

5.3. Method

5.3.1. Participants

Participants were 36 healthy full-term infants, 18 male (16 months, 27 days to 18 months, 3 days, *M* = 17 months, 14 days). Another 3 infants were excluded because they were inattentive (1) or active (1), or had a test looking time over 3 standard deviations from the mean of the condition (1). Equal numbers of infants were randomly assigned to each combination of condition (deception, silent-control) and test trial (matching, non-matching).

The infants' names in this and the following experiments were obtained from a university-maintained database of parents interested in participating in child-development research. Parents were offered reimbursement for their travel expenses but were not compensated for their participation. Each infant's parent gave written informed consent, and the protocol was approved by the Institutional Review Board at the University of Illinois at Urbana-Champaign.

5.3.2. Apparatus and stimuli

The apparatus consisted of a brightly lit display booth (201 cm high × 102 cm wide × 57 cm deep) with a large opening (46 cm × 95 cm) in its front wall; between trials, a supervisor lowered a curtain in front of this opening. Inside the apparatus, the walls were white, and the back wall and floor were covered with blue adhesive paper.

T wore a gray sweatshirt with a large (38 cm × 38 cm) kangaroo pocket on the front and sat behind a window (55.5 cm × 77 cm) in the back wall of the apparatus; a screen behind T hid the testing room. O wore a green shirt and knelt at a window (51 cm × 38 cm) in the right wall of the apparatus; this window was covered with a muslin curtain that could be drawn aside. Prior to the testing session, both agents introduced themselves to the infant; during the testing session, however, the agents did not make eye contact with the infant: they looked at the objects they acted on, followed the actions performed by the other agent, and otherwise kept their eyes on a neutral point on the apparatus floor.

Stimuli included ten cylindrical toys (each 9 cm high × 5.5 cm in diameter). The five rattling toys contained a marble and were built to rattle only when briskly shaken; the five silent toys were empty. The rattling toys were covered in one of five different adhesive papers: blue, black-and-white cow-patterned, green–blue-and-white striped, green, and yellow. The silent toys were covered in one of three different adhesive papers: green-and-white marble-patterned, yellow, and green. Stimuli also included a gray tray (2.3 cm × 12 cm × 15 cm), a red box (11.5 cm × 14.5 cm × 14.5 cm) with a lid (17.5 cm × 14.5 cm) decorated with dots, and a gray trashcan (25.5 cm × 20 cm × 22.5 cm).

During each testing session, a metronome beat softly to help the agents adhere to the events' second-by-second scripts. A camera captured an image of the events, and another camera captured an image of the infant. The two images were combined, projected onto a television set located behind the apparatus, and monitored by the supervisor to confirm that the events followed the prescribed scripts. Recorded sessions were also checked offline for accuracy.

5.3.3. Procedure

Infants sat on a parent's lap centered in front of the apparatus; parents were instructed to remain silent and close their eyes during the test trial.

Two naïve observers hidden on either side of the apparatus monitored each infant's looking behavior. Looking times during the initial and final phases of each trial were computed separately using the primary observer's responses. Interobserver agreement was measured for 101/108 infants in this report (only one observer was present for the other infants) and averaged 93% per trial per infant.

The six familiarization trials were administered in the following order: rattling (blue), silent (marble-patterned), silent (yellow), rattling (cow-patterned), silent (green), and rattling (striped). Infants were highly attentive during the initial phases of the trials; they looked, on average, for 97% of each initial phase. A similar high level of attention (95% of each initial phase) occurred in the two silent-toy familiarization trials involving the yellow and green toys, which served as the substitute toys in the test trial; thus, it seemed likely that infants knew both toys were in the trashcan. The final phase of each familiarization trial ended when the infant (a) looked away for 2 consecutive seconds after having looked for at least 5 cumulative seconds or (b) looked for a maximum of 60 cumulative seconds. Infants looked equally during the final phases of the rattling-toy ($M = 19.6$, $SD = 11.6$) and silent-toy ($M = 19.2$, $SD = 9.9$) familiarization trials, $t < 1$, indicating that they were attentive to both trial types.

Infants were highly attentive during the initial phase of the test trial; across conditions and trials, they looked, on average, for 98% of the initial phase. The final phase of the test trial ended when the infant (a) looked away for 1 consecutive second after having looked for at least 5 cumulative seconds or (b) looked for a maximum of 30 cumulative seconds.

Preliminary analyses of all test data in this report revealed no interactions of condition and trial with infants' sex or color of the test toy (green, yellow), all Fs < 1; the data were therefore collapsed across the latter two factors in subsequent analyses.

5.4. Results

The infants' looking times during the final phase of the test trial (Fig. 3) were analyzed using an analysis of variance (ANOVA) with condition (deception, silent-control) and trial (matching, non-matching) as between-subjects factors. The analysis yielded a significant main effect of condition,

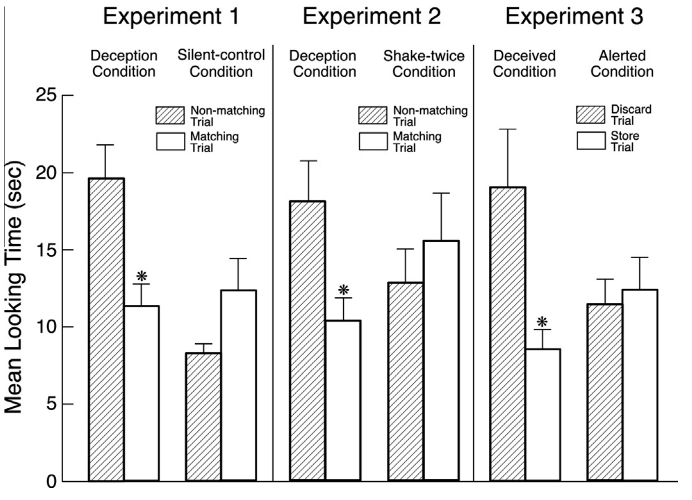

Fig. 3. Results from Experiments 1–3. Mean looking time (s) of the infants during the final phase of the test trial as a function of condition and trial. Errors bars represent standard errors, and an asterisk denotes a significant difference between the events within a condition ($p < .05$ or better).

$F(1, 32) = 9.15$, $p = .005$, and a significant Condition × Trial interaction, $F(1, 32) = 12.74$, $p = .001$. Planned comparisons revealed that in the deception condition, the infants who received the non-matching trial ($M = 19.6$, $SD = 6.7$) looked reliably longer than those who received the matching trial ($M = 11.3$, $SD = 4.3$), $F(1, 32) = 11.73$, $p = .002$, Cohen's $d = 1.48$; in the silent-control condition, the infants looked about equally whether they received the non-matching ($M = 8.3$, $SD = 1.93$) or the matching ($M = 12.3$, $SD = 6.2$) trial, $F(1, 32) = 2.64$, $p = .114$, $d = .85$. An analysis of covariance (ANCOVA) using as covariates the infants' averaged looking times during the final phases of the rattling-toy and silent-toy familiarization trials again revealed a significant Condition × Trial interaction $F(1, 30) = 10.20$, $p = .003$, and planned comparisons yielded equivalent results.

