Title
The relationship between parental mental-state language and 2.5-year-olds' performance on a nontraditional false-belief task

Authors
Roby, Erin
Scott, Rose M

Publication Date
2018-11-01

DOI
10.1016/j.cognition.2018.06.017

Peer reviewed

The relationship between parental mental-state language and 2.5-year-olds' performance on a nontraditional false-belief task

Erin Roby[a,b,*], Rose M. Scott[a]

[a]Psychological Sciences, University of California Merced, 5200 North Lake Road, Merced, California, 95343, United States of America
[b]Department of Pediatrics, NYU School of Medicine, 550 First Ave., New York, NY 10016

*Corresponding author: erin.roby@nyumc.org

Abstract

A growing body of evidence suggests that children succeed in nontraditional false-belief tasks in the first years of life. However, few studies have examined individual differences in infants' and toddlers' performance on these tasks. Here we investigated whether parental use of mental-state language (i.e. *think*, *understand*), which predicts children's performance on elicited-response false-belief tasks at older ages, also predicts toddlers' performance on a nontraditional task. We tested 2.5-year-old children in a verbal nontraditional false-belief task that included two looking time measures, anticipatory looking and preferential looking, and measured parents' use of mental-state language during a picture-book task. Parents' use of mental-state language positively predicted children's performance on the anticipatory-looking measure of the nontraditional task. These results provide the first evidence that social factors relate to children's false-belief understanding prior to age 3 and that this association extends to performance on nontraditional tasks. These findings add to a growing number of studies suggesting that mental-state language supports mental-state understanding across the lifespan.

Keywords: False-belief understanding, psychological reasoning, social cognition, mental-state language

1. Introduction

The ability to predict and interpret the behavior of other individuals in terms of their unobservable mental states (e.g., goals, preferences, beliefs) is essential for navigating the social world. Researchers have long been interested in when and how this critical psychological reasoning ability develops. In particular, considerable research has focused on when children understand that others can be mistaken, or hold false beliefs, about the world. Early investigations into this question relied on traditional elicited-response false-belief tasks, which require children to answer direct questions about the likely behavior of a mistaken agent (e.g., Baron-Cohen, Leslie, & Frith, 1985; Wimmer & Perner, 1983). Children's performance on such tasks led to the conclusion that the capacity to represent false beliefs does not emerge until at least age 4 (e.g., Wellman, Cross, & Watson, 2001).

However, this conclusion has been challenged by recent evidence from alternative, nontraditional tasks that do not require children to answer direct questions about a mistaken agent. In these tasks, researchers instead measure a variety of other responses that children produce as they observe or interact with a mistaken agent, including their looking behavior (e.g., Onishi & Baillargeon, 2005; Southgate, Senju, & Csbra, 2007), emotional expressions (e.g., Moll, Kane, & McGowan, 2016), neurological activity (e.g., Hyde, Simon, Ting, & Nikolaeva, in press; Kampis, Parise, Csibra, & Kovács, 2015; Southgate & Vernetti, 2014), and helping responses (e.g., Buttelmann, Carpenter, & Tomasello, 2009; Southgate, Chevallier & Csibra, 2010). Over 30 published studies using nontraditional false-belief tasks have now reported positive results with children between 6 months and 3 years of age (Scott & Baillargeon, 2017; Scott, Roby, & Smith, 2017). These findings have led many researchers to conclude that the capacity to represent false beliefs emerges in the first year of life (e.g., Barrett et al., 2013;

Buttelmann et al., 2009; Carruthers, 2013; Kovács, Téglás, & Endress, 2010; Luo, 2011; Scott, 2017; Southgate et al., 2007; Surian, Caldi, & Sperber, 2007).[1]

Despite the growing body of research using nontraditional false-belief paradigms, very few studies have examined individual variation in infants' and toddlers' performance on these tasks. Thus, little is known about individual differences in early false-belief understanding and what factors might be responsible for such differences. In particular, it is an open question whether the same factors that predict preschooler's performance on traditional elicited-response tasks also predict younger children's performance on nontraditional false-belief tasks.

Here we addressed this issue by examining whether toddlers' performance on a verbal nontraditional false-belief task was related to aspects of their social environment. Several decades of research have shown that a number of social factors predict preschool-aged children's performance on elicited-response false-belief tasks (e.g., Cutting & Dunn, 1999; Devine & Hughes, 2018; McAlister & Peterson, 2013; Meins et al., 2003; Perner, Ruffman, & Leekam, 1994; Ruffman, Slade, & Crow, 2002; Symons & Clark, 2000). In particular, there is a well-established positive relationship between parental use of mental-state language, terms that refer to psychological states such as *think*, *know*, and *remember*, and preschoolers' elicited-response performance (e.g., Adrián, Clemente, Villanueva & Rieffe, 2005; Brown, Donelan-McCall & Dunn, 1996; Devine & Hughes, 2018; Ensor & Hughes, 2008; Howard, Mayeux & Naigles, 2008; Nielson & Dissanayake, 2000; Ruffman et al., 2002). Parental mental-state language predicts their preschoolers' performance on elicited-response tasks both concurrently (e.g., Howard et al., 2008; LaBounty, Wellman, Olson, Lagattuta, & Liu, 2008; Peterson & Slaughter,

[1] Several researchers have questioned whether nontraditional false-belief tasks assess the same capacity as elicited-response tasks (e.g., Butterfill & Apperly, 2013; Heyes, 2014; Ruffman, 2014), as well as whether the results from nontraditional tasks are robust (Crivello & Poulin-Dubois, in press; Dörrenberg, Rakoczy & Liszkowski, in press; Powell, Hibbs, Bardis, Carey & Saxe, in press). We return to both of these issues in the General Discussion.

2003) and longitudinally (e.g., Adrián, Clemente, & Villanueva, 2007; Ensor & Hughes, 2008; Meins et al., 2003; Ruffman et al., 2002), and this association is evident when mental-state language is assessed in naturalistic (Howard et al., 2008) and laboratory settings (Ruffman et al., 2002).

Additional evidence for the relationship between parental mental-state language and elicited-response performance comes from three sets of findings. First, preschoolers whose parents were trained to engage in elaborative reminiscing, a style of discussing past events that involves rich memory cues and references to mental states (e.g., *remember*), performed better on elicited-response tasks than children whose parents had not received training (Taumoepeau & Reese, 2013). Second, deaf children raised by hearing parents, who hear fewer references to mental states than their hearing counterparts, exhibit deficits in performance on elicited-response tasks (Gale, de Villiers, de Villiers, & Pyers, 1996; Meristo et al., 2007; Moeller & Schick, 2006). Finally, in cultures where parents do not typically discuss mental states with their children (e.g., Taumoepeau, 2015), children pass elicited-response tasks at later ages (Mayer & Träuble, 2013). Together, these findings show a strong relationship between parental use of mental-state language and preschool children's performance on elicited-response false-belief tasks.

Could social factors such as parental use of mental-state language also be related to younger children's performance on nontraditional false-belief tasks? This depends on the causal mechanism underlying the relationship between social factors and performance on elicited-response tasks. One possibility is that social factors specifically facilitate children's ability to answer the direct questions used in elicited-response tasks. For instance, in one common false-belief scenario, an agent places an object in one of two locations; the object is then moved to the other location in her absence. In elicited-response tasks, children are asked a direct question such as "Where will she look for the toy?" It has been argued that pragmatic factors might cause children to misinterpret this question as asking something else, such as where the agent ought to look for the object or where the object is actually located (e.g., Hansen, 2010; Helming, Strickland, & Jacob, 2016; Siegal & Beattie, 1991; Westra & Carruthers, 2017; Yazdi, German, Defeyter, & Siegal, 2006). Frequently engaging in social interactions that involve discussions of mental states might help children overcome this pragmatic ambiguity, enabling them to interpret the question correctly and produce the appropriate response. If social factors specifically improve children's ability to answer direct questions about others' behavior, then we would not expect to see relationships between these social factors and performance on nontraditional tasks because they do not involve such questions.

An alternative possibility, however, is that the relationship between social factors and false-belief performance extends beyond facilitating children's ability to answer direct questions about the behavior of mistaken agents. In the scenario described above, several processes contribute to children's ability to understand where the mistaken agent will search for the object (e.g., Kampis, Fogd, & Kovács, 2017; Roby & Scott, 2016b). Children must attend to the agent and mental-state relevant information within the scene, such as which events the agent has or has not witnessed, and use this information to infer the agent's mental states (e.g., the agent holds a false belief about the object's location). When the agent returns to the scene, they must retrieve the agent's mental states and use them to infer the agent's likely actions. Critically, these processes are required regardless of whether children's understanding is ultimately assessed via their answers to direct questions, or alternative responses such as which of the two locations children look toward in anticipation of the agent's search for the object (e.g., anticipatory-looking; Southgate et al., 2007).

Children's social experiences could plausibly influence each of these processes (Roby & Scott, 2016b). For instance, social interactions that involve talk about others' mental states may heighten interest in agents and their mental states, increasing children's tendency to attend to agents over other aspects of a scene. Beyond drawing attention to agents, discussions involving others' mental states might help children learn how particular behaviors or situational cues are linked to particular mental states (e.g., Christensen & Michael, 2016; Scott et al., 2017), thereby improving children's ability to infer others' mental states and predict their subsequent actions in a range of situations. Finally, conversations involving mental-state language provide scaffolded practice remembering events and the mental states of those involved. Learning and using mental-state language also provides children with a useful tool for retrieving and holding in mind belief-relevant information (e.g., San Juan & Astington, 2012). Thus, social experiences could improve children's ability to retrieve mental-state relevant information when needed.

If the preceding analysis is correct, then we might see relationships between parental use of mental-state language and young children's performance on nontraditional false-belief tasks. Although no studies have directly examined this relationship, two sets of findings provide indirect support for this possibility. First, there is growing evidence that parental mental-state talk is related to infants' and toddlers' performance on a variety of social reasoning tasks (e.g., Drummond et al., 2014; Newton, Thompson, & Goodman, 2016; Taumoepeau & Ruffman, 2008). For instance, mothers' use of mental-state language, in particular their use of the terms *think* and *know*, with their 24-month-old children predicts children's emotion understanding at 33 months (Taumoepeau & Ruffman, 2008). In addition, maternal sensitivity and mother's use of mental-state language predicts 18-month-olds' tendency to engage in prosocial helping across a variety of contexts (Newton et al., 2016). These findings suggest that rather than being specifically related to preschool children's ability to answer direct questions about the behavior of mistaken agents, parental use of mental-state language is related to a range of social-cognitive skills even before the preschool years.

