Title
Think Fast! Mental-state Language is Related to the Speed of False-belief Reasoning in Adulthood

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 38(0)

Authors
Roby, Erin
Scott, Rose M.

Publication Date
2016

Peer reviewed

Think Fast! Mental-state Language is Related to the Speed of False-belief Reasoning in Adulthood

Erin Roby (eroby2@ucmerced.edu)
Psychological Sciences, 5200 Lake Rd.
Merced, CA 95343 USA

Rose M. Scott (rscott@ucmerced.edu)
Psychological Sciences, 5200 Lake Rd.
Merced, CA 95343 USA

Abstract

When tested appropriately, infants appear to demonstrate false-belief understanding in the first year of life. Some have argued that this is inconsistent with the well-established relationship between social experience and preschoolers' false-belief performance. We argue that these two sets of findings are not inconsistent because the ability to attribute false beliefs to others is necessary but not sufficient for false-belief performance, and we propose several ways that one social factor, hearing and using mental-state language, might relate to false-belief performance throughout the lifespan. We tested this account by examining the relationship between adults' use of mental-state language and their false-belief understanding. Participants' use of mental-state language was related to how quickly they could accurately predict the behavior of agents on the basis of desires and beliefs. These findings provide the first evidence that mental-state talk and false-belief performance are related into adulthood.

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

Introduction

As members of a social species, much of our everyday life involves predicting and interpreting the behavior of other individuals. Adults often do this by attributing to others unobservable mental states such as goals and beliefs. Researchers have long been interested in the nature and development of this psychological reasoning ability.

In particular, considerable research has focused on the ability to understand that others can be mistaken, or hold false beliefs, about the world. False-belief understanding provides evidence of the ability to recognize that mental states are internal representations of the world and thus they can be false. For several decades, false-belief understanding was primarily investigated using elicited-response tasks, which require children to answer direct questions about the behavior of a mistaken agent (for a review, see Wellman, Cross, & Watson, 2001). In one such task (Baron-Cohen, Leslie & Frith, 1985), children hear a story enacted with props: Sally puts a marble in a basket and then leaves; in her absence, Anne moves the marble to a nearby box. Children are then asked where Sally will look for her marble when she returns. Beginning around age 4, children typically indicate that Sally will look in the basket, where she falsely believes the marble to be. In contrast, younger children incorrectly respond that Sally will look in the marble's actual location. This widely replicated finding suggested that the capacity to attribute false beliefs to others did not emerge until at least age 4 (e.g., Wellman et al., 2001).

However, a growing body of evidence suggests that this ability might be present much earlier than previously thought: when tested via other means, infants appear to demonstrate false-belief understanding as early as 6 months of age (for a review, see Baillargeon, Scott, & Bian, 2016). For instance, infants visually anticipate where a mistaken agent will search for an object (e.g., Southgate, Senju, & Csibra, 2007; Surian & Geraci, 2012), look reliably longer when an agent's actions are inconsistent with her false belief (e.g., Onishi & Baillargeon, 2005; Träuble, Marinović, & Pauen, 2010), and use an agent's false belief to guide their own responses to that agent (e.g., Buttelmann, Carpenter, & Tomasello, 2009; Southgate, Chevallier, & Csibra, 2010). These positive findings have led many researchers to conclude that the capacity to attribute false beliefs to others emerges in the first year of life (e.g., Baillargeon et al., 2016; Buttelmann et al., 2009; Carruthers, 2013; Kovács, Téglás, & Endress, 2010, Southgate et al., 2007).

Some researchers have challenged accounts that ascribe false-belief understanding to infants on the grounds that they fail to address the role of social factors in the development of this ability (e.g., Ruffman, 2014; San Juan & Astington, 2012). Considerable research has shown that preschoolers' performance on elicited-response false-belief tasks is correlated with individual differences in their social experience (e.g., McAlister & Peterson, 2007; Meins et al., 2003; Ruffman, Slade & Crow, 2002). In particular, there is a robust relationship between preschooler's performance on elicited-response false-belief tasks and their exposure to and personal use of mental-state language – utterances that refer to psychological states such as *think, know,* and *believe* (e.g. Ensor & Hughes, 2008; Ruffman et al., 2002). It has been suggested that if the capacity to represent false beliefs is present in the first year of life, there is "little room for social factors to play a direct role" in children's false-belief performance (p. 110, San Juan & Astington, 2012), and thus the reason for these associations is unclear (e.g. Ruffman, 2014; San Juan & Astington, 2012).