5.5. Discussion

In the deception condition, the infants who saw T replace the rattling test toy with a non-matching silent toy looked reliably longer than those who saw her substitute a matching silent toy. This result suggests that the infants realized that (a) T had the goal of stealing the rattling test toy without O's knowledge and (b) T could achieve this deceptive goal by substituting the matching but not the non-matching silent toy: only the visually identical, matching silent toy could be mistaken by O for the rattling test toy she had left behind. In the silent-control condition, where T had no clear motivation for stealing the silent test toy, the infants had no expectation about which silent toy she would place on the tray. This negative result also ruled out the low-level interpretation that the infants in the deception condition merely responded to the change in the color of the toy on the tray in the non-matching trial.

Together, the results of Experiment 1 suggested that 17-month-olds can reason about one agent's attempt to implant in another agent a false belief about the identity of an object. These results supported the mentalistic as opposed to the minimalist account of early false-belief understanding.

6. Experiment 2

Experiment 2 had three goals. The first was to confirm the main result of Experiment 1 that 17-month-olds can reason about one agent's attempt to lure another agent into holding a false belief about the identity of an object. The second goal was to further explore 17-month-olds' understanding of the causal factors that determine whether a deceptive act is likely to be effective. In Experiment 1, T could secretly steal the rattling test toy by substituting the matching silent toy because O never shook the toy on the tray after she returned. In Experiment 2, we asked whether infants would realize that if O *did* routinely shake the toy on the tray after she returned, it would no longer matter whether T substituted the non-matching toy (O would detect the substitution when she saw the toy) or the matching toy (O would detect the substitution when she shook the toy).

Finally, the third goal of Experiment 2 was to address a possible alternative interpretation of the results of Experiment 1. It might be suggested that the infants detected a statistical regularity in the familiarization trials: after playing with a rattling toy, T always returned to the tray a toy that was visually identical to the one she had picked up. Thus, the infants in the deception condition might have looked longer in the non-matching trial because T deviated from this regularity and returned to the tray a visually distinct toy. Similarly, the infants in the silent-control condition might have looked equally in the non-matching and matching trials because T had never picked up a silent toy before, so that both trials deviated from her previous actions. The design of Experiment 2 allowed us to examine this regularity-based interpretation.

The infants were assigned to a shake-twice or a deception condition; both conditions were identical to the deception condition of Experiment 1, except that the familiarization trials differed. In the *shake-twice* condition, whenever O returned, she shook the toy on the tray before storing it in her box (rattling-toy trials) or discarding it in the trashcan (silent-toy trials). To accommodate O's new actions, the initial phase of the familiarization trials was lengthened from 36 s to 39 s. In the *deception* condition, O did not shake the toy when she returned in the familiarization trials; instead, she simply held the toy for a few seconds before storing or discarding it.

As in Experiment 1, the infants in the deception condition should realize that substituting the matching silent toy would serve T's goal of secretly stealing the rattling test toy, but substituting the non-matching silent toy would not, because O would be able to detect this substitution as soon as she saw the toy. The infants should thus look reliably longer if given the non-matching as opposed to the matching trial, as in Experiment 1. In contrast, the infants in the shake-twice condition should realize that neither silent toy could be effective in deceiving O, because she would be able to detect the substitution either when she saw the toy (non-matching trial) or when she shook the toy (matching trial). The infants should thus have no specific expectation about which silent toy T would place on the tray, and they should therefore look about equally whether they received the non-matching or the matching trial.

The shake-twice condition also addressed the regularity-based interpretation raised above. T performed exactly the same actions in the shake-twice condition as she did in the deception conditions of Experiments 1 and 2—only O's actions differed across conditions. If the infants in the deception conditions looked longer at the non-matching trial because T's actions deviated from those she had produced in the familiarization trials, then the infants in the shake-twice condition should do the same: they should look longer if they received the non-matching as opposed to the matching trial. Evidence that these infants instead looked equally whether they received the non-matching or the matching trial would thus rule out the regularity-based interpretation and support a richer interpretation of the results of the deception conditions.

6.1. Method

6.1.1. Participants

Participants were 36 healthy term infants, 18 male (16 months, 27 days to 18 months, 3 days, M = 17 months, 16 days). Another 7 infants were excluded because they were fussy (5) or active (1), or had a test looking time over 3 standard deviations from the mean of the condition (1). Equal numbers of infants were randomly assigned to each combination of condition (deception, shake-twice) and test trial (matching, non-matching).

6.1.2. Apparatus and procedure

The apparatus and procedure were identical to those used in the deception condition of Experiment 1. The infants were highly attentive during the initial phases of the familiarization trials and looked, on average, for 98% of each initial phase (97% for the silent-toy trials involving the yellow and green toys). The infants again looked equally during the final phases of the rattling-toy ($M = 20.6$, $SD = 9.0$) and silent-toy ($M = 20.5$, $SD = 10.3$) familiarization trials, $t < 1$, indicating that they were attentive to both trial types. Finally, the infants were highly attentive during the initial phase of the test trial and looked, on average, for 97% of the initial phase.

6.2. Results

The infants' looking times during the final phase of the test trial (Fig. 3) were analyzed using an ANOVA with condition (deception, shake-twice) and trial (matching, non-matching) as between-subjects factors. The analysis yielded only a significant Condition × Trial interaction, $F(1, 32) = 4.73$, $p = .037$. Planned comparisons revealed that in the deception condition, the infants who received the non-matching trial ($M = 18.3$, $SD = 7.8$) looked reliably longer than those who received the matching trial ($M = 10.5$, $SD = 4.4$), $F(1, 32) = 5.21$, $p = .029$, $d = 1.23$; in the shake-twice condition, the infants looked equally whether they received the non-matching ($M = 13.0$, $SD = 6.7$) or the matching ($M = 15.7$, $SD = 9.2$) trial, $F < 1$. As in Experiment 1, an ANCOVA revealed a significant Condition × Trial interaction, $F(1, 30) = 4.28$, $p = .047$, and planned comparisons yielded equivalent results.

6.3. Combined analyses of Experiments 1 and 2

In additional analyses, we combined the data from Experiments 1 and 2 in order to create a larger sample and compare the results of the two deception conditions ($n = 36$) to those of the two control conditions (silent-control and shake-twice, $n = 36$). The data were analyzed using an ANOVA with condition (combined-deception, combined-control) and trial (matching, non-matching) as between-subjects factors. The analysis yielded a marginal effect of condition, $F(1, 68) = 3.05$, $p = .085$, and a significant Condition × Trial interaction, $F(1, 68) = 14.703$, $p < .001$. Planned comparisons indicated that the infants in the combined-deception condition looked reliably longer if given the non-matching trial ($M = 18.9$, $SD = 7.1$) as opposed to the matching trial ($M = 10.9$, $SD = 4.2$), $F(1, 68) = 14.75$, $p < .001$, $d = 1.38$, whereas the infants in the combined-control condition looked about equally at the non-matching ($M = 10.7$, $SD = 5.3$) and matching ($M = 14.0$, $SD = 7.8$) trials, $F(1, 68) = 2.51$, $p = .12$, $d = .49$. Non-parametric Wilcoxon sum-rank tests confirmed the results of the combined-deception ($W = 226$, $p < .001$) and combined-control ($W = 294.5$, $p = .229$) conditions.

