



Are Infants' Theory-of-Mind Abilities Well Integrated? Implicit Understanding of Intentions, Desires, and Beliefs

Jessica Yott & Diane Poulin-Dubois

To cite this article: Jessica Yott & Diane Poulin-Dubois (2016) Are Infants' Theory-of-Mind Abilities Well Integrated? Implicit Understanding of Intentions, Desires, and Beliefs, Journal of Cognition and Development, 17:5, 683-698, DOI: [10.1080/15248372.2015.1086771](https://doi.org/10.1080/15248372.2015.1086771)

To link to this article: <http://dx.doi.org/10.1080/15248372.2015.1086771>



Accepted author version posted online: 23 Feb 2016.
Published online: 23 Feb 2016.



[Submit your article to this journal](#)



Article views: 328



[View related articles](#)



[View Crossmark data](#)

Are Infants' Theory-of-Mind Abilities Well Integrated? Implicit Understanding of Intentions, Desires, and Beliefs

Jessica Yott and Diane Poulin-Dubois

Concordia University, Canada

The development of theory of mind (ToM) in infancy has been mainly documented through studies conducted on a single age group with a single task. Very few studies have examined ToM abilities other than false belief, and very few studies have used a within-subjects design. During 2 testing sessions, infants aged 14 and 18 months old were administered ToM tasks based on the violation-of-expectation paradigm which measured intention, true belief, desire, and false-belief understanding. Infants' looking times at the congruent and incongruent test trials of each task were compared, and results revealed that both groups of infants looked significantly longer at the incongruent trial on the intention and true-belief tasks. In contrast, only 18-month-olds looked significantly longer at the incongruent trial of the desire task and neither age group looked significantly longer at the incongruent trial on the false-belief task. Additionally, intertask comparisons revealed only a significant relation between performance on the false-belief and intention task. These findings suggest that implicit intention and true-belief understanding emerge earlier than desire and false-belief understanding and that ToM constructs do not appear to be integrated, as is the case for explicit ToM.

Understanding others' mental states and how they differ from one's own is an important milestone in children's cognitive development and has many implications for social development (Wellman, 2014). That is, understanding someone's intentions, desires, and beliefs is an essential tool for humans to interact, understand, and predict other people's behavior. This ability, termed theory of mind (ToM), has been heavily researched for more than 30 years. The most common way of examining ToM understanding has been through the use of false-belief tasks, as these tasks require an individual to predict another's behavior based on a belief that they themselves know is false (Premack & Woodruff, 1978; Wellman, Cross, & Watson, 2001; Wimmer & Perner, 1983). It was traditionally believed that children did not understand false beliefs until 4 years of age when they pass verbal standard false-belief tasks (e.g., location-change false-belief task; Wellman et al., 2001). The ability to attribute false beliefs requires being able to differentiate what one knows from what another person knows to predict his or her behavior. For example, when observing a toy being moved during a person's absence, one must then inhibit this knowledge to understand that this person has a false belief about the toy's location. Consequently, one can predict and understand this person's behavior according to this false belief. For example, this person may search for the toy where he or she falsely believes it to be hidden, instead of where it is actually

hidden. When asked, children younger than 4 years of age make this error and respond that the person will search for the toy in its actual location, as they are unable to inhibit and differentiate their own knowledge from another person's knowledge.

Understanding others' false beliefs is one ability in a complex network of mental-state reasoning abilities that children develop in the first 5 years of life. To examine how and when children develop critical aspects of mental-state reasoning, Wellman and Liu (2004) developed a ToM scale based on the assumption that children's insights about the mind develop in a predictable sequence. To do this, they administered a set of five methodologically comparable ToM tasks to 3-, 4-, and 5-year-old children. Results revealed that children first understand that they themselves can have a different desire than someone else for the same object before they understand that they themselves can have a different belief than someone else about the same object. Next, children understand people's knowledge access, followed by false belief. The last ability to develop in this ToM scale was the ability to understand that a person may display one emotion but feel a different emotion. Thus, results from this study revealed that children's ToM progresses in a fixed developmental sequence from 3 to 5 years of age. More recently, this invariant sequence of ToM development has been replicated and extended with longitudinal designs, in many different cultures, and in atypical populations (Peterson, Wellman, & Slaughter, 2012; Wellman, Fang, & Peterson, 2011).