However, many researchers have argued that the capacity to represent false beliefs does not guarantee successful performance in false-belief tasks (e.g., Baillargeon, Scott, & He, 2010; Bloom & German, 2000; Carruthers, 2013; Roth & Leslie, 1998). Studies have identified several factors that impact children's performance in both elicited-response and non-elicited-response false-belief tasks (e.g., Lewis & Osborne, 1990; Rubio-Fernández & Geurts, 2013; Scott & Roby, 2015; Yazdi et al., 2006) If, as these findings suggest, the capacity to represent beliefs is not sufficient for successful false-belief reasoning, then there is opportunity for social factors to impact false-belief performance.

Specifically, we propose that hearing and using mental-state language might be related to false-belief performance in several ways. First, individuals who frequently engage in conversations about mental states may more readily attend to and consider the mental states of others. Second, such conversations provide practice inhibiting one's own beliefs and desires in order to focus on those of another. Third, individuals who often discuss others' mental states may be more practiced at inferring others' mental states based on the (sometimes limited) information available. Finally, individuals who frequently hear and use mental-state language may more rapidly retrieve information about others' mental states from memory when necessary.

This account predicts that the relationship between mental-state talk and false-belief understanding should not be confined to performance on elicited-response tasks in the preschool years: it should be evident across the lifespan. Support for this prediction comes from recent evidence that parents' use of mental-state language is associated with 2.5-year-olds' performance on an anticipatory-looking false-belief task (Roby & Scott, 2015). In addition, deaf infants raised by hearing parents are exposed to significantly fewer mental-state references than their hearing counterparts (Morgan et al., 2014) and fail anticipatory-looking false-belief tasks (Meristo et al., 2012). These findings suggest that exposure to mental-state language is related to false-belief understanding prior to the preschool years.

Here we sought complementary evidence for this account by examining the relationship between personal use of mental-state language and false-belief performance in adults. To our knowledge, no prior studies have examined this relationship in adulthood. Although adults are certainly able to represent mental states, their mental-state reasoning exhibits considerable within- and between-individual variability (e.g., Brown-Schmidt, 2009; Harkness et al., 2010; McKinnon & Moscovitch, 2007). For instance, adults are faster at predicting the behavior of an agent who wishes to approach rather than avoid an object (e.g., Apperly et al., 2011), and adults from collectivist cultures exhibit faster and more accurate perspective taking than adults from individualistic cultures (e.g., Wu & Keysar, 2007). If, as we argue, personal use of mental-state language is related to greater attention to and consideration of others' mental states, then adults' use of such language should be related to their performance on false-belief tasks.

The present study tested this prediction. Adult participants described images of people or physical objects and then completed speeded and unspeeded mental-state reasoning tasks. We predicted that the extent to which participants used more mental-state language when discussing people than when describing objects (i.e. situations in which mental states were not relevant) would be related to their false-belief performance, and that this relationship might be most apparent when participants were required to quickly and accurately predict another individual's behavior under time pressure.

Method

Participants

71 adults (62 females) participated for course credit. An additional 4 adults were tested but excluded, 1 because she withdrew from the study, 2 because their vocabulary scores were below the 10th percentile, and 1 because he did not describe the images aloud in the picture-description task.

Materials and Procedure

All participants first completed a picture-description task. Stimuli consisted of 27 images of people or objects. Images were presented one at a time on a computer screen. Participants were instructed to describe aloud what they thought was happening in each picture. Once participants finished describing an image, they pressed a button on the keyboard to advance to the next image. Participants were told to take as much time as they needed to complete the task. During the task, audio and video were recorded to disk using the computer's built-in microphone and camera.

Participants next completed a speeded belief-desire task, a Strange Stories task, a stroop task, and a digit-span task in a randomized order.

The speeded belief-desire task was administered using PsychoPy (Peirce, 2007). On each trial, a short story about an agent was presented in the center of the computer screen. Once participants read the story, they pressed any key. The story was then replaced by a question about what the agent would do next. After 1s, two response options appeared, one in each of the bottom corners of the screen. Participants pressed either the left or right arrow key to select a response. Participants were told to respond as quickly and accurately as possible.