Finally, we also examined infants' responses in each trial across conditions. A planned comparison focusing on the non-matching trial revealed that the infants in the combined-deception condition ($M = 18.9$, $SD = 7.1$) looked reliably longer than did those in the combined-control condition, ($M = 10.7$, $SD = 5.3$), $F(1, 68) = 15.57$, $p < .001$, $d = 1.32$. In contrast, a planned comparison focusing on the matching trial revealed no reliable difference between the responses of the infants in the combined-deception ($M = 10.9$, $SD = 4.2$) and combined-control ($M = 14.0$, $SD = 7.8$) conditions, $F(1, 68) = 2.19$, $p = .14$, $d = .49$.

6.4. Discussion

The positive result of the deception condition in Experiment 2 replicated that from the deception condition in Experiment 1: the infants attributed to T the goal of stealing the rattling test toy without O's knowledge, and they understood that T could do so by substituting the matching but not the non-matching silent toy. In contrast, the infants in the shake-twice condition had no expectation about which silent toy T would place on the tray, because neither toy could deceive O: she would be able to detect the substitution of the non-matching toy when she saw it, and she would be able to detect the substitution of the matching toy when she shook it. This negative result also ruled out the possibility that the infants in the deception conditions of Experiments 1 and 2 looked longer in the non-matching trial merely because T deviated from her previous actions by placing a visually distinct toy on the tray. T performed exactly the same actions in the deception and shake-twice conditions, and yet these conditions yielded reliably different results.

Together, the results of Experiments 1 and 2 indicated that the infants understood that T could lure O into mistaking the silent toy on the tray for the rattling toy she had left behind only if (a) the silent toy was visually identical to the rattling toy (Experiments 1 and 2) and (b) O did not routinely shake her toy when she returned (Experiment 2). These results supported the mentalistic account of early false-belief understanding, but cast doubt on the minimalist account.

6.4.1. Consistency or efficiency violations?

One other facet of the results of Experiments 1 and 2 deserves mention. Csibra and Gergely proposed that early psychological reasoning is constrained by a principle of rationality (e.g., Csibra et al., 1999; Gergely & Csibra, 2003; Gergely et al., 1995; see also Dennett, 1987), and in their work with infants they focused mainly on one corollary of this principle, *efficiency*: agents should expend as little effort as possible to achieve their goals (see also Scott & Baillargeon, 2013). Baillargeon and her colleagues recently proposed that many findings in the early psychological-reasoning literature can be taken to demonstrate infants' sensitivity to another corollary of the rationality principle, *consistency*: agents should act in a manner consistent with their mental states (e.g., Baillargeon et al., 2015, in press). Up to this point, we have offered a consistency-based interpretation of the positive results of the deception conditions: the infants viewed T's actions in the non-matching trial as inconsistent with her goal of secretly stealing the rattling test toy. However, an efficiency-based interpretation could also be offered for these results: the infants viewed T's actions in the non-matching trial as an inefficient means of achieving her goal.

Although we acknowledge that the infants in the deception conditions could have detected either a consistency or an efficiency violation in the non-matching trial, we prefer the former description because (a) the infants in Experiments 1 and 2 understood at least some of the causal conditions under which T's actions could deceive O, and (b) the infants in the deception conditions, in particular, recognized that T's substitution of the non-matching silent toy could not deceive O. Therefore, it seemed more intuitive to describe this substitution as inconsistent with T's goal, rather than as merely inefficient (i.e., substituting a silent green toy for a yellow rattling toy is not just an inefficient means of secretly stealing the rattling toy, it is inconsistent with this deceptive goal).

7. Experiment 3

Experiment 3 had two goals. The first was to demonstrate that infants would expect O to be deceived if she returned after T stole the rattling test toy and substituted the matching silent toy on the tray. According to the mentalistic account, which holds that an abstract capacity for false-belief understanding emerges early in development, infants should be able to understand both T's deceptive substitution of the silent matching toy and O's mistaken belief that this toy was the rattling toy she had left behind. The second goal was to further explore infants' understanding of the circumstances under which O could or could not be deceived, by asking whether infants would realize that O could be deceived by the substitution of the silent matching toy only if she did not witness this substitution.

The infants were assigned to a deceived or an alerted condition. In both conditions, the infants received the same familiarization trials as in Experiment 1; only the test trial differed. In the *deceived* condition, the 36-s initial phase of the test trial began just like that of the matching trial in the deception condition of Experiment 1 (Fig. 4): O brought in the rattling test toy, shook it, and left; in her absence, T substituted the matching silent toy on the tray and hid the rattling test toy in her pocket. After T completed these actions, the initial phase continued: while T watched, O knocked, opened the curtain, picked up the toy on the tray, and either stored it in her box (*store* trial) or discarded it in the trashcan (*discard* trial).

From a mentalistic perspective, as mentioned earlier, infants should expect O to mistake the matching silent toy on the tray for the (visually identical) rattling test toy she had left there. Infants should thus expect O to store the matching silent toy alongside her rattling toys, and they should detect a violation when she discarded it instead (even though this was precisely how she had acted on it before). Infants should thus look reliably longer if given the discard as opposed to the store trial.