Second, recent evidence suggests that deaf children of hearing parents, who have difficulty with traditional false-belief tasks, also exhibit deficits on nontraditional false-belief tasks (e.g., Meristo et al., 2012; Meristo, Strid, & Helmquist, 2016). For instance, Meristo et al. (2012) found that when tested in a nonverbal anticipatory-looking false-belief task, hearing 23-month-old infants successfully anticipated the actions of a mistaken agent, but the deaf infants of hearing parents did not. Although Meristo et al. (2012) did not directly assess the mental-state talk of the parents of the infants in their study, other investigations have shown that deaf toddlers of hearing parents hear significantly fewer references to mental states than same-aged hearing peers raised with hearing parents (Morgan et al., 2014). This raises the possibility that the deaf infants performed worse on the anticipatory-looking task in part due to a lack of exposure to mental-state talk. These findings thus provide suggestive, albeit indirect, evidence that parental mental-state talk might be related to children's performance on nontraditional false-belief tasks prior to the preschool years.

2. The present research

The evidence reviewed in the previous section suggests that parental mental-state language could be related to young children's performance on nontraditional false-belief tasks. The present study directly tested this possibility. Our experimental approach was motivated by a recent study in which we investigated how increasing task demands impacted toddlers' performance on a verbal nontraditional false-belief task (Scott & Roby, 2015). In that study, 3-year-olds were tested in a verbal preferential-looking task adapted from Scott, He, Baillargeon, and Cummins (2012) in which they heard a false-belief story accompanied by pictures, and the linguistic ambiguity of the story (i.e. whether key story lines were open to multiple interpretations) varied across children. The final page of the story showed two images: one consistent with the main character acting on her false belief, and the other consistent with reality. Replicating Scott et al. (2012), when the story was unambiguous children looked longer at the belief-consistent image, successfully demonstrating false-belief understanding. When the story was ambiguous children performed at chance as a group, and performance was positively correlated with children's verbal ability: those with better verbal abilities looked longer at the belief-consistent image, while those with low verbal abilities looked longer at the reality-consistent image. No such correlation occurred when the story was unambiguous. These results suggest that this preferential-looking task offers a promising way of investigating individual differences in young children's performance on nontraditional false-belief tasks.

We therefore tested 2.5-year-olds in a verbal nontraditional false-belief task adapted from Scott et al.'s (2012) preferential-looking paradigm. Given the indirect evidence from Meristo et al. (2012) suggesting that parental mental-state language might relate to anticipatory-looking performance, we modified the task to include both an anticipatory-looking trial and a preferential-looking trial. Our task thus included two measures that have been used to document individual differences in young children's false-belief performance, increasing the likelihood that we might observe a correlation with parental mental-state language. Based on the contrasting results of the two conditions in Scott and Roby (2015), we reasoned that if parental mental-state language were related to young children's performance on nontraditional false-belief tasks, we would be more likely to observe this relationship under higher task demands. We therefore (a) included younger children than have previously demonstrated successful performance in verbal anticipatory-looking tasks or preferential-looking tasks on the assumption that they would find the task more difficult and (b) made several procedural modifications intended to make the task more challenging (see Section 3.2.1). After completing the false-belief task, parent-child dyads completed a picture-book task in which the parents described pictures to their child (Taumoepeau & Ruffman, 2006). We then examined whether parents' use of mental-state terms during the picture-book task was related to their children's performance on either of the looking-time measures in the false-belief task.

An additional group of 2.5-year-olds were tested in a true-belief version of the nontraditional task. The primary purpose of this condition was to rule out potential low-level interpretations of children's performance in the false-belief condition (e.g., Heyes, 2014; Ruffman, 2014). A secondary goal was to explore whether children's true-belief performance was related to parental mental-state talk in the picture-book task. To our knowledge, none of the prior studies on the relationship between parental mental-state talk and performance on elicited-response tasks have included a true-belief condition. The present study thus provided an opportunity to explore whether parental mental-state talk specifically predicts children's ability to reason about situations in which an agent holds a false belief, or whether it relates more generally to children's ability to predict and interpret how an agent will behave, given their mental states. Because parental mental-state language appears to be broadly related to young children's social cognition, as reviewed in the previous section, and many of the processes involved in false-belief tasks – attending to, inferring, and retrieving mental-state relevant information – are also involved in true-belief tasks, we predicted that similar patterns of relationships between parental mental-state language and performance might emerge in the false- and true-belief conditions.

3. Method

3.1. Participants

Seventy-six 2.5-year-olds (27.3 – 32.5 months, *M* = 29.3 months[2]), 38 male and 38 female, participated in the study. An additional 9 children completed all tasks but were excluded because they were inattentive (4) or fussy (1) during the nontraditional task, because the difference between their test looking times in the preferential-looking trial of the nontraditional task were over 3 standard deviations away from the mean of their condition (3), or because the number of utterances that their parent used in the picture-description task was over 3 standard deviations away from the mean of the sample (1) (the decision to exclude outliers was made prior to data collection). A further 9 children were excluded because they failed to complete one or more tasks: 4 children failed to complete the nontraditional task, 2 failed to complete the picture-description task, and 3 did not successfully complete either task. Equal numbers of males and females were randomly assigned to the false-belief (*N* = 38, *M* = 29.6 months) and the true-belief (*N* = 38, *M* = 29.1 months) condition of the nontraditional task.

All children were native English speakers. Children's vocabulary was measured using the short form of the MacArthur-Bates Communicative Development Inventory, Level 2 (Fenson et al., 2000). Vocabulary scores ranged from 12 to 100 with a median of 71.5.

The majority of children completed the picture-description task with their mother (*N* = 62); the remainder completed the task with their father (*N* = 14). We recorded the education level of the participating parent: 5 had less than a high school diploma, 24 completed high school, 14 had an Associate's degree, 20 had a Bachelor's degree, 12 had a Master's degree, and 1

[2] The youngest child tested by Scott et al. (2012) was 29.9 months old; the youngest child tested in He, Bolz, and Baillargeon (2012), the only prior study to report positive results in a verbal anticipatory-looking task with 2.5-year-olds, was 29.0 months old. Thus, the mean age of our sample was comparable to the youngest age tested in these prior studies.

completed a professional degree such as an MD or PhD. Parents were asked to indicate their child's race and ethnicity: 51 of the children were identified as White, 6 as Asian, 3 as African American and 1 as American Indian or Alaska Native; 6 parents chose 'other race', 4 selected 'more than one race', and 5 chose not to respond. 36 children were identified as Hispanic or Latino, 38 as Not Hispanic or Latino, and 2 parents chose not to respond.

Prior research has shown that children with siblings, and in particular those with an older sibling, show superior performance on elicited-response false-belief tasks (e.g., Lewis, Freeman, Kyriakidou, Maridaki-Kassotaki, & Berridge, 1996; McAlister & Peterson, 2007; McAlister & Peterson, 2013; Perner et al., 1994). We therefore included two sibling variables in our analyses in an exploratory fashion: the number of siblings the children had (sibling-number) and whether or not the child had an older sibling (older-sibling). 19 children did not have a sibling, 27 had one sibling, 13 had 2 siblings, 10 had 3 siblings, and 7 had 4 or more siblings. 41 children had at least one older sibling, whereas 35 children did not have an older sibling.

The toddlers' names were obtained from birth records provided by the California Department of Public Health and from parents who had previously expressed interest in participating in research studies with their children. Parents provided written informed consent for participation.

3.2. Apparatus, Materials, and Procedure

All children first completed the nontraditional task. The children and their parents were then taken to a separate room where they were left alone to complete the picture-book task.

3.2.1. Nontraditional task

Children sat on their parent's lap in front of a large television screen (68.5 × 122 cm) approximately 91.5 cm away. A camera hidden at the base of the television (centered, 89 cm high) recorded children's eye movements during the study. Parents were asked to close their eyes or look down to prevent them from biasing their children's responses during the video.

Materials consisted of color images of two male actors. The images were presented on the television screen and accompanied by a soundtrack recorded by a native English speaker. The video procedure consisted of three phases: introduction, setup, and test (see Appendix A for pictures and script).

To start, the television screen was blank and children heard, "This is a story about a boy named Lucas." Children then saw the first introduction trial: an image (30.5 cm x 39 cm) of a male actor waving was shown on one side of the screen for 6 seconds and children heard, "Look! There's Lucas!" (introduction-1). The other side of the screen remained blank. There was then a 6-s blank-screen interval in which children heard, "Lucas has a friend named Jacob." Jacob was then introduced on the other side of the screen in the same manner (introduction-2).

Next, children saw four 8-s setup trials. Each trial was preceded by an 8-s blank-screen interval in which children heard a line of the story. Two images were then displayed simultaneously (25 cm apart) and the story line was repeated. In each trial, one image matched the story and the other was irrelevant.

In the setup trials in the *false-belief* condition, children first heard that Lucas and Jacob liked to play together (setup-1). Next, children were told that the Lucas and Jacob were going to play hide and seek and Jacob was going to hide (setup-2). Lucas peeked and saw Jacob hiding in one of two locations (in a tent or under a table; setup-3). Lucas then closed his eyes and Jacob moved to the other hiding location (setup-4).

The setup trials in the *true-belief condition* were identical to those in the false-belief condition with two exceptions (see Appendix). In setup-3, Lucas closed his eyes while Jacob hid in one of the two possible hiding locations. Lucas then peeked in the next trial (setup-4) and saw Jacob move to the other hiding location. Thus, in both conditions Lucas saw one hiding event but not the other; what differed was *which* hiding event he witnessed. In the false-belief condition, Lucas witnessed the first but not the second hiding event and therefore held a false belief about Jacob's location. In the true-belief condition, Lucas did not witness the first hiding event, but he did see the second hiding event and hence he held a true belief about Jacob's location.

The side of the matching image in the introduction and setup trials was counterbalanced across trials and across children: in each trial, half the children in each belief condition saw the matching image on the left, and half saw it on the right. The order in which Jacob hid in the two locations (tent then table, or table then tent) was counterbalanced across sex, belief condition, and side of the matching image in the introduction and setup trials. These two counterbalancing factors are henceforth referred to as hide-order and setup-side condition, respectively.

The test phase consisted of two trials separated by an 8-s blank-screen interval. First children saw an 8-s anticipatory-looking trial. In the blank-screen interval preceding the trial, children heard, "Lucas is ready to look for Jacob." Next, children saw a single large image (69 cm x 90 cm) centered on the screen. The image showed Lucas standing between the two hiding locations facing the camera (the hiding locations were 32 cm apart on the screen). While viewing the image, children heard, "Lucas says, 'Ready or not! Here I come!'" Next, children received an 8-s preferential-looking test trial. In the blank screen-interval preceding the trial, children heard, "Lucas goes to find Jacob." Two images were then presented simultaneously: one image showed Lucas approaching the tent and the other showed Lucas approaching the table. Children heard, "Lucas is looking for Jacob." The sides of the images in the preferential-looking trial were counterbalanced across sex, belief condition, hide-order, and setup-side condition.