Participants first viewed two practice trials in which they read a story about an agent (e.g., Dennis loves pizza and hates vegetables. He is hungry and goes to the store.) and then answered a question about what the agent would do next (e.g., What will he buy?). Participants then completed 12 test trials; the stories in these trials were adapted from prior research (e.g., Bennett & Galpert, 1992; Leslie & Polizzi, 1998; Roby & Scott, 2015; Scott & Roby, 2015; Setoh, Scott, & Baillargeon, 2011; Yazdi et al., 2006). In each story, an agent held either a true or a false belief about the location of an object. In half the stories, the agent wanted to approach the object and in the other half the agent wanted to avoid the object. We varied the agent's desire because, as mentioned previously, adults find situations in which an agent wants to avoid an object more challenging to reason about than those in which an agent wants to approach an object (Apperly et al., 2011). We therefore expected that we would be maximally likely to see individual differences in performance in avoidance situations. Belief (true, false) and desire (approach, avoid) were crossed to create four story types; participants read three stories of each type.

The Strange Stories Task (White, Hill, Happé, 2009) consisted of 14 stories: 8 described social situations (mental-state stories), 3 described events with physical causes (e.g. an old woman slips on her icy doorstep, causing her to need an X-ray; physical stories), and 3 described natural phenomena (e.g. Sun melting snow causes the formation of puddles; nature stories). Stories were presented in one of four orders; each order was randomized with the constraint that no more than 2 stories of a given type occurred in a row. Once participants finished reading the story, it was removed and they were asked to provide a written response to a question about what happened in the story. For example, in one mental-state story, a nervous woman mistakes an innocent man for a robber. She tells him to take her purse and asks him not to hurt her. The question asked, "Why did she say that?" Participants were told to take as much time as they needed to complete the task.

The stroop task (Stroop, 1935) was administered using Inquisit 4.0.3 (2004). The task consisted of 56 experimental and 28 control trials. In experimental trials, participants saw one of 4 color words (red, blue, green, black); in half the trials the color of the text matched the word (congruent trials) while in the remaining trials it did not (incongruent trials). In control trials, participants saw a box displayed in one of the four colors. Participants were asked to respond to the color of the stimulus as quickly and accurately as possible by pressing a corresponding arrow key.

The digit-span task was also administered using Inquisit (e.g., Woods et al., 2011). On each trial, participants viewed a series of digits one at a time on the screen. The digits then disappeared, and after 1s a box appeared. Participants were instructed to type the digits they had seen in either forwards or backwards order using the keyboard. They completed 14 forward trials followed by 14 backwards trials.

The stroop and digit-span tasks were included to control for differences in inhibitory control and working memory, which correlate with adults' mental-state reasoning (e.g. German & Hehman, 2006, McKinnen & Moskovitch).

Finally, participants completed the Peabody Picture Vocabulary Task (PPVT-4; Dunn & Dunn, 2012). On each trial, participants indicated which of four images corresponded to a spoken word. This measure was included to control for differences in English fluency, which could impact participants' picture descriptions or comprehension of the stories in the mental-state reasoning tasks.

Coding and Analysis

Picture-description task We transcribed participants' descriptions of each image and then used the Linguistic Inquiry and Word Count (LIWC; Pennebaker, Booth, & Francis, 2007) program to identify utterances that included mental-state terms. These terms were divided into 5 categories: cognition (e.g., *think*, *know*), desire (e.g., *want*, *need*), emotion (e.g., *happy*, *angry*), modulations of assertion (e.g., *maybe*, *might*), and other mental-state words (e.g., *remember*, *wonder*). The mental-state words included words from previous work investigating parental mental-state talk during parent-child interactions (e.g., Ruffman et al., 2002) as well as synonyms for those words that are not generally used in child-directed speech (e.g., *devastated*, *pondering*). We omitted the word "like" because adults use it in several ways that are unrelated to mental states (e.g. "It's, like, two people at the beach").