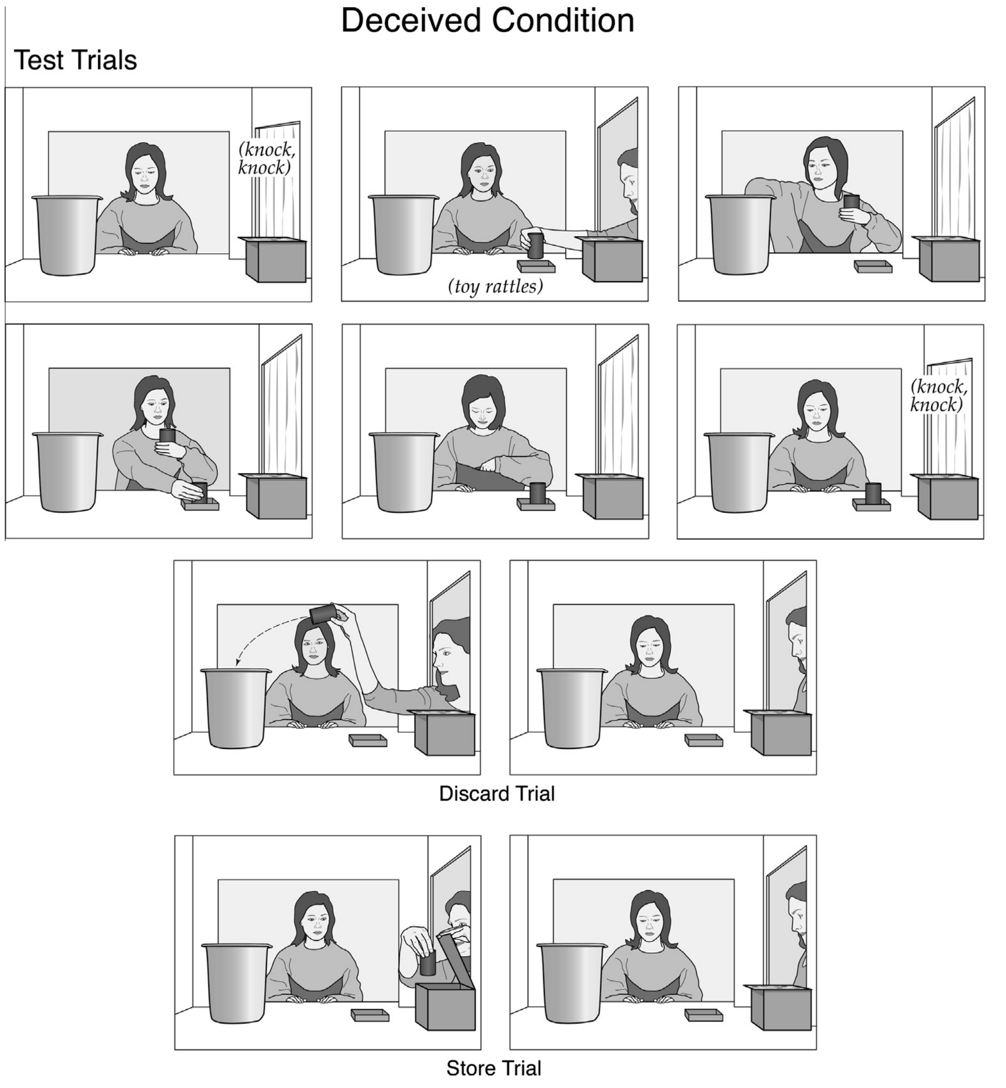

Fig. 4. Schematic depiction of the events presented in the test trials of the deceived condition of Experiment 3. Infants in the *alerted* condition saw similar test trials except that O returned earlier: she caught T with a visually identical toy in each hand and watched as T placed one toy (the silent toy) on the tray and the other toy (the rattling toy) in her pocket.

From a minimalist perspective, however, the opposite prediction held. In the test trial, the early-developing system could reason that O had registered the matching silent toy in the trashcan and the rattling test toy on the tray, but had not registered T's substitution of the matching silent toy for the rattling test toy. However, the early-developing system could not take into account *how* O was likely to construe the toy on the tray: false beliefs about identity fall beyond the purview of this system. Thus, when O returned in the test trial, infants should expect her to register the toy on the tray for what it really *was*, the matching silent toy. Because O always discarded the silent toys (in fact, she had previously discarded that very same toy), infants should expect her to throw the toy in the trashcan. Infants should therefore look reliably longer at the store than at the discard trial.

The *alerted* condition was identical to the deceived condition except that O returned 3 s earlier in the test trial and caught T with a visually identical toy in each hand. O then watched as T placed one toy (the matching silent toy) on the tray and the other toy (the rattling test toy) in her pocket. According to the mentalistic account, the infants in the alerted condition should respond differently from those in the deceived condition: because O could not know which toy was on the tray, the infants in the alerted condition should have no expectation about her actions on the toy, and they should therefore look about equally whether they received the discard or the store trial. In contrast, the minimalist account predicted that the infants in the alerted condition should respond similarly to those in the deceived condition: in either case, the infants should expect O to register the toy on the tray as the silent toy, and hence they should look reliably longer if they received the store as opposed to the discard trial.

If negative results were obtained in the alerted condition, as predicted by the mentalistic account, this would also address a possible alternative interpretation of positive results in the deceived condition. Perhaps the infants in this condition detected a statistical regularity in the familiarization trials—O always stored toys following rattling—and thus looked longer in the discard trial because it deviated from this regularity: O discarded the toy on the tray even though the last toy she had manipulated rattled. Because O performed exactly the same actions on the toys in the deceived and alerted conditions, evidence that the infants in the latter condition looked equally at the discard and store trials would rule out this regularity-based interpretation.

7.1. Method

7.1.1. Participants

Participants were 36 healthy full-term infants, 19 male (16 months, 26 days to 18 months, 5 days, $M = 17$ months, 12 days). Another 5 infants were excluded because they were inattentive (3), looked the maximum time allotted in the familiarization and test trials (1), or had a test looking time over 3 standard deviations from the mean of the condition (1). Equal numbers of infants were randomly assigned to each combination of condition (deceived, alerted) and test trial (store, discard).

7.1.2. Apparatus and procedure

The apparatus and procedure were identical to those used in the deception condition of Experiment 1, with one exception: the final phase of the test trial ended when the infant (a) looked away for 1.5 consecutive seconds (as opposed to 1 consecutive s) after having looked for at least 5 cumulative seconds or (b) looked for a maximum of 30 cumulative seconds. The initial phase of the test trial in Experiment 3 was longer than that in Experiment 1 (36 s vs. 27 s) and required infants to reason about both T's deceptive actions and O's responses to these actions; a slightly longer look-away criterion allowed infants greater opportunity to process all of the events they had seen before the trial could end.

The infants were highly attentive during the initial phases of the familiarization trials and looked, on average, for 99% of each initial phase (98% for the silent-toy trials involving the yellow and green toys). The infants again looked about equally during the final phases of the rattling-toy ($M = 21.5$, $SD = 8.3$) and silent-toy ($M = 19.6$, $SD = 9.2$) familiarization trials, $t(35) = 1.34$, $p = .19$, indicating that they were attentive to both trial types. Finally, the infants were highly attentive during the initial phase of the test trial and looked, on average, for 99% of the initial phase.

7.2. Results

The infants' looking times during the final phase of the test trial (Fig. 3) were analyzed using an ANOVA with condition (deceived, alerted) and trial (store, discard) as between-subjects factors. The analysis yielded a marginal effect of trial, $F(1, 32) = 4.02$, $p = .053$, and a significant Condition $\times$ Trial interaction, $F(1, 32) = 5.81$, $p = .022$. Planned comparisons revealed that in the deceived condition, the infants who received the discard trial ($M = 19.0$, $SD = 11.4$) looked reliably longer than those who received the store trial ($M = 8.5$, $SD = 3.9$), $F(1, 32) = 9.75$, $p = .004$, $d = 1.24$; in the alerted condition, the infants looked about equally whether they received the discard ($M = 11.4$, $SD = 4.7$) or the store (M = 12.4, SD = 6.2) trial, $F < 1$. An ANCOVA also revealed a significant Condition × Trial interaction, $F(1, 30) = 4.82$, $p = .036$, and planned comparisons yielded equivalent results.