In the only prior study that tested 2.5-year-olds in a verbal anticipatory-looking task (He et al., 2012), researchers analyzed children's anticipatory looking during the first 5s following the anticipatory prompt. We used a longer 8-s anticipatory-looking trial for two reasons. First, our sample included younger children (i.e. 27- and 28-month-olds) than were tested in He et al. (2012). Second, in He et al.'s task, children had time to inspect the scene prior to the anticipatory prompt. In the current task, the image appeared immediately after the anticipatory prompt ("Lucas is ready to look for Jacob."). It was therefore possible that the children in our sample might require additional time to process the scene, retrieve Lucas' belief, and identify the location he should approach given that belief.

As stated in Section 2, the design of this task included several modifications intended to increase the demands of the task. First, unlike most of the nontraditional tasks used with infants and toddlers (e.g., Buttelmann et al., 2009; Kovács et al., 2010; Onishi & Baillargeon, 2005; Southgate & Vernetti, 2014), including all previous preferential-looking tasks (Scott et al., 2012; Scott & Roby, 2015) and verbal anticipatory-looking tasks (Clements & Perner, 1994; Clements, Rustin, & McCallum, 2000; Garnham & Perner, 2001; He et al., 2012), the agent (Lucas) was present throughout the entire story. Thus, in our task children could not use the agent's presence or absence to infer his belief about Jacob's location. Instead, they had to monitor the agent's attention to determine which events he had witnessed. Second, the majority of anticipatory-looking false-belief tasks used with infants and toddlers have been low-inhibition tasks: the agent believed an object was in one of two locations, when in fact it had been removed and was no longer in either location (e.g., He et al., 2012; Senju, Southgate, Snape, Leonard, & Csibra, 2011; Southgate et al., 2007). In contrast, we used a high-inhibition task: Lucas believed Jacob was in one location when he was actually in the other location. Note that the anticipatory-looking task used by Meristo et al., (2012), on which deaf infants of hearing parents exhibited difficulty, was also a high-inhibition task. Finally, as described above, the image in the anticipatory-looking trial did not appear until after the anticipatory prompt. This prevented children from inspecting the scene and identifying Lucas' likely search location prior to the prompt. This timing increased the likelihood that individual differences in children's belief-based anticipations would be captured during the anticipatory-looking trial.

3.2.2. Picture-book task

Parents sat in a chair with their child seated on their lap. A hidden camera centered in front of the chair recorded the parents' verbal utterances. A second camera mounted above and behind the parent-child dyad captured the stimuli. A small, inconspicuous baby monitor was placed on a table behind the parent's chair.

Materials consisted of 23 (20 cm x 28 cm) color photos used in previous research investigating parental use of mental-state language (Taumoepeau & Ruffman, 2006, 2008). The photos displayed images of adults and children engaging in a variety of activities (e.g., going to the doctor, building with blocks, feeding ducks, etc.) and reflected a variety of emotions (e.g., anger, happiness, sadness). The pictures were placed into 20 cm x 28 cm clear plastic sleeves and assembled in a three-ring binder to create a picture book. The pages did not contain any words.

Parents were instructed to go through each picture with their child as though they were looking at a story at home. The experimenter then left the parent and child in a room alone until they finished looking at the images. The experimenter monitored the interaction through a baby monitor in the lobby.

3.3. Coding

3.3.1. Nontraditional task

All two-image trials were coded frame-by-frame from silent video by a trained coder who was blind to the side of the images in each trial and the hide-order that the child received. The coder indicated where children looked (left image, right image, away) during the 8s that the images were visible. The anticipatory-looking test trial was also coded frame-by-frame from silent video by a trained coder who was naïve to the hide-order that the child received. The coder indicated where children looked (left hiding location, right hiding location, agent, or away) during the 8s that the image was visible.

All test trials, as well as a randomly selected 50% of the setup trials, were coded by a second naïve coder. The two coders agreed on 95% of coded video frames (setup trials: 97%; anticipatory-looking trial: 95%; preferential-looking trial: 95%). Trials in which agreement was less than 85% were resolved by a third coder (7/304 coded trials: 2 setup trials, 3 anticipatory-looking trials, 2 preferential-looking trials). With the exception of 4 trials in which the third coder agreed with the second coder (1 setup trial, 2 anticipatory-looking trials, 1 preferential-looking trial), the primary coder's data were used in all analyses.

For the setup trials, we calculated the time that children spent looking at the matching and non-matching images, averaged across the 4 trials. We also computed children's preference for the matching image by subtracting their average looking time to the non-matching image from their average looking time to the matching image. Positive difference scores thus indicated that children successfully followed the verbal story.

For the anticipatory-looking trial, we calculated children's looking time to Jacob's original hiding location and Jacob's current hiding location during the 8-s trial. Consistent with Scott et al. (2012), the preferential-looking test trial was divided into a 2-s preview followed by a 6-s analysis window. For the 6-s analysis window, we computed children's looking times to each of the two images: Lucas searching in Jacob's original hiding location (original-location image) and Lucas searching in Jacob's current hiding location (current-location image). For both test trials, we calculated children's preference for the location/image that was consistent with Lucas' belief by subtracting their looking time to the belief-inconsistent location/image from their looking time to the belief-consistent location/image (false-belief condition: original location – current location; true-belief condition: current location – original location). Thus, for both test trials, positive difference scores indicated correct belief-based responses.

3.3.2. Picture-book task

Parent language was transcribed verbatim. Non-word sounds or exclamations (e.g., gasps, sighs, etc.) were not included in the transcription. Because this task was intended as a measure of parent language, child language was not transcribed or analyzed.

Table 1

Mental-state and Physical Words Used by Parents in the Picture-Book Task

Category	Words and Phrases
Cognition	concentrating, confused, curious, focus, forgot, guess, interested, know, pretend, remember, remind, think, trust, wonder
Desire	want
Emotion	angry, bored, care, cranky, disgusting, enjoying, excited, feeling, frustrated, fun, grumpy, happy, hurt, like, love, mad, miss, nervous, proud, sad, scare, shocked, surprised, upset, worry
Physical	asleep, awake, crying, hungry, hurt, laughing, sick, sleeping, smiling, tired

Note: All possible forms of a given word (e.g., *think, thinks, thinking*, etc.) were included

We used the CLAN program (MacWhinney, 2000) to identify parent utterances that included mental-state terms. Specifically, we examined the three categories of mental-state terms outlined by Ensor, Hughes, and colleagues (e.g., Ensor, Devine, Marks, & Hughes, 2014; Ensor

& Hughes, 2008): cognition, desire, and emotion. Table 1 shows a complete listing of the terms that parents used in each of these categories. The cognition category included words that referred to thoughts, knowledge, or memory. The most commonly used words in this category were *think* and *know* along with variations of these words. Given the frequency of these two cognition terms, as well as prior evidence that parents' use of these terms in particular predicts preschoolers' false-belief performance (e.g., Ruffman et al., 2002), we also created a separate *think/know* subcategory that included uses of these terms and their variations. The desire category included the word *want* and all the variations of this word. None of the other desire words or phrases identified by Hughes and colleagues (e.g., *wish, hope, would like, would love*) occurred in our transcripts. Finally, the emotion category included all words that referred to an emotional state, with the most frequent terms being *happy*, *sad*, and *mad*. Note that in prior studies by Ruffman and colleagues, *like* and *love* were coded as desire terms rather than emotion terms (e.g., Ruffman et al., 2002). If this alternative coding is used, the patterns of significant relationships between parent talk and performance on the nontraditional task described in the Results section remain the same.

Consistent with previous research on mental-state language (e.g., Bartsch & Wellman, 1995; Ensor & Hughes, 2008; Ruffman et al. 2002; Taumoepeau & Ruffman, 2008), parent utterances that were perfect repetitions of their child's utterance or that contained only the phrase "I don't know" (i.e. utterances consisting of only these 3 words that did not explain or expand upon what was unknown) were not counted as containing mental-state references.

A recent meta-analysis of the relationship between parent mental-state language and preschoolers' elicited-response performance found that studies that controlled for parent verbosity by analyzing the percentage of parent utterances that contained mental-state terms yielded more modest effect sizes than those that examined raw frequency of mental-state terms (Devine & Hughes, 2018). We therefore chose to examine the percentage of utterances that contained mental-state terms because this would provide a more conservative estimate of the relationship between parental mental-state language and toddlers' performance in the nontraditional task (analyses using the raw number of mental state terms produced the same pattern of significant results). For each parent, we calculated percentage scores for overall use of mental-state terms as well as separate percentage scores for each category of mental-state terms.

Our initial coding of mental-state terms did not distinguish between mental-state uses of these words (e.g., *The little girl likes kittens*) and alternative uses that did not refer to mental states (e.g., *That looks like a kitty to me*). Prior studies have differed in their approach to this issue: while some researchers exclude non-genuine or conversational uses (e.g., Bartsch & Wellman, 1995), others include all occurrences of these terms (e.g. Taumoepeau & Ruffman, 2008). It is possible that young children initially do not distinguish between mental-state and non-mental-state uses of the terms we coded, and thus any use of these words could potentially impact their performance on the nontraditional task. However, it is also possible that by 2.5 years of age, children do appropriately distinguish between these uses and that only genuine mental-state references predict their mental-state reasoning abilities.

To explore these possibilities, we performed a second coding pass to identify genuine mental-state uses of the terms in Table 1. CLAN was used to extract all utterances from each transcript that contained a mental-state word. Two coders independently reviewed each utterance and indicated whether it contained a genuine mental-state reference. Agreement between the two coders was 98.5% (Cohen's kappa = .98); disagreements were resolved through discussion. The term that was most frequently used in an alternative, non-mental-state fashion was *like*, which was used both in a comparative sense (e.g., *It looks like a frog*) and as a filler (e.g., *He's like uh surprised.*). For each parent, we calculated an additional set of scores reflecting the percentage of their utterances that contained genuine mental-state terms (overall, and for each category of terms).

Finally, we also coded the percentage of parent utterances that contained references to physical states such as *sick*, *crying*, *laughing*, and *smiling*. While these terms are often used to describe people and behavior, they focus on observable expressions of physical states rather than internal mental states. Prior research suggests that these terms are not related to preschoolers' performance on elicited-response tasks (e.g., Ruffman et al., 2002). These terms thus served as a measure of discriminant validity for any possible associations between toddlers' performance and parents' use of mental-state terms.

4. Results

We begin by analyzing children's performance on the nontraditional task. We then provide descriptive information regarding parents' language in the picture-description task. Finally, we examine the relationships between children's performance on the nontraditional task and parents' language in the picture-description task.

4.1. Performance on the Nontraditional Task

Preliminary analyses of children's looking times in each of the three trial types (setup, anticipatory-looking, preferential-looking) revealed no significant interactions of belief condition and image/location with any of the counterbalancing factors (setup-side condition, hide-order, side of matching image in the preferential-looking trial), all Fs < 1.20, all ps > .27. We therefore collapsed across these factors in all analyses.