During the last decade, researchers have begun to examine ToM abilities in infancy using implicit nonverbal tasks. These tasks have also been coined indirect or spontaneous response tasks, as opposed to direct, explicit, or elicited response tasks (Onishi & Baillargeon, 2005; Yott & Poulin-Dubois, 2012). These studies use infants' spontaneous responses (e.g., looking-time durations, anticipatory looking) to infer implicit ToM understanding. Using such tasks, a large number of studies have demonstrated a more precocious false-belief understanding (or other ToM understanding) than is indicated by standard tasks. For example, in 2005, Onishi and Baillargeon published an influential article in which they demonstrated that 15-month-olds have an understanding of both true and false beliefs using the violation-of-expectation (VOE) paradigm. More specifically, they demonstrated that infants look longer at a scene when the actor's behavior is incongruent with their true or false belief about a toy's location. According to the authors, these results suggest that infants expected the actor to behave according to where he or she believed the toy to be hidden and not where the toy was actually hidden, indicating the presence of belief understanding. Träuble, Marinović, and Pauen (2010) extended these findings by using a within-subjects design comparing looking times in a false- and true-belief task in 15-month-olds and provided evidence for flexible-belief understanding in infancy. In the last 20 years, numerous studies have shown belief understanding in children younger than 4 years of age, even in infants as young as 7 months of age, using a variety of looking-time measures and contexts (Buttelmann, Carpenter, & Tomasello, 2009; Clements & Perner, 1994; Kovács, Téglás, & Endress, 2010; Scott, Baillargeon, Song, & Leslie, 2010; Senju, Southgate, Snape, Leonard, & Csibra, 2011; Southgate, Senju, & Csibra, 2007; Surian, Caldi, & Sperber, 2007; Yott & Poulin-Dubois, 2012).

Similarly, research based on implicit response paradigms has been used to demonstrate other types of mental states in infancy such as intentions and goals (Olineck & Poulin-Dubois, 2007; Phillips & Wellman, 2005; Woodward, 1998). For example, Phillips and Wellman (2005) habituated 12-month-old infants to an actor reaching over a barrier to retrieve a toy. Next, the barrier was removed from the scene and the infants observed two types of test trials: a direct and an indirect reach for the toy. Results revealed that the 12-month-olds looked significantly longer

when the actor reached indirectly for the toy, as opposed to directly, once the barrier was removed, indicating an understanding of goal directedness when a successful intentional action is observed. Understanding desires or the preferences of others has been well researched in toddlers and preschoolers using interactive and explicit tasks (Repacholi & Gopnik, 1997; Wellman & Liu, 2004; Wright et al., 2005). However, studies often vary in terms of their task demands as well as in how they measure desire/preference understanding. In a landmark study of desire understanding in infancy, Repacholi and Gopnik (1997) found that infants begin to take into account the subjective desires of others at 14 to 18 months of age. Infants observed an actor express a preference for one type of food and distaste for another. Results revealed that infants reliably offered the actor their preferred food, even when the food desired was at odds with their own preference. Although not an explicit measure, this task requires the infant to give the food that he or she thinks the actor prefers. Tasks examining preference and/or desire understanding using looking-time measures have demonstrated that these abilities emerge in the 1st year of life (Henderson & Woodward, 2012). However, to our knowledge, there are no implicit versions of the Repacholi and Gopnik task, demonstrating implicit knowledge of desires. Therefore, one goal of the current study was to develop an implicit desire-understanding task, based on the task designed by Repacholi and Gopnik.

An important issue with regard to data obtained with nonverbal, implicit tasks is whether these tasks measure what they are assumed to measure—that is, the representation of others' mental states. For example, do the longer looking times associated with incongruent trials on the belief task reflect full-fledged belief understanding or a more primitive concept that is a precursor of belief understanding? Recent research has begun to address this question with longitudinal designs and has confirmed that performance on implicit ToM tasks is related to later, explicit ToM knowledge more generally (Low, 2010; Olineck & Poulin-Dubois, 2007; Thoermer, Sodian, Vuori, Perst, & Kristen, 2012). These findings suggest some continuity between early ToM abilities and the more explicit ToM understanding that develops later. Such continuity is compatible with the views that implicit tasks are measuring some kind of early ToM understanding. However, more research is needed, including research examining concurrent relations between ToM tasks to determine if a similar pattern of integration is present for implicit and explicit forms of these tasks.