7.3. Discussion

In the deceived condition, T completed her deceptive actions before O returned, and the infants expected O to mistake the matching silent toy on the tray for the rattling toy she had left there. The infants therefore expected O to store the toy and detected a violation when she discarded it instead. In the alerted condition, O caught T in the act, and the infants realized that O could not know whether the toy on the tray was the matching silent toy or the rattling test toy. The infants thus tended to look equally whether O stored or discarded the toy. This negative result also ruled out the possibility that the infants in the deceived condition looked longer in the discard trial merely because T deviated from her previous actions by discarding a toy following rattling.

Together, the results of Experiment 3 indicated that the infants in the deceived condition expected O to hold a false belief about the identity of the matching silent toy on the tray. Could minimalist researchers offer an object-type alternative interpretation (as was discussed in Section 3.2) for these results? We think not. In the present experiments, there were no predictive visual cues distinguishing the rattling and silent toys: until O shook each toy, one could not know whether it would rattle or not. Thus, the infants could not have expected O to store the toy she found on the tray when she returned because misleading visual cues made it *appear* to be a rattling type of toy; they could only have expected her to store the toy because they understood that she was likely to mistake it for the visually identical rattling toy she had left there—in other words, because they attributed to her a false belief about the identity of the toy.

Perhaps another object-type interpretation might be suggested: O expected two types of toys to be present in the scene, a rattling type of toy on the tray and a silent type of toy in the trashcan, and her registrations of the toys' locations were not updated because these changed in her absence. Therefore, O should reach for the tray to retrieve the rattling type of toy she had placed there. Notice, however, that this interpretation essentially concedes that the early-developing system would predict that O would mistake the silent matching toy on the tray for the visually identical rattling toy she had left there, which is precisely what the minimalist account claims the early-developing system cannot do.

8. General discussion

The present results provide the first experimental demonstration that infants in the 2nd year of life can understand deceptive intentions to implant false beliefs in others. When a thief attempted to secretly steal a desirable object during its owner's absence by replacing it with a less desirable object, infants realized that this substitution could elude detection only if the substitute object was visually identical to the desirable object (deception conditions of Experiments 1 and 2) and the owner did not regularly check the desirable object's non-obvious properties when she returned (shake-twice condition of Experiment 2). When these two conditions were met, infants expected the owner to be deceived by the substitution (deceived condition of Experiment 3), unless she returned before it was completed (alerted condition of Experiment 3). Finally, infants held no expectation about the thief's actions when she inexplicably chose to steal an undesirable object (silent-control condition of Experiment 1).

These results provide robust evidence against the minimalist account of early psychological reasoning. As was discussed in Section 2.2, three signature limits of the early-developing system are that (a) it cannot handle false beliefs about identity, (b) it cannot track complex goals, such as goals that reference another agent's mental states; and (c) it cannot handle complex causal structures involving interlocking mental states. To succeed in the deception conditions of Experiments 1 and 2, however, infants had to understand that by placing the matching silent toy on the tray, T sought to lure O into holding a false belief about the identity of the toy. To succeed in the deceived condition of Experiment 3, infants had to appreciate that O would be deceived by this substitution and would mistake the toy on the tray for the rattling test toy she had left there. Thus, contrary to minimalist claims, (a) infants could reason about T's efforts to lure O into holding a false belief about the identity of the toy on the tray as well as about O's actions when she held such a false belief; (b) infants understood T's goal of secretly stealing the rattling test toy by anticipating and manipulating O's representation of the substitute toy; and (c) infants could attribute to T a causally coherent set of interlocking mental states that included her goal of secretly stealing the rattling test toy by implanting in O a false belief about the identity of the toy on the tray. Our results thus indicate that at least by 17 months of age, infants' psychological reasoning does not exhibit the signature limits thought to characterize the early-developing system.

Do our findings call into question the broader claim by minimalist researchers that two distinct systems underlie human psychological reasoning? Not necessarily: it might be possible to identify new signature limits for the early-developing system, or it might be suggested that the original signature limits identified for this system apply only to psychological reasoning in the first year of life. For our part, however, we believe that our results are more consistent with a one-system view in which psychological reasoning is mentalistic from the start, allowing infants to make sense of agents' actions by representing their motivational, epistemic, and counterfactual states. This is not to say, of course, that no important developments take place in psychological reasoning during infancy and childhood. For example, there is obviously vast improvement with age in the ease and rapidity with which psychological assessments are performed as well as in the ability to distinguish subtly different mental states and appreciate their causal implications. There are also significant changes in the ability to reflect explicitly on issues pertinent to psychological reasoning. As Carruthers (in press) pointed out, the fact that these various changes occur does not necessarily mean that two different psychological-reasoning systems must be involved. It could be "that there is just a single mindreading system that exists throughout, but which undergoes gradual conceptual enrichment through infancy and childhood" (p. 1). Recent neuroimaging findings with adults showing that the same core brain regions are recruited in intuitive and explicit false-belief tasks also support this one-system view (e.g., Hyde, Aparicio Betancourt, & Simon, in press; Kovács, Kühn, Gergely, Csibra, & Brass, 2014).

8.1. Failures to attribute false beliefs about identity in preschoolers and adults

Our findings that 17-month-olds can reason about the actions of a deceptive agent who wants to implant a false belief about an object's identity as well as about the actions of a deceived agent who holds such a false belief are consistent with the findings of Buttelmann et al. (2015), Song and Baillargeon (2008), and Scott and Baillargeon (2009) reviewed in Section 3. Together, these findings present converging evidence that a robust ability to reason about false beliefs about identity is present in the 2nd year of life. As such, these results stand in sharp contrast to recent results by Low and his colleagues (Low & Watts, 2013; Low et al., 2014) that preschoolers and even adults fail at anticipatory-looking tasks tapping false beliefs about identity. As explained below, however, these negative results are open to alternative interpretations that have little to do with limitations in false-belief understanding.

In the task used by Low and Watts (2013), 3- and 4-year-olds and adults received four familiarization trials and one test trial involving videotaped events. At the start of the first familiarization trial, a male agent stood centered behind a screen with two windows; next to each window was a box whose front and sides were covered with fringe. A blue boat traveled from the right box to the left box, and then a red boat traveled from the left box to the right box. Next, a beep sounded, the windows lit up, and after 1.75 s the agent reached through the left window and retrieved the blue boat. In the other familiarization trials, blue and red cars, ducks, and buggies were used, and the initial side of the blue object was counterbalanced; the agent consistently reached for the blue object, indicating that he preferred blue. The test trial involved a dog-robot toy that was blue on one side and red on the other. The dog first traveled from the left box to the right box with its blue side facing the agent. Inside the right box, and visible only to the participants, the dog spun several times, revealing its two sides. Finally, the dog returned to the left box, with its red side now facing the agent. The beep sounded, the windows lit up, and during the next 1.75 s anticipatory looks toward the two sides of the TV screen were measured. (For other participants the dog was initially in the right box, and in other conditions the agent preferred red in the familiarization trials; for ease of communication, however, we use the version of the task described above).