For each participant, we calculated the percentage of utterances in each description that contained each category of word. For each category, we then averaged across all the pictures of a given type (i.e. people, objects). Next, we calculated difference scores for each category of mental-state talk by subtracting the percentage of mental-state utterances made in reference to object scenes from the percentage of mental-state utterances made in reference to images of people (see Table 1). This difference score thus reflected the extent to which participants selectively used mental-state language when discussing other individuals.

Table 1: Mean (SD) percentage of utterances in the picture-description task containing mental-state terms by image type, as well as difference scores (people – object)

	People	Object	Difference Score
Cognitive	5.6 (4.4)	4.8 (5.9)	.9 (6.0)
Desire	1.3 (1.7)	.2 (.6)	1.2 (1.8)
Emotion	11.2 (5.7)	.8 (2.2)	10.4 (5.7)
Modulations of assertion	22.3 (12.2)	18.0 (15.0)	4.3 (11.9)
Other	2.9 (3.0)	1.3 (2.9)	1.5 (3.2)
Total	36.2 (13.0)	23.8 (15.4)	12.5 (14.2)

Belief-desire Task One story was eliminated because more than 50% of participants responded incorrectly, suggesting a problem with the item. For the remaining 11 items, we eliminated responses with reaction times less than .15s or more than 3 standard deviations above the mean (n = 20). For each participant, we calculated the percentage of correct responses and average reaction time across all items, as well as separately for each story type (true-belief approach, true-belief avoid, false-belief approach, false-belief avoid).

Strange Stories Task Consistent with previous research (e.g. White et al., 2009), responses were scored on a 3-point scale. Participants received 0 points for responses that involved irrelevant or incorrect facts, 1 point for responses that were factually correct but failed to acknowledge key elements that were necessary for interpreting the story, and

2 points for fully correct responses that were factually correct and referenced the components of the story necessary for accurate interpretation. We calculated participant's average score for each story type (i.e. mental state, physical, and nature).

Inhibitory Control Trials in the stroop task with reaction times less than .15s or 2.5 standard deviations above the mean were excluded. We then subtracted participants' average reaction time on congruent trials from their average reaction time on incongruent trials.

Working Memory Consistent with prior work (German & Hehman, 2006), we used participants' backwards digit span (i.e. the highest number of digits participants recalled backwards correctly before producing two consecutive errors) as a measure of their working memory.

Receptive Vocabulary The PPVT was scored according to published procedures for this measure (Dunn & Dunn, 2012). Participants' raw scores were converted to standardized scores for analysis.

Results

Table 2: Mean (SD) percentage of correct responses and reaction times in seconds on the belief-desire task, separately by story type

	Percent Correct		Reaction Time	
	True Belief	False Belief	True Belief	False Belief
Approach	90 (21)	95 (30)	1.4 (.6)	1.4 (.7)
Avoid	79 (23)	78 (28)	1.7 (.9)	1.6 (.9)

Table 2 provides descriptive statistics for participants' performance in the belief-desire task. An ANOVA on the percentage of correct responses with desire (approach, avoid) and belief (true, false) as within-subject factors revealed a significant effect of desire, $F(1, 69) = 20.78$, $p <.001$. No other effects were significant, all Fs < 1. A corresponding ANOVA on participants' reaction times revealed a significant effect of desire, $F(1, 69) = 9.92$, $p = .002$; no other effects were significant, all Fs < 1. Consistent with prior findings (Apperly et al., 2011), participants were slower and less accurate when the agent wished to avoid rather than approach an object.

An ANOVA on participants' performance on the Strange Stories task with story type as a within-subject factor revealed no effect of story type, $F(2, 140) = 1.31$, $p = .27$. Participants performed equally well on mental-state stories ($M = 1.52$, $SD = .26$), physical stories ($M = 1.61$, $SD = .39$), and nature stories ($M = 1.54$, $SD = .34$).

To examine relationships between tasks, we computed partial correlations controlling for participants' inhibitory control, working memory, and receptive vocabulary. We report p-values corrected for multiple comparisons ($\alpha = .05$ one-tailed; Benjamini & Hochberg, 1995).

Participants' percentage of correct responses on the belief-desire task was significantly partially correlated with their performance on the mental-state stories, $r = .31$, $p = .03$, but not the physical stories, $r = .20$, $p = .15$, or nature stories, $r = .12$, $p = .25$. This suggests that the belief-desire task and mental-state stories assessed a common mental-state reasoning ability.