4.1.1. Setup trials

We first examined children's performance during the setup trials. These trials were not intended as a measure of belief reasoning. Rather, they served as a measure of whether children successfully followed the verbal story that preceded the test trials. Preliminary analyses of children's looking to the matching and non-matching images in the setup trials revealed no interactions of belief condition and image with sex or older-sibling, all Fs < 1. We therefore collapsed across these factors in subsequent analyses of the setup trials. Children's preference for the matching image in the setup trials was not correlated with sibling-number, $r = -.12$, $p = .29$, but it was significantly positively correlated with their age, $r = .29$, $p = .01$ and vocabulary $r = .25$, $p = .03$. We therefore controlled for age and vocabulary in the following analyses of the setup trials.

Children's average looking times in the setup trials were analyzed using an analysis of covariance (ANCOVA) with image (matching, non-matching) as a within-subject factor, belief condition (false belief, true belief) as a between-subjects factor, and age and vocabulary as covariates. This analysis revealed a significant effect of image, $F(1,72) = 4.75$, $p = .03$, $\eta^2 = .06$: children looked longer to the matching image than the non-matching image (see Figure 1). There was no effect of belief condition and no interaction of belief condition with image, both Fs < 1. These findings suggest that children in both conditions successfully followed the verbal story during the setup trials.

4.1.2. Anticipatory-looking trial

Preliminary analyses indicated that children's preference for the belief-consistent location was not correlated with their age, vocabulary, sibling-number, or their preference for the matching image during the setup trials, all ps > .47. We therefore collapsed across these factors in all subsequent analyses of this measure.

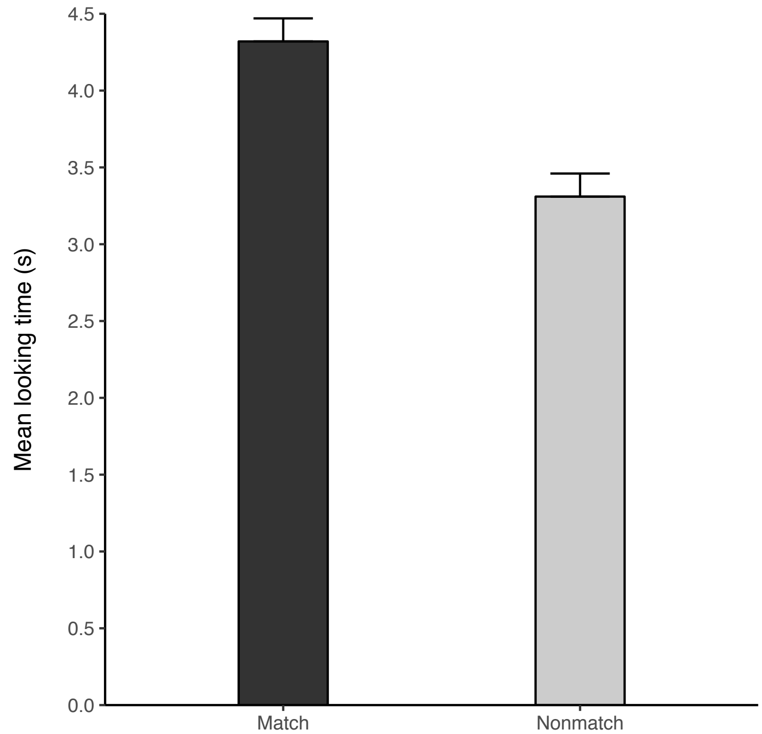


Figure 1. Results from the setup trials of the nontraditional task. Estimated marginal mean looking time in seconds to the matching and non-matching images, averaged across the four setup trials. Error bars represent one standard error of the mean.

Children's looking times to the two hiding locations during the anticipatory-looking trial were analyzed using an ANOVA with location (original, current) as a within-subject factor and belief condition (false belief, true belief), sex, and older-sibling as between-subjects factors. This analysis revealed no effect of location, $F(1, 68) = .04$, $p = .84$, $\eta^2 = .001$, and no interaction of belief condition and location, $F(1, 68) = .27$, $p = .61$, $\eta^2 = .004$. This suggests that as a group, children failed to show a significant preference for the belief-consistent location during the anticipatory-looking trial (see Figure 2). However, there was a significant three-way interaction of belief condition, location, and whether the child had an older sibling, $F(1, 68) = 6.71$, $p = .01$, $\eta^2 = .09$. Children with an older sibling ($M = .40$, $SD = 1.16$) showed a larger preference for the belief-consistent location than did children without an older sibling ($M = -.29$, $SD = 1.00$). There was also a marginal three-way interaction of belief condition, location, and sex, $F(1, 68) = 3.82$, $p = .06$, $\eta^2 = .05$. Exploration of this interaction effect showed that females ($M = .30$, $SD = 1.06$) demonstrated a larger preference for the belief-consistent location than did males ($M = -.14$, $SD = 1.18$). There were no other significant effects, all Fs < 2.32, all ps > .13.

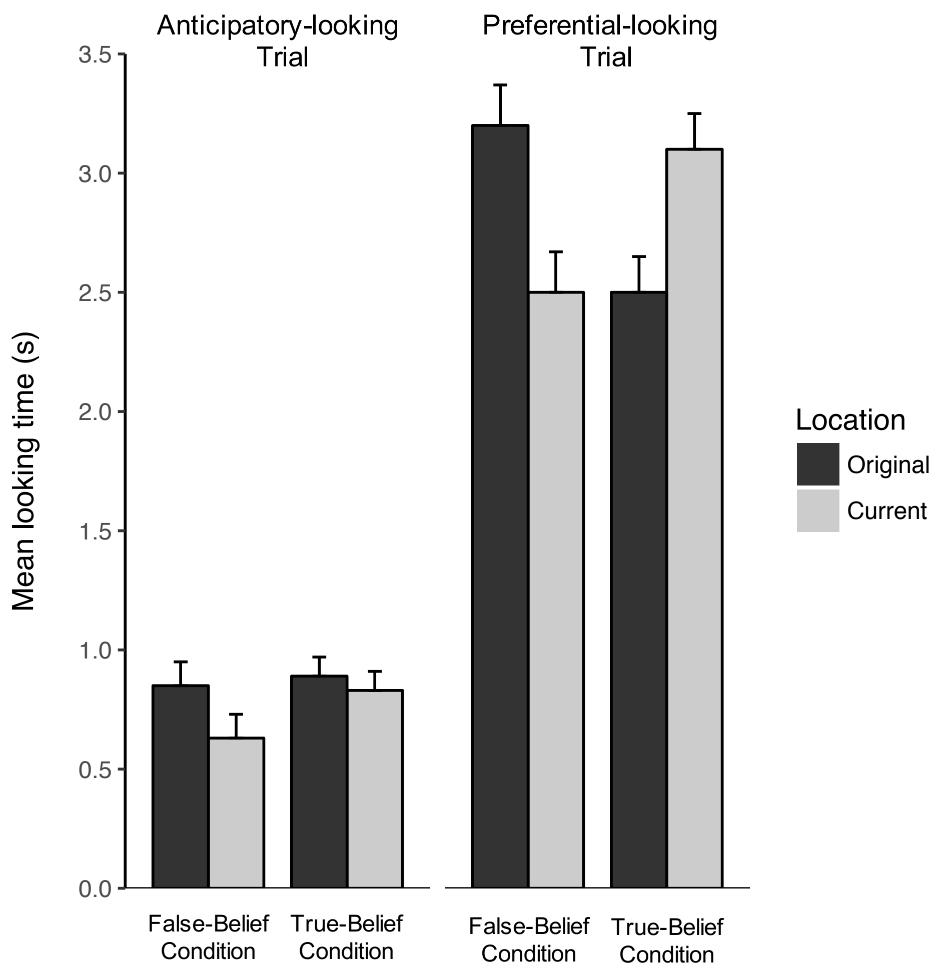


Figure 2. Results from the anticipatory-looking and preferential-looking trials of the nontraditional task. Estimated marginal mean looking time in seconds to the original-location and current-location, separately for the false-belief and true-belief conditions. Error bars represent one standard error of the mean.

4.1.3. Preferential-looking trial

Preliminary analyses of children's looking times to the two images in the preferential-looking trial revealed no significant interactions of belief condition and image with older-sibling, $F < 1$. We collapsed across this factor in subsequent analyses. Children's preference for the belief-consistent image in the preferential-looking trial was not correlated with their vocabulary, sibling-number, or their preference for the matching image during the setup trials, all *p*s > .28. However, children's preference for the belief-consistent location was significantly correlated with their age, $r = -.25$, $p = .03$. We therefore controlled for age in subsequent analyses of the preferential-looking trial.

An ANCOVA on children's looking times during the preferential-looking trial with image (original-location, current-location) as a within-subjects factor, condition (false belief, true belief) and sex as between-subject factors, and age as a covariate revealed a significant interaction of image and condition, $F(1, 71) = 8.17$, $p = .006$, $\eta^2 = .10$ (Figure 2). Planned simple effect comparisons revealed that children in the false-belief condition looked significantly longer at the original-location image than the current-location image, $F(1, 71) = 4.69$, $p = .03$, $\eta^2 = .06$. In contrast, children in the true-belief condition looked marginally longer at the current-location image than the original-location image, $F(1, 71) = 3.62$, $p = .06$, $\eta^2 = .05$. Thus, in both conditions children looked longer at the image that matched the final line of the story.

There was also a marginal three-way interaction of image, condition, and sex, $F(1, 71) = 3.64$, $p = .06$, $\eta^2 = .05$. Exploration of this interaction revealed that males ($M = 1.01$, $SD = 1.73$) exhibited a larger preference for the belief-consistent image than females ($M = .19$, $SD = 2.07$) during the preferential-looking trial. No other effects were significant, all *F*s < 1.20, all *p*s > .28

In an additional analysis, we examined whether the children in the two belief conditions differed in their preference for the belief-consistent image during the preferential-looking trial.

An ANCOVA on children's looking times with image (belief-consistent, belief-inconsistent) as a within-subjects factor, condition (false belief, true belief) and sex as between-subject factors, and age as a covariate revealed a significant effect of picture, $F(1, 71) = 6.52$, $p = .01$, $\eta^2 = .08$. Children looked longer at the belief-consistent ($M = 3.14$, $SD = 1.08$) than the belief-inconsistent image ($M = 2.54$, $SD = 1.03$). There was also a significant interaction of picture and sex, $F(1, 71) = 3.97$, $p = .05$, $\eta^2 = .05$. There was no effect of belief condition, $F(1, 71) = .06$, $p = .81$, $\eta^2 = .001$ and no interaction of belief condition and image, $F(1, 71) = .34$, $p = .56$, $\eta^2 = .005$. There were no other significant effects, all $Fs < 1.45$, all $ps > .23$.