Nevertheless, there is a hot debate in the literature regarding whether or not infants have a true ToM understanding (Ruffman, 2014; Sodian, 2011). Researchers have proposed a range of explanations for why infants demonstrate ToM understanding, and they have proposed developmental theories to support their claims. One possible explanation for why children younger than 4 years of age fail the standard false-belief task and yet infants pass implicit versions of such a task is that an understanding of false belief emerges earlier in development but methodological aspects of the traditional false-belief task mask competence. More specifically, it has been proposed that children fail this type of task because of task demands that require a range of cognitive skills, such as language competence and executive functioning, which are not yet fully developed (Baillargeon, Scott, & He, 2010; Carlson, Moses, & Claxton, 2004). As such, these accounts propose that by the end of the 1st year of life, infants can understand and attribute motivational states and reality-congruent informational states, like goals and true beliefs (Baillargeon et al., 2010; Leslie, Friedman, & German, 2004). In addition, they propose that it is only in the 2nd year of life that infants develop the understanding of reality-incongruent informational states like false beliefs. Moreover, they explain that the reason infants pass these

implicit tasks and not those with explicit demands in the 2nd year of life is that all language and executive-functioning requirements have been removed. However, recent research has demonstrated that some executive-functioning abilities are related to performance on implicit belief tasks, suggesting that implicit tasks may not completely eliminate executive-functioning requirements but rather decrease these demands (Yott & Poulin-Dubois, 2012).

Apperly and Butterfill (2009) have also proposed a two-system account of ToM development, although they posited that infants are able to solve some ToM tasks due to the development of an efficient yet limited and inflexible system. They argue that these limitations are only overcome when children acquire psychological concepts like belief and desire related to the development of language and executive functioning, which provides them with a new system for ToM reasoning that is flexible but inefficient. In contrast to Baillargeon and colleagues (2010), they do not believe that the reason that infants pass implicit ToM tasks is simply due to the removal of executive-functioning and language demands. Instead, they proposed that a system is in place to reason quickly about ToM but that this system is inflexible and fragile. Apperly and Butterfill's theory fits well with the adult, chimpanzee, and infant literatures on belief understanding and is consistent with results demonstrating fragile and inconsistent ToM reasoning in children (for a review, see Low, Drummond, Walmsley, & Wang, 2014; Low & Watts, 2013; Poulin-Dubois, Polonia, & Yott, 2013; Surtees, Butterfill, & Apperly, 2012). In contrast, others have also proposed that ToM understanding in infancy as demonstrated by implicit tasks may not reflect any mental-state attribution but could, in principle, reflect novelty effects or low-level processing (Heyes, 2014; Perner, 2014; Ruffinan, 2014; Ruffinan & Perner, 2005). That is, it has been proposed that infants might simply respond to something unusual in the display or react according to violation of behavioral rules.

In sum, our current understanding of ToM development in infancy is incomplete, and research has focused primarily on belief understanding. Keeping these limitations in mind, it is of foremost importance to understand what ToM abilities are present in infancy, how they relate to one another, and when they develop. One way to contribute to the debate on the nature of implicit ToM reasoning in infancy is to examine the development of a wide range of ToM abilities across more than one age group. Additionally, examining intertask correlations among ToM tasks in infancy and how they compare to those found in older children might help to clarify this debate. Research on intertask correlations among ToM tasks in children is limited, and much of the extant literature has focused on the preschool period. Results from these studies vary considerably as they have shown anywhere from weak to strong intertask correlations. For example, some studies have demonstrated weak to moderate intertask relations among ToM tasks in 3- to 4-year-olds (Carlson, Mandell, & Williams, 2004; Hughes & Ensor, 2007), whereas other studies including children in the same age range have reported strong correlations (Gopnik & Astington, 1988; Rakoczy, Bergfeld, Schwartz, & Fiske, 2015).