The rationale of the experiment was that if participants could attribute to the agent the false belief that the red robot was a different toy than the blue robot, then they should expect the agent to believe the blue robot was still inside the right box, and they should therefore produce anticipatory looks toward the right side of the screen. Contrary to this prediction, however, most preschoolers and adults looked first toward the left side of the screen. Low and Watts (2013) took these negative results to support the minimalist claim that looking responses are controlled by the early-developing system, which "eschews consideration of the particular way in which an object is represented by an agent" (p. 310). The results are open to an alternative, and much simpler, interpretation, however. Prior evidence indicates that looking responses can be influenced by multiple factors: in any scene, unless special steps are taken to constrain participants' responses, looks toward different portions of the scene can occur for different reasons (e.g., Ferreira, Foucart, & Engelhardt, 2013). Thus, in the test-trial scene used by Low and Watts, preschoolers and adults could have looked first toward the left side of the screen simply to see whether the dog would spin in the left box, as it had in the right box (for different deflationary interpretations of these results, see Carruthers, in press; Jacob, 2012).

In the task of Low et al. (2014), the test-trial scene again involved a screen with two windows. Centered in front of the screen was an animal cutout that was a duck on one side and a rabbit on the other; on either side of the cutout, under the windows, were snacks appropriate for the duck (bread) and the rabbit (carrots), with sides counterbalanced. After participants saw both sides of the cutout, the agent arrived and stood behind the screen, facing the duck (for other participants, the agent faced the rabbit, but we use the duck version here). Next, the beep sounded, the windows lit up, and during the next 1.75 s anticipatory looks were measured to determine which snack participants expected the agent to select. The rationale of the experiment was that if participants could take into account which animal the agent saw (the duck), then they should expect him to reach for the snack appropriate for that animal (the bread). Contrary to this prediction, however, most preschoolers and adults looked first toward the carrots. Low et al. concluded that participants' early-developing system was unable to take into account the specific way in which the agent perceived the cutout. This interpretation is questionable on two grounds, however. First, it is unclear why this task is characterized as involving false-belief understanding: all participants had to do to succeed was to track which side of the cutout the agent could see and choose the related snack. This amounts to a "level-1" perspective-taking task, and there is considerable evidence that toddlers and even infants can succeed at such simple epistemic tasks (e.g., Luo & Baillargeon, 2007; Luo & Beck, 2010; Masangkay et al., 1974; Moll & Tomasello, 2004). Second, participants might have looked first toward the carrots, not because they did not understand that the agent faced the duck, but because they thought first about which snack was appropriate for the animal *they* faced, the rabbit, before going on to think about which snack was appropriate for the animal the agent faced, the duck. This interpretation reinforces the caution expressed above that looking responses unambiguously reveal reasoning processes only when adequate constraints are in place; without these, participants may look toward different portions of the scene at different times for different reasons (for an alternative deflationary account of these results, see Jacob, 2014).

8.2. The behavioral-rule account of early psychological reasoning

Although we have focused in this article on the minimalist account of prior psychological-reasoning findings, our research also bears on the *behavioral-rule* account of these same findings (e.g., Mandler, 2012; Paulus et al., 2011; Perner, 2010; Perner & Roessler, 2012; Perner & Ruffman, 2005; Ruffman, Taumoepeau, & Perkins, 2012). A key assumption of this account is that early expectations about agents' actions are statistical rather than mentalistic in nature: in everyday life, infants gather information—in the form of statistical regularities or behavioral rules—about the actions agents typically perform in specific situations. When infants observe an agent in one of these situations in a laboratory task, they retrieve the appropriate behavioral rule to interpret or predict the agent's actions. Examples of behavioral rules that have been invoked to explain prior findings include: an agent will follow the shortest route available to a target (e.g., Gergely et al., 1995), and an agent will search for an object where it was last seen (e.g., Onishi & Baillargeon, 2005) or where it is usually placed (e.g., Surian et al., 2007). Because such rules seem plausible and could conceivably be abstracted by infants from everyday observable behaviors, the behavioral-rule account is often presented as a compelling alternative to the mentalistic account, which grants infants rich psychological interpretations laden with unobservable mental states.

Could the behavioral-rule account explain the present results? To do so, this account would need to assume that infants in the second year of life have repeated opportunities to observe various forms of deception, including deceptive actions intended to implant false beliefs in others. One possible prediction from this approach might be that infants with one or more older siblings, who presumably have more opportunities to observe (or be the victims of) deceptive actions, are more likely to possess statistical rules related to surreptitious-theft situations. To explore this possibility, we returned to the combined-deception and combined-control conditions of Experiments 1 and 2 and compared the responses of infants with one or more older siblings ($n = 33$) to those of infants without an older sibling ($n = 37$); sibling information was unavailable for two infants, who were excluded from this analysis. Infants' looking times were compared by means of an ANOVA with condition (combined-deception, combined-control), trial (matching, non-matching), and sibling (yes, no) as between-subjects factors. Only the Condition × Trial interaction was significant, $F(1, 62) = 12.99$, $p = .001$. There were no main effects or interactions involving sibling as a factor, all $Fs < 1.38$, all $ps > .244$. Infants without an older sibling looked reliably longer in the non-matching trial of the combined-deception condition ($n = 17$, $F(1, 33) = 5.29$, $p = .027$, $d = 1.07$), but looked about equally in the matching and non-matching trials of the combined-control condition ($n = 20$, $F(1, 33) = 1.27$, $p = .268$). Similarly, infants with one or more older siblings looked reliably longer in the non-matching trial of the combined-deception condition ($n = 18$, $F(1, 29) = 6.88$, $p = .013$, $d = 1.67$), but looked about equally in the two trials of the combined-control condition ($n = 15$, $F(1, 29) = 1.66$, $p = .208$). Thus, whether infants had an older sibling or not had no appreciable effect on their performance in our task. Of course, infants without an older sibling might have other opportunities to observe deceptive actions, such as in daycare interactions, play dates, and so on. Still, these results provide no support for the notion that infants in the present experiments brought to bear statistical rules about deception to make sense of O's actions.

8.3. Understanding social acting

Recent comparative reviews of social cognition suggest that chimpanzees understand motivational and epistemic states and can produce acts of tactical deception aimed at keeping others uninformed about their actions; nevertheless, chimpanzees cannot understand false beliefs (they treat misinformed agents as though they were uninformed), nor can they produce more sophisticated acts of strategic deception aimed at implanting false beliefs in others (e.g., Call & Tomasello, 2008; Hare, Call, & Tomasello, 2006; Tomasello & Moll, 2013; Whiten, 2013). These findings stand in sharp contrast to those obtained with human infants, who not only can understand false beliefs, as shown in prior research, but also can make sense of acts of strategic deception intended to implant false beliefs, as shown here. The infants in Experiments 1–3 were able to judge under what conditions T's substitution of a silent toy was likely to be effective at deceiving O. When this substitution was judged to be effective, the infants expected O to hold a false belief about the substitute toy's identity and to act accordingly. Had O been expected to be merely ignorant or uninformed about the toy's identity, then the infants in the deceived condition of Experiment 3 would have looked equally whether O stored or discarded the toy, as an ignorant O could have performed either action. This is in fact what happened in the alerted condition of Experiment 3, where O caught T in the act and was ignorant about which toy T had placed on the tray, the rattling test toy or the silent matching toy from the trashcan. In the deceived condition, in contrast, the infants expected O to be appropriately fooled and to store the silent matching toy in her box. The infants were thus able to reason about both T's effective act of strategic deception and O's resulting false belief in the identity of the toy on the tray.