We next examined relationships between participants' mental-state language and their performance on the belief-desire task. We specifically examined correlations with the cognitive and total difference scores, as these categories are related to preschoolers' false-belief performance.

There were no significant relationships between participants' percentage of correct responses on the belief-desire task (overall or by story type) and their difference scores for either cognitive or total mental-state terms, all ps > .45. However, both difference scores were significantly negatively correlated with participants' reaction times on false-belief avoid stories (Table 3). Participants who used more cognitive and total mental-state terms in situations where mental states were relevant (to discuss people rather than objects) were significantly faster to respond correctly in situations where they needed to simultaneously consider an agent's false belief and desire to avoid an object.

Table 3: Partial correlations between mental-state difference scores and reaction times (correct responses only) on the belief-desire task

	True Belief		False Belief	
	Approach	Avoid	Approach	Avoid
Cognitive	-.09	-.17	-.07	-.28*
Total	.09	-.01	.00	-.25*

Note: In line with Apperly et al., 2011, only reaction times for correct responses were included. $df = 63$; * $p < .05$.

Finally, there were no significant relationships between participants' difference scores for cognitive or total mental-state terms and their performance on the Strange Stories task, all ps > .18.

General Discussion

Recent evidence suggests that the ability to attribute false beliefs to others may emerge in the first year of life. However, some have challenged this conclusion on the grounds that it is at odds with the well-established associations between social factors and preschoolers' false-belief performance (e.g. Ruffman, 2014; San Juan & Astington, 2012). If infants understand false beliefs, then why would social experience predict preschooler's performance on elicited-response false-belief tasks? We proposed that these two sets of findings are not inconsistent because the ability to attribute false beliefs to others is *necessary* but not *sufficient* for false belief performance, and thus social experience might very well facilitate the latter throughout the lifespan. For instance, individuals who hear and use more mental-state language might more readily attend to, infer, and retrieve others' mental states.

Here we tested this claim by examining the relationship between mental-state language and false-belief performance in adults. Participants described images of people and objects and then completed two mental-state reasoning tasks. Participants who exhibited a greater tendency to use mental-state terms (and cognitive terms in particular) when discussing people rather than objects were significantly faster at correctly predicting the behavior of an agent who held both a false belief and a desire to avoid an object.

These findings constitute the first evidence that personal use of mental-state language is related to false-belief performance in adulthood. The fact that this relationship was specific to false-belief avoidance situations likely reflects the fact that these situations impose greater cognitive load than those involving true beliefs or approach desires (e.g., Apperly et al., 2011; Leslie & Polizzi, 1998). In particular, false-belief avoidance stories require 'double inhibition': participants must inhibit their own knowledge about the location of an object in order to appreciate the agent's false belief while also recognizing the agent's desire to *not* approach that object. Individuals who engage in more discussions of others' mental states may have more practice inhibiting their own beliefs and desires while focusing on those of another, allowing them to more rapidly predict an agent's behavior in this particular context.

Adults' use of mental-state language was not related to the accuracy of their responses on either mental-state reasoning task. This suggests that use of mental-state language is related to how quickly one attends to, encodes, and retrieves mental-state information rather than one's ability to generate correct predictions or explanations based on that mental-state information. This might explain why several prior studies have failed to find associations between school-aged children's use of mental-state language and their performance on the Strange Stories task (e.g. Charman & Shmueli-Goetz, 1998; Meins, et al., 2006).

This pattern of findings is also consistent with recent work examining the relationship between parental mental-state talk and toddlers' false-belief performance. Roby and Scott (2015) tested 2.5-year-olds on a false-belief task with two consecutive test trials: an anticipatory-looking trial followed by a preferential-looking trial. Parental use of mental-state language predicted children's performance on the anticipatory-looking trial but not the preferential-looking trial. This difference could reflect the fact these two trials had different requirements for successful performance (prediction vs. post hoc analysis). However, our results suggest that this difference may instead reflect trial order. If exposure to mental-state language is related to how quickly children are able to attend to, encode, and retrieve mental-state information, then this would lead to stronger associations on the first test trial than the second.