These results indicate that children's preference for the belief-consistent image did not differ across belief conditions. When Lucas witnessed the first but not the second hiding event, children attributed to him a false belief that Jacob was in his original hiding location; upon hearing, "Lucas is looking for Jacob," they looked longer at the image that showed Lucas acting on this false belief and searching in Jacob's original hiding location. In contrast, when Lucas witnessed the second rather than the first hiding event, children attributed to him a true belief and looked longer at the image in which Lucas searched for Jacob in his current hiding location.

Together these results suggest that children followed the story, tracked Lucas' belief about Jacob's location, and were able to use this belief to identify the belief-consistent image in the preferential-looking trial. However, as a group, children failed to use Lucas' belief to anticipate his search behavior in the anticipatory-looking trial.

4.2. Parental Talk in the Picture-book Task

Table 2 shows descriptive statistics for parent talk in the picture-book task. Parents used emotion terms more frequently than any of the other types of term. Comparison of parents' percentage scores for overall uses and genuine uses of mental-state terms revealed significant differences for cognition, $t(75) = 3.01$, $p = .004$, emotion, $t(75) = 8.78$, $p < .001$, and total use of mental-state language, $t(75) = 8.80$, $p < .001$. In each case, parents' scores were lower when only genuine mental-state uses were considered. No significant differences were found for parents' use of desire terms or the *think/know* subcategory.

Table 2

Descriptive Statistics for Parental Talk in the Picture-book Task

	Mean		**Range**	
Percentage of Utterances				
Mental-state terms	Overall uses	Genuine uses	Overall uses	Genuine uses
Cognition	4.08 (3.76)	3.94 (3.70)	0 – 16.42	0 – 16.42
Think/Know	3.39 (3.42)	3.37 (3.41)	0 – 12.22	0 – 12.22
Desire	1.42 (1.90)	1.42 (1.90)	0 – 10.15	0 – 10.15
Emotion	13.63 (6.54)	10.71 (6.10)	1.56 – 39.13	1.18 – 39.13
Total	18.20 (7.74)	15.30 (6.91)	1.56 – 39.13	1.56 – 39.13
Physical terms	6.49 (4.03)		0 – 22.00	
Number of Parental Words and Utterances				
Number of Utterances	128.13 (57.67)		13 – 299	
Number of Words	503.45 (242.99)		40 – 1163	

As shown in Table 3, parents' use of cognition terms (both overall and genuine), was significantly positively correlated with their child's vocabulary; this relationship also held for the *think/know* subcategory. Parents' total use of mental-state terms was also positively correlated with their children's vocabulary, although this relationship was only marginal for genuine mental-state uses. There were no significant correlations between parent mental-state language and child age. Finally, we conducted a multivariate analysis of variance (MANOVA) with percentage of utterances that contained cognition, desire, and emotion terms as dependent variables and child sex (male, female) as a between-subjects factor. The effect of sex was not significant for either overall uses or genuine uses, both *Fs* < 1. A MANOVA with the number of words and utterances that parents used in the task as dependent variables also revealed no significant effect of sex, $F(2, 73) = 1.20$, $p = .31$. Thus, boys and girls heard relatively equal amounts of talk, and equal amounts of mental-state talk, in our picture-book task.

Table 3

Correlations Between Children's Age, Children's Vocabulary, and the Percentage of Parent Utterances Containing Mental-state or Physical Terms

	Child Age	**Child Vocabulary**
All mental-state terms		
Cognition	.11	.31**
Think/Know	.08	.31**
Desire	-.04	-.14
Emotion	.05	.16
Total	.06	.23*
Genuine mental-state terms		
Cognition	.11	.31**
Think/Know	.08	.30**
Desire	-.04	-.14
Emotion	.04	.09
Total	.06	.19^
Physical terms	-.08	-.05

Note: **$p < .01$; *$p < .05$; ^ $p < .1$

4.3. Relationships Between the Nontraditional Task and Parent Talk

We obtained the same patterns of significant correlations between children's performance and parent talk for both overall and genuine uses of mental-state terms (see Table 4). Therefore, we focus here on parents' overall uses of mental-state language for the sake of brevity.

Table 4

Bivariate Correlations Between Children's Difference Scores in the Nontraditional Task and the Percentage of Parent Utterances Containing Mental-state or Physical Terms

	Anticipatory-looking trial	**Preferential-looking trial**
All mental-state terms		
Cognition	.26*	.02
Think/Know	.29*	.02
Desire	-.04	.06
Emotion	.13	.14
Total	.20^	.17
Genuine mental-state terms		
Cognition	.26*	.01
Think/Know	.28*	.02
Desire	-.04	.06
Emotion	.14	.01
Total	.22^	.07
Physical-state terms	.15	.05

Note: * $p < .05$, ^ $p < .10$

4.3.1. Anticipatory-looking trial

Table 4 shows bivariate correlations between children's difference scores for the anticipatory-looking trial and parental mental-state language. Children's performance was significantly positively correlated with the percentage of parent utterances containing cognition terms; the same pattern was found for the *think/know* subcategory. Children's performance was marginally positively correlated with the total percentage of parent utterances that contained mental-state terms. This suggests that children who heard more mental-state terms, especially *think* and *know*, were better able to anticipate Lucas' behavior in the anticipatory-looking trial. In contrast, children's performance was not correlated with the percentage of parent utterances that contained physical terms, suggesting that the relationship between children's performance and parent talk was specific to parents' use of mental-state language.

Table 5

Hierarchical Multiple Regression Predicting Children's Anticipatory-Looking Difference Scores in the Nontraditional Task from Sex, Older-Sibling, Belief Condition, and Percentage of Parent Utterances Containing Cognition Terms (Overall Uses)

Predictor	ΔR^2	ß
Step 1	.14*	
Sex		.21^
Older-sibling		.32*
Belief condition		.02
Step 2	.06*	
Sex		.18^
Older-sibling		.33**
Belief condition		.06
Cognition terms		.25*
Step 3	.00	
Sex		.18^
Older-sibling		.34**
Belief condition		.07
Cognition terms		.25*
Belief condition X Cognition terms		-.02

Note: $^{**}p < .01$; $^{*}p < .05$; ^ $p < .1$

Recall that children's performance in the anticipatory-looking trial differed based on sex and whether the child had an older sibling. To determine whether parents' use of mental-state language contributed to children's performance above and beyond these other factors, we performed a series of hierarchical multiple regressions with children's difference scores in the anticipatory-looking trial as the dependent variable. In each model, sex (male, female), whether the child had an older-sibling (yes, no), and belief condition (false belief, true belief) were entered at Step 1. As shown in Table 5, these accounted for 14% of the variance in children's anticipatory-looking difference scores, $F(3, 72) = 3.75$, $p = .02$.

Table 6

Hierarchical Multiple Regression Predicting Children's Anticipatory-Looking Difference Scores in the Nontraditional Task from Sex, Older-Sibling, Belief Condition, and Percentage of Parent Utterances Containing Think/Know (Overall Uses)

Predictor	ΔR^2	ß
Step 1	.14*	
Sex		.21^
Older-sibling		.32*
Belief condition		.02
Step 2	.06*	
Sex		.18
Older-sibling		.31*
Belief condition		.05
Think/Know		.24*
Step 3	.001	
Sex		.18
Older-sibling		.31*
Belief condition		.09
Think/Know		.24*
Belief condition X Think/Know		-.04

Note: *$p < .05$; ^ $p < .1$

In the first model, we entered the percentage of utterances containing cognition terms at Step 2 (Table 5). This variable accounted for an additional 6% of the variance in children's anticipatory-looking difference scores, $F_{change}(1, 71) = 5.23$, $p = .025$. Children who heard a higher percentage of cognition terms exhibited a stronger preference for the belief-consistent location. To determine whether this effect differed across belief conditions, we next entered the interaction term (belief condition X cognition terms) at Step 3. The addition of this term did not significantly increase the variance accounted for, $F_{change}(1, 70) = .008$, $p = .927$. This suggests that the relationship between parents' use of cognition terms and children's anticipatory-looking performance did not differ across belief conditions: children who heard more cognition terms were better able to anticipate Lucas' behavior, regardless of whether he held a true or false belief about Jacob's location.

The same pattern emerged in a second model in which the percentage of utterances that contained the terms *think* and *know* was entered at Step 2 (see Table 6): parents use of these terms accounted for an additional 6% of the variance in children's anticipatory-looking difference scores, $F(1, 71) = 5.03$, $p = .03$. Once again, the interaction term (belief condition x *think/know*) was not significant when added at Step 3, $F_{change}(1, 70) = .081$, $p = .777$.

Table 7

Hierarchical Multiple Regression Predicting Children's Anticipatory-Looking Difference Scores in the Nontraditional Task from Sex, Older-Sibling, Belief Condition, and Percentage of Parent Utterances Containing Total Mental-State Terms (Overall Uses)

Predictor	ΔR^2	ß
Step 1	.14*	
Sex		.21^
Older-sibling		.32*
Belief condition		.02
Step 2	.03	
Sex		.21^
Older-sibling		.30*
Belief condition		.01
Total		.17

Note: *$p < .05$; ^ $p < .1$

Finally, we entered the total percentage of utterances containing mental-state terms at Step 2 (see Table 7). Although this term accounted for an additional 3% of the variance in children's anticipatory-looking difference scores, this did not reach significance, $F_{change}(1, 71) =$

2.53, p = .12. This suggests that parent's total mental-state talk did not predict children's anticipatory-looking performance.

4.3.2. Preferential-looking trial

As shown in Table 4, there were no significant relationships between children's difference scores for the preferential-looking trial and parental mental-state language. This suggests that parent use of mental-state language did not predict children's ability to identify the belief-consistent image in the preferential-looking trial. We therefore did not analyze this relationship further.

5. General Discussion

Substantial evidence suggests that preschooler's performance on traditional false-belief tasks is related to a variety of social factors, including parental use of mental-state language (e.g., Devine & Hughes, 2018). The present study investigated whether this relationship extended to younger children's performance on nontraditional false-belief tasks. We addressed this question by testing 2.5-year-old children in a verbal nontraditional task that included two looking-time measures, anticipatory looking and preferential looking, and assessing their parent's use of mental-state language in a picture-book task.