This marked gap between the psychological-reasoning capacities of chimpanzees and human infants raises interesting questions about the functions of false-belief understanding in everyday life. Why might humans have evolved the capacity to attribute false beliefs? Why does false-belief understanding matter? Our capacity for understanding and implanting false beliefs no doubt serves us well in a variety of competitive situations (e.g., hunting, sports, war, politics, and corporate dealings). This same capacity may also be critical in everyday cooperative situations, however. According to a recent hypothesis (Baillargeon et al., 2013; Yang & Baillargeon, 2013), one important function of our abstract ability to represent false beliefs, pretense, and other counterfactual mental states is that it makes possible *social acting*, the well-intentioned social pretense we routinely produce in the form of white lies, tactful omissions, feigned interest, hidden disappointments, and false cheer. Social acting helps to maintain positivity within groups: it prevents aggressive confrontations, avoids hurt or embarrassed feelings, smoothes over awkward situations, and bolsters feelings of trust and acceptance. From an evolutionary standpoint, it does not seem implausible that during the millions of years our ancestors lived in small bands of hunter-gatherers, selective pressures supported the acquisition of various pro-group biases, including positivity; after all, positivity would facilitate cooperation within a group and as such would contribute to the group's long-term prosperity and survival (e.g., Baillargeon et al., 2015; Brewer, 1999).

According to the social-acting hypothesis, it is thus no accident that human infants can interpret the actions of agents who hold false beliefs as well as those of agents who seek to implant false beliefs; both abilities are essential for social acting. Of course, many years of experience are necessary before children become adept at producing and interpreting social acting. Skillful, nuanced, and context-sensitive social acting is a staggering accomplishment, not fully achieved until late in development, and profoundly shaped by familial, social, and cultural practices (e.g., Ma, Xu, Heyman, & Lee, 2011; Xu, Bao, Fu, Talwar, & Lee, 2010).

8.4. Conclusion

In sum, the present findings provide new evidence that 17-month-olds can represent and reason about false beliefs about identity. These findings call into question the minimalist account of early psychological reasoning, and they support the mentalistic claim that an abstract capacity for false-belief understanding emerges early in human development.

Acknowledgments

This research was supported by NICHD Grant HD-021104 to RB and by grants from the University of California Merced Graduate Research Council and Hellman Fellows Fund to RMS. We thank the staff of the University of Illinois Infant Cognition Laboratory for their help with the data collection, as well as the parents and infants who participated in the research.

References

Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review, 116*, 953–970.

Baillargeon, R., Scott, R. M., & Bian, L. (in press). Psychological reasoning in infancy. *Annual Review of Psychology*.

Baillargeon, R., He, Z., Setoh, P., Scott, R. M., Sloane, S., & Yang, D. Y. J. (2013). False-belief understanding and why it matters: The white-lie hypothesis. In M. Banaji & S. Gelman (Eds.), *Navigating the social world: What infants, children, and other species can teach us* (pp. 88–95). New York: Oxford University Press.

Baillargeon, R., Scott, R. M., & He, Z. (2010). False-belief understanding in infants. *Trends in Cognitive Sciences, 14*, 110–118.

Baillargeon, R., Scott, R. M., He, Z., Sloane, S., Setoh, P., Jin, K., ... Bian, L. (2015). Psychological and sociomoral reasoning in infancy. In M. Mikulincer, P. R. Shaver (Eds.), E. Borgida, & J. A. Bargh (Assoc. Eds.), *APA handbook of personality and social psychology: Vol. 1. Attitudes and social cognition* (pp. 79–50). Washington, DC: American Psychological Association.

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

Barrett, C., Broesch, T., Scott, R. M., He, Z., Baillargeon, R., Wu, D., ... Laurence, S. (2013). Is early false-belief reasoning universal? Evidence from three non-western cultures. *Proceedings of the Royal Society B, 280*, 1755.

Brewer, M. B. (1999). The psychology of prejudice. Ingroup love or outgroup hate? *Journal of Social Issues, 55*, 429–444.

Buttelmann, D., Carpenter, M., & Tomasello, M. (2009). 18-month-old infants show false-belief understanding in an active helping paradigm. *Cognition, 112*, 337–342.

Buttelmann, D., Over, H., Carpenter, M., & Tomasello, M. (2014). Eighteen-month-olds understand false beliefs in an unexpected-contents task. *Journal of Experimental Child Psychology, 119*, 120–126.

Buttelmann, F., Suhrke, J., & Buttelmann, D. (2015). What you get is what you believe: Eighteen-month-olds demonstrate belief understanding in an unexpected-identity task. *Journal of Experimental Child Psychology, 131*, 94–103.

Butterfill, S., & Apperly, I. A. (2013). How to construct a minimal theory of mind. *Mind and Language, 28*, 606–637.