Our findings complement recent findings suggesting that parental mental-state talk is associated with infants' and toddlers' false-belief performance (e.g. Meristo et al., 2012; Morgan et al., 2014; Roby & Scott, 2015). Together, these results suggest that the relationship between mental-state language and false-belief performance is not constrained to the preschool years or to elicited-response tasks: hearing and using mental-state talk facilitates the ability to reason about mental states throughout life.

References

Apperly, I. A., Warren, F., Andrews, B. J., Grant, J., & Todd, S. (2011). Developmental continuity in theory of mind: Speed and accuracy of belief–desire reasoning in children and adults. *Child Development*, *82*, 1691-1703.

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

Baillargeon, R., Scott, R. M., & Bian, L. (2016). Psychological reasoning in infancy. *Annual Review of Psychology*, *67*, 159-186.

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

Benjamini, Y. & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B, 57,* 289-300.

Bennett, M., & Galpert, L. (1992). Complex belief-desire reasoning in children. *Social Development*, *1*, 201-210.

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

Brown-Schmidt, S. (2009). The role of executive function in perspective taking during online language comprehension. *Psychonomic Bulletin & Review*, *16*, 893-900.

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

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

Charman, T., & Shmueli-Goetz, Y. (1998). The relationship between theory of mind, language and narrative discourse: An experimental study. *Cahiers de Psychologie Cognitive, 17,* 245-271.

Draine, S. C. (2004). Inquisit 2.0. 50401 [Computer software]. *Seattle, WA: Millisecond Software*.

Dunn, L. M., & Dunn, D. M. (2012). Peabody Picture Vocabulary Test, (PPVT-4). *Johannesburg: Pearson Education Inc*.

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

German, T. P., & Hehman, J. A. (2006). Representational and executive selection resources in 'theory of mind':

Evidence from compromised belief-desire reasoning in old age. *Cognition*, *101*, 129-152.

Harkness, K. L., Jacobson, J. A., Duong, D., & Sabbagh, M. A. (2010). Mental state decoding in past major depression: Effect of sad versus happy mood induction. *Cognition and Emotion*, *24*, 497-513.

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.

Leslie, A. M., & Polizzi, P. (1998). Inhibitory processing in the false belief task: Two conjectures. *Developmental Science*, *1*, 247-253.

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

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

McKinnon, M. C., & Moscovitch, M. (2007). Domain-general contributions to social reasoning: Theory of mind and deontic reasoning re-explored. *Cognition*, *102*, 179-218.

Meins, E., Fernyhough, C., Johnson, F., & Lidstone, J. (2006). Mind-mindedness in children: Individual differences in internal-state talk in middle childhood. *British Journal of Developmental Psychology*, *24*, 181-196.

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., 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.

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.

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

Peirce, JW (2007). PsychoPy - Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*, 8-13.

Pennebaker, J. W., Booth, R. J., & Francis, M. E. (2007). Linguistic inquiry and word count: LIWC [Computer software]. *Austin, TX: liwc.net*.

Roby, E. M. & Scott, R. M. (2015, March). *How does social input influence false-belief reasoning? Parent mental-state language and toddlers' false-belief understanding.* Paper presented at the biennial meeting of the Society for Research in Child Development, Philadelphia, PA.

Roth, D., & Leslie, A. M. (1998). Solving belief problems: Toward a task analysis. *Cognition*, *66*, 1-31.

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.

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.

Setoh, P. Scott, R. M., & Baillargeon, R. (March, 2011). *False-belief reasoning in 2.5-year-olds: evidence from an elicited-response low-inhibition task.* Paper presented at the biennial meeting of the Society for Research in Child Development, Montreal, Canada.

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.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643-662.

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

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

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.

White, S., Hill, E., Happé, F., & Frith, U. (2009). Revisiting the strange stories: revealing mentalizing impairments in autism. *Child Development*, *80*, 1097-1117.

Woods, D. L., Kishiyama, M. M., Yund, E. W., Herron, T. J., Edwards, B., Poliva, O., ... & Reed, B. (2011). Improving digit span assessment of short-term verbal memory. *Journal of Clinical and Experimental Neuropsychology*, *33*, 101-111.

Wu, S., & Keysar, B. (2007). The effect of culture on perspective taking. *Psychological Science*, *18*, 600-606.

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.