Results revealed that parental use of mental-state language positively predicted children's performance in the anticipatory-looking trial of the nontraditional task; this relationship did not vary as a function of whether the agent's belief was true or false. This relationship was specific to parents' use of cognition terms, especially *think* and *know*. This is consistent with prior findings that cognition terms are especially predictive of preschoolers' performance on elicited-response false-belief tasks (e.g., Adrián et al., 2007; Howard et al., 2008; Ruffman et al., 2002). Our findings constitute the first evidence that this same facet of parental mental-state language is related to children's performance on a nontraditional false-belief task, as well as the first evidence of a relationship between parental mental-state language and false-belief understanding prior to age 3. These results add to the recent findings that deaf infants of hearing parents, who hear fewer mental-state references than hearing children of hearing parents (Morgan et al., 2014), exhibit deficits on nontraditional false-belief tasks (Meristo et al., 2012; Meristo et al., 2016). To our knowledge, our results also provide the first evidence that parental mental-state language shows similar patterns of relationships with both false-belief and true-belief conditions of the same task. Together, these findings suggest that the relationship between mental-state language and belief understanding is not specific to performance on elicited-response false-belief tasks during the preschool years. Instead, mental-state language appears to relate broadly to mental-state understanding, including false-belief understanding, in early childhood.

Our study also revealed a relationship between performance on the anticipatory-looking trial and a second social factor: whether or not the child had an older sibling. This factor was included in an exploratory fashion because preschoolers with siblings, in particular older siblings, show superior performance on elicited-response false-belief tasks (e.g., Devine & Hughes, 2018; Lewis et al., 1996; McAlister & Peterson, 2007; McAlister & Peterson, 2013; Perner et al., 1994). Consistent with these prior findings, children who had an older sibling performed better on the anticipatory-looking trial of the nontraditional task than children without an older sibling. This finding thus provides additional support for the notion that performance on nontraditional tasks is related to aspects of children's social experience.

In contrast to the anticipatory-looking trial, children's performance in the preferential-looking trial was not related to parent's use of mental-state language. Children succeeded as a group in the preferential-looking trial, looking longer at the image that was consistent with the agent's belief. Children's successful performance in the preferential-looking trial thus replicates the findings from Scott et al. (2012) and provides additional evidence that children can demonstrate belief understanding in nontraditional tasks prior to age 4 (e.g., Scott, 2017; Scott & Baillargeon, 2017).

One unexpected finding was the presence of marginal sex effects in both test trials: females showed better performance than males on the anticipatory-looking measure, whereas males performed better than females on the preferential-looking measure. One possibility is that females were faster to predict the behavior of the agent in the anticipatory-looking trial, resulting in less interest or persistence in the preferential-looking trial. However, we raise this possibility with some hesitation because these effects were marginal and no prior studies using nontraditional tasks have reported sex effects. It is therefore possible that this effect represents sampling error rather than genuine sex differences in early belief understanding.

Why might parental mental-state language have been related to performance in the anticipatory-looking trial but not the preferential-looking trial? There are at least three possible explanations for this pattern of findings. First, children succeeded as a group in the preferential-looking but not the anticipatory-looking trial, suggesting that the anticipatory-looking trial was globally more difficult. Perhaps children who regularly engage in conversations about mental states are better at engaging in belief reasoning under challenging conditions and hence performed better on this more difficult trial. Second, children who frequently hear mental-state language might be faster at retrieving beliefs when necessary. In our task, the anticipatory-looking trial always preceded preferential-looking trial. Perhaps children who heard more cognitive talk retrieved the agent's belief more quickly, allowing them to demonstrate belief understanding in the first test trial, whereas children who heard less cognitive talk took longer to retrieve the agent's belief and hence only succeeded in the second test trial. Finally, it is possible that hearing more cognitive talk is specifically related to children's ability to predict agents' actions. Because the preferential-looking trial did not involve prediction, and instead required a post hoc analysis of the agent's behavior, children's performance on this measure was not correlated with cognitive talk.

With regards to this last possibility, one might argue that anticipatory looking is a better or more exact measure of mental-state understanding than preferential looking because it requires prediction rather than looking longer at a particular action after it has occurred. If so, then perhaps the preferential-looking trial was not correlated with parental mental-state language because it does not involve the same mental-state understanding as the anticipatory-looking trial. Although possible, we are skeptical that either of our measures is better or more revealing of children's genuine mental-state understanding than the other. After all, traditional elicited-response tasks vary in whether children have to predict how an agent will act (Baron-Cohen et al., 1985; Wimmer & Perner, 1983) or explain an agent's actions after they occur (Bartsch & Wellman, 1989; Dunn, Brown, Slomkowski, Tesla, & Youngblade, 1991). Both measures are regarded as indicative of false-belief understanding. Likewise, infants and toddlers succeed on a range of nontraditional tasks, some involving prediction (e.g., He et al., 2012; Moll et al., 2016; Surian & Geraci, 2012) and others involving post-hoc interpretation (Buttelmann et al., 2009; Onishi & Baillargeon, 2005; Scott et al., 2012).

For these reasons, we argue that both types of measures assess mental-state understanding, and that the lack of a correlation between parental mental-state language and preferential-looking performance is due to one of the three alternatives outlined above. Additional research is needed to disentangle these possibilities. If our pattern of results reflects differences in the relative difficulty or order of the two trials, then one might observe relationships between parental mental-state language and 2.5-year-olds' performance on a more difficult preferential-looking task or one that involves only a single test trial. If, however, engaging in conversations about mental states is specifically related children's ability to predict agents' actions, then parental mental-state language should not correlate with preferential-looking performance even under more challenging conditions.

Finally, an alternative, less-interesting explanation for our findings is that children whose parents used more mental-state talk were simply better at comprehending the verbal story in our task. We find this explanation unlikely for several reasons. First, if this were the case, then parental mental-state talk should have been related to performance on both test trials, since both required comprehension of the verbal story. Yet parental mental-state talk was not related to performance on the preferential-looking trial. Moreover, children looked longer at the matching than the non-matching image during the setup trials, suggesting that they were capable of following the story. Second, this would not explain why a second social factor, whether or not children had an older sibling, would also be associated with children's performance in our task. Third, this interpretation also cannot explain previous evidence that deaf infants show deficits in false-belief performance on an entirely nonverbal task (Meristo et al., 2012). Finally, data from our laboratory suggests that similar associations emerge between parental use of mental-state language and toddlers' performance on a completely non-verbal nontraditional false-belief task (Roby & Scott, 2016a). This suggests that it was not the verbal nature of the task that accounted for the relationships we found between children's performance and parental use of mental-state language. It seems more likely that social factors support belief reasoning, as they do in the context of elicited-response false-belief tasks.

5.1. Expression vs. emergence

In Section 1, we discussed ways in which parental mental-state language might influence the expression of belief understanding during the preschool years. Similarly, in the preceding section we offered several interpretations of our findings, all of which assume that parental mental-state language influences the expression, rather than the emergence, of toddlers' false-belief understanding. This is because we assume that nontraditional tasks, like traditional elicited-response tasks, assess the capacity to represent beliefs and hence positive results from these tasks indicate that the capacity to represent beliefs emerges early in infancy. If this is the case, then it is unlikely that social factors facilitate the emergence of belief representation during the preschool or toddler years.

However, the interpretation of the findings from nontraditional tasks has been the subject of considerable debate. Some researchers have argued that children's responses in nontraditional tasks do not reflect belief understanding. Instead, they suggest that children's performance on nontraditional tasks is driven by more rudimentary capacities, such as responses to perceptual novelty (e.g., Heyes, 2014), learned behavioral rules (e.g., Ruffman, 2014), or an early-developing system for tracking belief-like states (e.g., Butterfill & Apperly, 2013). These accounts maintain that the ability to represent beliefs does not emerge until the preschool years, as indicated by successful performance on elicited-response tasks. Advocates of such accounts would likely argue that our results do not speak to either the expression or emergence of false-belief understanding during the toddler years. Rather, our results merely indicate that parental mental-state language is related to some aspect of toddlers' performance on nontraditional tasks, such as how children respond to low-level properties of the scenes (e.g., Heyes, 2014) or the types of behavioral rules they bring to bear in our task (e.g., Ruffman, 2014). This perspective is consistent with the notion that social factors facilitate the emergence of the capacity to represent beliefs during the preschool years, leading to successful performance on elicited-response false-belief tasks (e.g. Heyes & Frith, 2014; de Villiers, 2005; Ruffman, 2014).

Other researchers have suggested that nontraditional and elicited-response tasks assess implicit and explicit forms of belief representation, respectively (e.g., Low, 2010; San Juan & Astington, 2012, 2017). This view assumes some continuity between the two forms of representation, with children gradually transitioning from implicit to explicit false-belief reasoning. Mental-state language has been argued to play a key role in this transition (e.g., San Juan & Astington, 2012, 2017). From this perspective, the relationship we observed between children's anticipatory-looking performance and parental use of mental-state language could indicate that social factors impact the expression of implicit belief understanding. Alternatively, our results might reflect the transition between implicit and explicit representational capacities, with mental-state language facilitating the emergence of the latter form of false-belief reasoning. Thus, our results might indicate that mental-state language plays a role in both the expression and the emergence of belief understanding.

We have argued against various alternative interpretations of nontraditional tasks elsewhere (e.g., Scott, 2017; Scott & Baillargeon, 2014, 2017; Scott et al., 2017), and thus we prefer an expression interpretation of our current findings. However, we acknowledge that additional evidence is needed to test these alternative interpretations and clarify the nature of the relationship between social factors and toddlers' performance on nontraditional belief tasks. We also note that even if an expression interpretation of the current findings is correct, this does not preclude the possibility that social factors support the emergence of belief understanding at an earlier point in time. It is possible that social experiences in the early months of life have an essential impact on the emergence of mental-state understanding. That said, recent neurological studies have demonstrated belief representation abilities in 6- to 8-month-olds (e.g., Hyde et al., in press; Kampis et al., 2015; Southgate & Vernetti, 2014) who are unlikely to appreciate or comprehend mental-state terms such as *think* and *know*. Thus, if parent-child interactions do play a role in the emergence of belief understanding at an early age, they would have to do so through via factors other than parental mental-state language.

5.2. Replication of findings from nontraditional tasks

In their recent review, Scott and Baillargeon (2017) identified over 30 published studies that reported positive evidence of false-belief understanding in children less than 3 years of age. Additional positive results have since been produced (e.g., Hyde et al., in press), including the findings from the preferential-looking trial of the present study. This accumulating body of evidence comes from a broad array of nontraditional tasks that assess a range of behavioral and neurological responses. Yet several recent studies have reported negative findings in nontraditional tasks (e.g., Crivello & Poulin-Dubois, in press; Dörrenberg et al., in press; Kulke, Reiß, Krist, Rakoczy, in press; Powell et al., in press), leading some researchers to question whether the findings from such tasks are reliable. Such researchers might view children's performance in the anticipatory-looking trial of the current study as a non-replication of prior anticipatory-looking results, providing another demonstration that the findings from such tasks are not robust.