Call, J., & Tomasello, M. (2008). Does the chimpanzee have a theory of mind? 30 years later. *Trends in Cognitive Sciences, 12*, 187–192.
Carruthers, P. (2013). Mindreading in infancy. *Mind & Language, 28*, 141–172.
Carruthers, P. (in press). Two systems for mindreading? *Review of Philosophy and Psychology*.
Christensen, W. & Michael, J. (in press). From two systems to a multi-systems architecture for mindreading. *New Ideas in Psychology*.
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.
Dennett, D. C. (1987). *The intentional stance*. Cambridge, MA: MIT Press.
Ferreira, F., Foucart, A., & Engelhardt, P. E. (2013). Language processing in the visual world: Effects of preview, visual complexity, and prediction. *Journal of Memory and Language, 69*, 165–182.
Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: The naıve 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.
Hamlin, J. K., Mahajan, N., Liberman, Z., & Wynn, K. (2013). Not like me = bad: Infants prefer those who harm dissimilar others. *Psychological Science, 24*, 589–594.
Hamlin, J. K., Wynn, K., Bloom, P., & Mahajan, N. (2011). How infants and toddlers react to antisocial others. *Proceedings of the National Academy of Science, 108*, 19931–19936.
Hare, B., Call, J., & Tomasello, M. (2006). Chimpanzees deceive a human competitor by hiding. *Cognition, 101*, 495–514.
Heyes, C. (2014). False belief in infancy: A fresh look. *Developmental Science, 17*, 647–659.
Hyde, D. C., Aparicio Betancourt, M., & Simon, C. E. (in press). Human temporal-parietal junction automatically tracks other's beliefs: An fNIRS study. *Human Brain Mapping*.
Jacob, P. (2012, July 24). Do we use different tools to mindread a defendant and a goalkeeper? *Cognition and Culture Blog*. Retrieved from <http://www.cognitionandculture.net/home/blog/44-pierre-jacobs-blog/2455-do-we-use-different-tools-to-mindread-a-defendant-and-a-goalkeeper>.
Jacob, P. (2014, October 25). Another look at the two-systems model of mindreading. *Cognition and Culture Blog*. Retrieved from <http://www.cognitionandculture.net/home/blog/44-pierre-jacobs-blog/2646-another-look-at-the-two-systems-model-of-mindreading>.
Knudsen, B., & Liszkowski, U. (2012). Eighteen- and 24-month-old infants correct others in anticipation of action mistakes. *Developmental Science, 15*, 113–122.
Kovács, Á. M., Kühn, S., Gergely, G., Csibra, G., & Brass, M. (2014). Are all beliefs equal? Implicit belief attributions recruiting core brain regions of theory of mind. *PLOS ONE, 9*, e106558.
Kovács, Á. M., Téglás, E., & Endress, A. D. (2010). The social sense: Susceptibility to others' beliefs in human infants and adults. *Science, 330*, 1830–1834.
Kushnir, T., Xu, F., & Wellman, H. M. (2010). Young children use statistical sampling to infer the preferences of other people. *Psychological Science, 21*, 1134–1140.
Liszkowski, U., Carpenter, M., & Tomasello, M. (2008). Twelve-month-olds communicate helpfully and appropriately for knowledgable and ignorant partners. *Cognition, 108*, 732–739.
Low, J. (in press). Two-systems view of children's theory-of-mind understanding. *Emerging Trends in the Social and Behavioral Sciences*.
Low, J., Drummond, W., Walmsley, A., & Wang, B. (2014). Representing how rabbits quack and competitors act: Limits on preschoolers' efficient ability to track perspectives. *Child Development, 85*, 1519–1534.
Low, J., & Watts, J. (2013). Attributing false beliefs about object identity reveals a signature blind spot in humans' efficient mind-reading system. *Psychological Science, 24*, 305–311.
Luo, Y. (2011). Do 10-month-old infants understand others' false beliefs? *Cognition, 121*, 289–298.
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., & Beck, W. (2010). Do you see what I see? Infants' reasoning about others' incomplete perceptions. *Developmental Science, 13*, 134–142.
Ma, F., Xu, F., Heyman, G. D., & Lee, K. (2011). Chinese children's evaluations of white lies: Weighing the consequences for recipients. *Journal of Experimental Child Psychology, 108*, 308–321.
Mandler, J. M. (2012). On the spatial foundations of the conceptual system and its enrichment. *Cognitive Science, 36*, 421–451.
Masangkay, Z. S., McCluskey, K. A., McIntyre, C. W., Sims-Knight, J., Vaughn, B. E., & Flavell, J. H. (1974). The early development of inferences about the visual percepts of others. *Child Development, 45*, 357–366.
Moll, H., & Tomasello, M. (2004). 12- and 18-month-old infants follow gaze to spaces behind barriers. *Developmental Science, 7*, F1–F9.
Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science, 308*, 255–258.
Paulus, M., Hunnius, S., van Wijngaarden, C., Vrins, S., van Rooij, I., & Bekkering, H. (2011). The role of frequency information and teleological reasoning in infants' and adults' action prediction. *Developmental Psychology, 47*, 976–983.
Perner, J. (2010). Who took the cog out of cognitive science? Mentalism in an era of anti-cognitivism. In P. A. Frensch & R. Schwarzer (Eds.), *Cognition and neuropsychology: International perspectives on psychological science* (pp. 241–261). Hove, UK: Psychology Press.
Perner, J., & Roessler, J. (2012). From infants' to children's appreciation of belief. *Trends in Cognitive Sciences, 16*, 519–525.
Perner, J., & Ruffman, T. (2005). Infants' insight into the mind: How deep? *Science, 308*, 214–216.
Ruffman, T., Taumoepeau, M., & Perkins, C. (2012). Statistical learning as a basis for social understanding in children. *British Journal of Developmental Psychology, 30*, 87–104.
Scott, R. M., & Baillargeon, R. (2009). Which penguin is this? Attributing false beliefs about object identity at 18 months. *Child Development, 80*, 1172–1196.
Scott, R. M., & Baillargeon, R. (2013). Do infants really expect others to act efficiently? A critical test of the rationality principle. *Psychological Science, 24*, 466–474.

Scott, R. M., Baillargeon, R., Song, H., & Leslie, A. M. (2010). Attributing false beliefs about non-obvious properties at 18 months. *Cognitive Psychology, 61*, 366–395.

Scott, R. M., Roby, E., & Smith, M. (in press). False-belief understanding in the first years of life. In J. Kiverstein (Ed.), *The Routledge handbook of philosophy of the social mind.*

Senju, A., Southgate, V., Snape, C., Leonard, M., & Csibra, G. (2011). Do 18-month-olds really attribute mental states to others? A critical test. *Psychological Science, 22*, 878–880.

Song, H., & Baillargeon, R. (2008). Infants' reasoning about others' false perceptions. *Developmental Psychology, 44*, 1789–1795.

Song, H., Onishi, K. H., Baillargeon, R., & Fisher, C. (2008). Can an agent's false belief be corrected by an appropriate communication? Psychological reasoning in 18-month-old infants. *Cognition, 109*, 295–315.

Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science, 18*, 587–592.

Surian, L., Caldi, S., & Sperber, D. (2007). Attribution of beliefs by 13-month-old infants. *Psychological Science, 18*, 580–586.

Thompson, J. R. (2014). Signature limits in mindreading systems. *Cognitive Science, 38*, 1432–1455.

Tomasello, M., & Haberl, K. (2003). Understanding attention: 12- and 18-month-olds know what is new for other persons. *Developmental Psychology, 39*, 906–912.

Tomasello, M., & Moll, H. (2013). Why don't apes understand false beliefs? In M. R. Banaji & S. A. Gelman (Eds.), *Navigating the social world: What infants, children, and other species can teach us* (pp. 81–88). New York: Oxford University Press.

Träuble, B., Marinović, V., & Pauen, S. (2010). Early theory of mind competencies: Do infants understand others' beliefs? *Infancy, 15*, 434–444.

Wellman, H. M., & Bartsch, K. (1988). Young children's reasoning about beliefs. *Cognition, 30*, 239–277.

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.

Whiten, A. (2013). Humans are not alone in computing how others see the world. *Animal Behaviour, 86*, 213–221.

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. (1999). Infants' ability to distinguish between purposeful and non-purposeful behaviors. *Infant Behavior and Development, 22*, 145–160.

Woodward, A. L., & Sommerville, J. A. (2000). 12-month-old infants interpret action in context. *Psychological Science, 11*, 73–77.

Xu, F., Bao, X., Fu, G., Talwar, V., & Lee, K. (2010). Lying and truth-telling in children: From concept to action. *Child Development, 81*, 581–596.

Yang, D. Y. J., & Baillargeon, R. (2013). Difficulty in understanding social acting (but not false beliefs) mediates the link between autistic traits and ingroup relationships. *Journal of Autism and Developmental Disorders, 43*, 2199–2206.

Zawidzki, T. W. (2011). How to interpret infant socio-cognitive competence. *Review of Philosophy and Psychology, 2*, 483–497.