We suspect that recent negative findings in nontraditional tasks stem, at least in part, from procedural differences across paradigms. Many psychological methods are sensitive to procedural variations – small changes in things like the nature or timing of stimuli can impact the results. This has been well documented for elicited-response false-belief tasks, where subtle changes, such as the addition of the word *first* to the test question (e.g., Siegal & Beattie, 1991; Yazdi, German, Defeyter, & Siegal, 2006), impact the age at which children succeed (e.g., Bartsch, 1996; Chandler, Fritz, & Hala, 1989; Friedman & Leslie, 2005; Hansen, 2010; Lewis & Osborne, 1990; Mitchell & Lacohée, 1991; Rubio-Fernández & Geurts, 2013; Setoh, Scott, & Baillargeon, 2016; Westra, 2017). Nontraditional tasks are no exception, as we have demonstrated in our own work: modifications in the linguistic ambiguity of a false-belief story impacted whether 3-year-olds succeeded or failed as a group in a preferential-looking task, and whether or not their performance was correlated with their verbal ability (Scott & Roby, 2015). Thus, careful examination of negative findings from nontraditional tasks should help identify when and how procedural variations impact performance (see Baillargeon, Buttelmann, & Southgate, in press).

With regards to the present study, our task was not designed as a direct replication of any prior anticipatory-looking or preferential-looking task. On the contrary, we deliberately attempted to increase the demands of our task because our prior work suggested we would be more likely to observe relationships between social factors and children's performance under challenging conditions (e.g., Scott & Roby, 2015). For this reason, we feel it is inappropriate to treat children's performance in the anticipatory-looking trial as a non-replication of any prior anticipatory-looking task.

More importantly, our results suggest that although as a group children looked equally at the two hiding locations during the anticipatory-looking trial, this was not due to chance or random responding. Rather, children's tendency to look at the belief-consistent location varied systematically with the amount of cognition talk that their parents produced, as well as whether they had an older sibling. This suggests that our task tapped into meaningful variation in children's belief reasoning skills. Our findings thus offer a second possible interpretation of recent negative findings in nontraditional tasks: perhaps at least some of these group-level failures are the product of interesting individual variability in young children's false-belief reasoning abilities. We hope that our results encourage researchers to examine this possibility more closely by including a variety of individual differences measures when studying infants' and toddlers' belief reasoning.

5.3. Concluding remarks

In the present research, we showed that 2.5-year-olds' performance on true- and false-belief conditions of a nontraditional task was related to their parents' use of mental-state language. These results are consistent with the robust associations that have been found between parental use of mental-state language and children's performance on traditional false-belief tasks during the preschool years. The current findings show that these associations extend to an earlier age range than previously shown and are apparent when utilizing a nontraditional false-belief task that measures anticipatory looking. Our findings thus add to a growing number of studies suggesting that exposure to and use of mental-state language is related to false-belief understanding across the lifespan (e.g., Bianco, Lecce, & Banerjee, 2016; Grazzani & Ornaghi, 2012; Lecce, Bianco, Devine, Hughes, & Banerjee, 2014; Lecce, Bottiroli, Bianco, Rosi, & Cavallini, 2015; Roby & Scott, 2016c). Our results also demonstrate that parental use of mental-state language is associated with both false- and true-belief understanding. Together with other recent findings (e.g., Drummond et al., 2014; Newton et al., 2016; Taumoepeau & Ruffman, 2008), these results suggest that parent use of mental-state language relates broadly to children's early social-cognitive skills.

Acknowledgements

This research was supported by a fellowship from the University of California Office of the President to ER and grants from the Hellman Fellows Fund and the University of California Merced Graduate Research Council to RMS. We thank the staff of the University of California Merced Center for Early Cognition and Language for their assistance with data collection, Renée Baillargeon and Eric Walle for helpful discussions, and the parents and children who participated in this research.

References

Adrián, J. E., Clemente, R. A., & Villanueva, L. (2007). Mothers' use of cognitive state verbs in picture-book reading and the development of children's understanding of mind: A longitudinal study. *Child Development*, *78*, 1052-1067.

Adrián, J. E., Clemente, R. A., Villanueva, L., & Rieffe C. (2005). Parent-child picture-book reading, mothers' mental state language and children's theory of mind*. Journal of Child Language, 32,* 673-686.

Baillargeon, R., Buttelmann, D., & Southgate, V. (in press). How to interpret failed replications of early false-belief findings: methodological and theoretical considerations. *Cognitive Development.*

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

Barrett, H. C., Broesch, T., Scott, R. M., He, Z., Baillargeon, R., Wu, D., ... & Laurence, S. (2013). Early false-belief understanding in traditional non-Western societies. *Proceedings of the Royal Society of London B: Biological Sciences*, *280*, 20122654.

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

Bartsch, K., & Wellman, H. (1989). Young children's attribution of action to beliefs and desires. *Child Development*, *60*, 946-964.

Bartsch, K., & Wellman, H. M. (1995). *Children talk about the mind*. Oxford, UK: Oxford University Press.

Bianco, F., Lecce, S., & Banerjee, R. (2016). Conversations about mental states and theory of mind development in middle childhood: A training study. *Journal of Experimental Child Psychology, 149*, 41-61.

Brown, J. R., Donelan-McCall, N., & Dunn, J. (1996). Why talk about mental states? The significance of children's conversations with friends, siblings, and mothers. *Child Development*, *67*, 836-849.

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

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

Carruthers, P. (2013). Mindreading in infancy. *Mind & Language*, *28*, 141-172.

Chandler, M., Fritz, A. S., & Hala, S. (1989). Small-scale deceit: Deception as a marker of two-, three-, and four-year-olds' early theories of mind. *Child Development*, *60*, 1263-1277.

Christensen, W., & Michael, J. (2016). From two systems to a multi-systems architecture for mindreading. *New Ideas in Psychology*, *40*, 48-64.

Clements, W. A., & Perner, J. (1994). Implicit understanding of belief. *Cognitive Development*, *9*, 377-395.

Clements, W. A., Rustin, C. L., & McCallum, S. (2000). Promoting the transition from implicit to explicit understanding: A training study of false belief. *Developmental Science*, *3*, 81-92.

Crivello, C. & Poulin-Dubois, D. (in press). Infants' false belief understanding: A non-replication of the helping task. *Cognitive Development*.

Cutting, A. L., & Dunn, J. (1999). Theory of mind, emotion understanding, language, and family background: Individual differences and interrelations. *Child Development*, *70*, 853-865.

de Villiers, J. G. (2005). Can language acquisition give children a point of view? In J. W. Astington and J. A Baird (Eds.), *Why language matters for theory of mind* (pp. 186-219). New York, NY: Oxford University Press.

Devine, R. T., & Hughes, C. (2018). Family correlates of false belief understanding in early childhood: A meta-analysis. *Child Development, 89,* 971-987.

Dörrenberg, S. Rakoczy, H., & Liszkowski, U., (in press). How (not) to measure infant Theory of Mind: Testing the replicability and validity of four non-verbal measures. *Cognitive Development*.

Drummond, J., Paul, E. F., Waugh, W. E., Hammond, S. I., & Brownell, C. A. (2014). Here, there and everywhere: Emotion and mental state talk in different social contexts predicts empathic helping in toddlers. *Frontiers in Psychology*, *5*, 361.

Dunn, J., Brown, J., Slomkowski, C., Tesla, C., & Youngblade, L. (1991). Young children's understanding of other people's feelings and beliefs: Individual differences and their antecedents. *Child Development*, *62*, 1352-1366.

Ensor, R., Devine, R. T., Marks, A., & Hughes, C. (2014). Mothers' cognitive references to 2-year-olds predict theory of mind at ages 6 and 10. *Child Development*, *85*, 1222-1235.

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

Fenson, L., Pethick, S., Renda, C., Cox, J. L., Dale, P. S., & Reznick, J. S. (2000). Short-form versions of the MacArthur communicative development inventories. *Applied Psycholinguistics*, *21,* 95–116.

Friedman, O., & Leslie, A. M. (2005). Processing demands in belief-desire reasoning: inhibition or general difficulty? *Developmental Science*, *8*, 218-225.

Gale, E., de Villiers, P., de Villiers, J., & Pyers, J. (1996). Language and theory of mind in oral deaf children. In A. Stringfellow, D. Cahana-Amitay, C. Hughes, and A. Zukowski (Eds.), *Proceedings of the 20th Annual Boston University Conference on Language Development* (pp. 213-224). Sommerville, MA: Cascadilla Press.

Garnham, W.A., & Perner, J. (2001). Actions really do speak louder than words—But only implicitly: Young children's understanding of false belief in action. *British Journal of Developmental Psychology*, *19*, 413-432.

Grazzani, I., & Ornaghi, V. (2012). How do use and comprehension of mental-state language relate to theory of mind in middle childhood? *Cognitive Development*, *27*, 99-111.

Hansen, M. B. (2010). If you know something, say something: Young children's problem with false beliefs. *Frontiers in Psychology*, *1,* 23.

He, Z., Bolz, M., & Baillargeon, R. (2012). 2.5-year-olds succeed at a verbal anticipatory-looking false-belief task. *British Journal of Developmental Psychology*, *30*, 14-29.

Helming, K. A., Strickland, B., & Jacob, P. (2016). Solving the puzzle about early belief-ascription. *Mind & Language*, *31*, 438-469.

Heyes, C. (2014). False belief in infancy: a fresh look. *Developmental Science, 17,* 647-659.

Heyes, C. M., & Frith, C. D. (2014). The cultural evolution of mind reading. *Science*, *344*, 1243091.

Howard, A. A., Mayeux, L., & Naigles, L. R., (2008). Conversational correlates of children's acquisition of mental verbs and a theory of mind. *First Language*, *28*, 375-402.

Hyde, D., Simon, C. E., Ting, F., & Nikolaeva, J. (in press). Functional organization of the temporal-parietal junction for theory of mind in preverbal infants: A near-infrared spectroscopy study. *The Journal of Neuroscience.*

Kampis, D., Fogd, D., & Kovács, Á. M. (2017). Nonverbal components of theory of mind in typical and atypical development. *Infant Behavior and Development*, *48*, 54-62.

Kampis, D., Parise, E., Csibra, G., & Kovács, Á. M. (2015). Neural signatures for sustaining object representations attributed to others in preverbal human infants. *Proceedings of the Royal Society B*, *282*, 20151683.

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.

Kulke, L., Reiß, M., Krist, H., & Rakoczy, H. (in press). How robust are anticipatory looking measures of Theory of Mind? Replication attempts across the life span. *Cognitive Development*.

LaBounty, J., Wellman, H. M., Olson, S., Lagattuta, K., & Liu, D. (2008). Mothers' and fathers' use of internal state talk with their young children. *Social Development*, *17*, 757-775.

Lecce, S., Bianco, F., Devine, R. T., Hughes, C., & Banerjee, R. (2014). Promoting theory of mind during middle childhood: A training program. *Journal of Experimental Child Psychology*, *126*, 52-67.

Lecce, S., Bottiroli, S., Bianco, F., Rosi, A., & Cavallini, E. (2015). Training older adults on Theory of Mind (ToM): Transfer on metamemory. *Archives of Gerontology and Geriatrics*, *60*, 217-226.

Lewis, C., Freeman, N. H., Kyriakidou, C., Maridaki-Kassotaki, K., & Berridge, D. M. (1996). Social influences on false belief access: Specific sibling influences or general apprenticeship? *Child Development*, *67*, 2930-2947.

Lewis, C., & Osborne, A. (1990). Three-year-olds' problems with false belief: Conceptual deficit or linguistic artifact? *Child Development*, *61*, 1514-1519.

Low, J. (2010). Preschoolers' implicit and explicit false-belief understanding: relations with complex syntactical mastery. *Child development*, *81*, 597-615.

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

MacWhinney, B. (2000). *The CHILDES Project: Tools for Analyzing Talk. 3rd Edition*. Mahwah, NJ: Lawrence Erlbaum Associates.

Mayer, A., & Träuble, B. E. (2013). Synchrony in the onset of mental state understanding across cultures? A study among children in Samoa. *International Journal of Behavioral Development*, *37*, 21-28.

McAlister, A., & Peterson, C. (2007). A longitudinal study of child siblings and theory of mind development. *Cognitive Development*, *22*, 258-270.

McAlister, A. R., & Peterson, C. C. (2013). Siblings, theory of mind, and executive functioning in children aged 3–6 years: New longitudinal evidence. *Child Development*, *84*, 1442-1458.

Meins, E., Fernyhough, C., Wainwright, R., Clark-Carter, D., Das Gupta, M., Fradley, E., & Tuckey, M. (2003). Pathways to understanding mind: Construct validity and predictive validity of maternal mind-mindedness. *Child Development*, *74*, 1194-1211.

Meristo, M., Falkman, K. W., Hjelmquist, E., Tedoldi, M., Surian, L., & Siegal, M. (2007). Language access and theory of mind reasoning: Evidence from deaf children in bilingual and oralist environments. *Developmental Psychology*, *43*, 1156.

Meristo, M., Morgan, G., Geraci, A., Iozzi, L., Hjelmquist, E., Surian, L., & Siegal, M. (2012). Belief attribution in deaf and hearing infants. *Developmental Science*, *15*, 633-640.

Meristo, M., Strid, K., & Hjelmquist, E. (2016). Early conversational environment enables spontaneous belief attribution in deaf children. *Cognition*, *157*, 139-145.

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

Moeller, M. P., & Schick, B. (2006). Relations between maternal input and theory of mind understanding in deaf children. *Child Development*, *77*, 751-766.

Moll, H., Kane, S., & McGowan, L. (2016). Three-year-olds express suspense when an agent approaches a scene with a false belief. *Developmental Science*, *19*, 208-220.

Morgan, G., Meristo, M., Mann, W., Hjelmquist, E., Surian, L., & Siegal, M. (2014). Mental state language and quality of conversational experience in deaf and hearing children. *Cognitive Development*, *29*, 41-49.

Newton, E. K., Thompson, R. A., & Goodman, M. (2016). Individual differences in toddlers' prosociality: Experiences in early relationships explain variability in prosocial behavior. *Child Development*, *87*, 1715-1726.

Nielsen, M., & Dissanayake, C. (2000). An investigation of pretend play, mental state terms and false belief understanding: In search of a metarepresentational link. *British Journal of Developmental Psychology*, *18*, 609-624.

Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, *308*, 255-258.

Perner, J., Ruffman, T., & Leekam, S. R. (1994). Theory of mind is contagious: You catch it from your sibs. *Child Development*, *65*, 1228-1238.

Peterson, C., & Slaughter, V. (2003). Opening windows into the mind: Mothers' preferences for mental state explanations and children's theory of mind. *Cognitive Development*, *18*, 399-429.

Powell, L. J., Hobbs, K., Bardis, A., Carey, S., & Saxe, R. (in press). Replications of implicit theory of mind tasks with varying representational demands. *Cognitive Development*.

Roby, E. M. & Scott, R. M. (2016a). Parent mental-state talk predicts 2.5-year-olds' performance on a non-verbal anticipatory-looking false-belief task. Poster presented at the International Conference on Infant Studies, New Orleans, Louisiana.

Roby, E., & Scott, R. M. (2016b). Rethinking the relationship between social experience and false-belief understanding: A mentalistic account. *Frontiers in Psychology*, *7,* 1721.

Roby, E. M. & Scott, R. M. (2016c). "Think Fast! Mental-state language is related to the speed of false-belief reasoning in adulthood," in *Proceedings of the 38th Annual Conference of the Cognitive Science Society* (pp. 2525-2530). Austin, TX: Cognitive Science Society.

Rubio-Fernández, P. & Geurts, B. (2013). How to pass the false-belief task before your fourth birthday. *Psychological Science*, *24*, 27-33.

Ruffman, T. (2014). To belief or not belief: Children's theory of mind. *Developmental Review*, *34*, 265-293.

Ruffman, T., Slade, L., & Crowe, E. (2002). The relation between children's and mothers' mental state language and theory-of-mind understanding. *Child Development*, *73*, 734-751.

San Juan, V., & Astington, J. W. (2012). Bridging the gap between implicit and explicit understanding: How language development promotes the processing and representation of false belief. *British Journal of Developmental Psychology*, *30*, 105-122.

San Juan, V., & Astington, J. W. (2017). Does language matter for implicit theory of mind? The effects of epistemic verb training on implicit and explicit false-belief understanding. *Cognitive Development*, *41*, 19-32.

Scott, R. M. (2017). The developmental origins of false-belief understanding. *Current Directions in Psychological Science*, *26*, 68-74.

Scott, R. M., & Baillargeon, R. (2014). How fresh a look? A reply to Heyes. *Developmental Science*, *17*, 660-664.

Scott, R. M., & Baillargeon, R. (2017). Early false-belief understanding. *Trends in Cognitive Sciences, 21*, 237–249.

Scott, R. M., He, Z., Baillargeon, R., & Cummins, D. (2012). False-belief understanding in 2.5-year-olds: Evidence from two novel verbal spontaneous-response tasks. *Developmental Science*, *15*, 181-193.

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

Scott, R. M., Roby, E., & Smith, M. (2017). False-belief understanding in the first years of life. In J. Kiverstein (Ed.), *The Routledge Handbook of Philosophy of the Social Mind* (pp. 152-171). New York, NY: Routledge.

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.

Setoh, P., Scott, R. M., & Baillargeon, R. (2016). 2.5-year-olds succeed at traditional false-belief tasks with reduced executive demands. *Proceedings of the National Academy of Sciences*, *113,* 13360-13365.

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

Southgate, V., Chevallier, C., & Csibra, G. (2010). Seventeen-month-olds appeal to false beliefs to interpret others' referential communication. *Developmental Science*, *13*, 907-912.

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

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

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

Surian, L. & Geraci, A. (2012). Where will the triangle look for it? Attributing false beliefs to geometric shapes at 17 months. *British Journal of Developmental Psychology, 30,* 30-44.

Symons, D. K., & Clark, S. E. (2000). A longitudinal study of mother-child relationships and theory of mind in the preschool period. *Social Development*, *9*, 3-23.

Taumoepeau, M. (2015). From talk to thought: Strength of ethnic identity and caregiver mental state talk predict social understanding in preschoolers. *Journal of Cross-Cultural Psychology*, *46*, 1169-1190.

Taumoepeau, M., & Ruffman, T. (2006). Mother and infant talk about mental states relates to desire language and emotion understanding. *Child Development*, *77*, 465-481.

Taumoepeau, M., & Ruffman, T. (2008). Stepping stones to others' minds: Maternal talk relates to child mental state language and emotion understanding at 15, 24, and 33 months. *Child Development*, *79*, 284-302.

Taumoepeau, M., & Reese, E. (2013). Maternal reminiscing, elaborative talk, and children's theory of mind: An intervention study. *First Language*, *33*, 388-410.

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.

Westra, E. (2017). Pragmatic development and the false belief task. *Review of Philosophy and Psychology*, *8*, 235-257.

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

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.

Yazdi, A. A., German, T. P., Defeyter, M. A., & Siegal, M. (2006). Competence and performance in belief-desire reasoning across two cultures: The truth, the whole truth and nothing but the truth about false belief? *Cognition*, *100*, 343-368.

Appendix A: Pictures and Script for the False-belief Condition of the Nontraditional Task

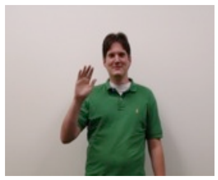
"This is a story about a boy named Lucas."

Introduction-1

"Look! There's Lucas!"

"Lucas has a friend named Jacob."

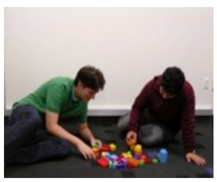
Introduction-2

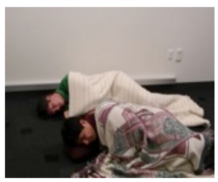


"Look! There's Jacob

"Lucas and Jacob like to play together."

Setup-1

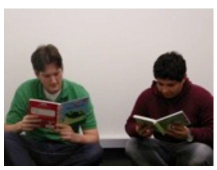


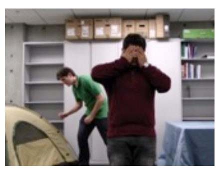


"Look! They're playing together."

"Lucas and Jacob are going to play hide and seek."

Setup-2

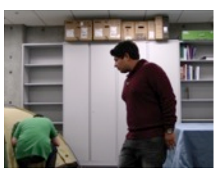


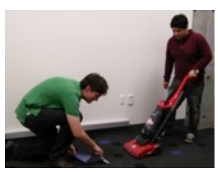


"See? Jacob is going to hide."

"Lucas peeks and sees Jacob go into the tent."

Setup-3

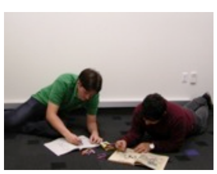




"See? Lucas is watching Jacob go into the tent."

"When Lucas closes his eyes Jacob comes out of the tent and goes under the table."

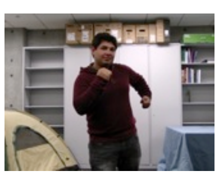
Setup-4



"See? Jacob is going under the table."

"Lucas is ready to look for Jacob."

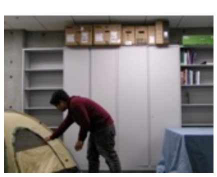
Anticipatory-Looking Trial

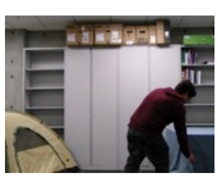


"Lucas says, 'Ready or not, here I come!'"

"Lucas goes to find Jacob.

Preferential-Looking Trial





"See? Lucas is looking for Jacob."

"Lucas is looking for Jacob."