In infancy, even fewer studies have examined intertask correlations. Chiarella, Kristen, Poulin-Dubois, and Sodian (2013) reported no significant correlations among scores on ToM tasks in 30- to 38-month-old toddlers. More specifically, children completed two or three ToM tasks, including a visual perspective-taking task, a desire-understanding task, and an emotional perspective-taking task, and results revealed no significant correlations in either the Canadian or German sample. In contrast, some studies have revealed weak intertask relations in 24-month-old children (Carlson, Mandell, et al., 2004; Hughes & Ensor, 2005). Lastly, in 18-month-olds, Yott and Poulin-Dubois (2012) demonstrated a trend-level association between an interactive intention task and an implicit

false-belief task. Taken together, these few results suggest that ToM abilities in young children are weakly, if at all, related, and may only integrate later in development, but only a few tasks have been compared.

To date, there have been no studies examining the intertask correlations of multiple ToM abilities in infancy using a within-subjects design. Additionally, there is a paucity of research examining developmental trends of multiple ToM abilities in infancy. Based on these important gaps in the literature, the goals of the present study were threefold: a) to investigate the development of ToM abilities during the ages of 14 to 18 months, b) to examine how concurrent ToM abilities relate to one another as assessed with the VOE paradigm, and c) to document desire understanding using a novel implicit desire task.

METHOD

Participants

A group of 43 14-month-old infants (24 boys) and 53 18-month-old infants (32 boys) participated in the study. The mean age for the 18-month-old sample was 1;6 (range = 1;4–1;8) and the mean age for the 14-month-olds was 1;2 (range = 1;1–1;3). The sample was ethnically diverse and included families who identified themselves as being of Caucasian descent ($n = 57$), Asian descent ($n = 8$), African descent ($n = 6$), Arab descent ($n = 8$), South American descent ($n = 5$), or Caribbean descent ($n = 4$). Eight did not report their ethnicity. On the basis of parental reports, infants had no visual or auditory impairments and had a minimum 35-week gestational period. All infants were recruited from birth records provided by a governmental health services agency and were exposed to primarily English or French.

An additional 6 (12%) 14-month-old infants participated but were excluded from the sample due to fussiness ($n = 4$), parental interference ($n = 1$), or missing their second appointment ($n = 1$). Similarly, an additional 19 (26%) 18-month-olds infants participated but were excluded from the study due to fussiness ($n = 8$), parental interference ($n = 2$), technical difficulties ($n = 2$), a reported developmental delay ($n = 1$), or missing one of the two testing sessions ($n = 6$).

Materials

A stage-like apparatus (107 cm × 61 cm × 211 cm) was used to administer all four tasks. The experimenter was visible through a window (86 cm × 91 cm) located approximately 80 cm from the bottom of the front panel. Visible through the window was a flat surface, which was used as a stage where the toys were placed for any given task. Just below the surface top was a small circular opening where a digital camera was placed to record infants' looking behavior. Infants observed the experimenter from a high chair ($n = 61$) placed approximately 110 cm from the display or from their parent's lap ($n = 35$). All parents were asked to remain silent during the testing procedures. A white curtain that was operated by the experimenter covered the window of the display. Infants' looking patterns were coded live by a second experimenter using the Habit 2000 program (University of Texas) on a Mac G4 computer.

Receptive Vocabulary

To measure receptive vocabulary, parents completed the MacArthur-Bates Short Form Vocabulary checklist: Level 1, which includes 89 words. Parents were asked to indicate the words that their child understood.

Intention Task

For the intention task, a black barrier (30 cm × 25 cm) was used and placed on the right side of the stage. At the beginning of the task, a small yellow duck (12 cm × 12 cm) was placed on the far side of the barrier, facing the infant.

Belief Tasks

A red cup (7.5 cm in diameter, 10.5 cm high) or a yellow duck (11 cm × 11 cm) was placed on the surface top directly between a yellow box and a green box (14 cm × 14 cm × 14 cm). The distance between the boxes was 18 cm. Each box had an opening on the side facing the cup or duck (14 cm × 14 cm) that was covered with a fabric fringe. A rectangular opening underneath each box allowed for the attraction between a magnet located inside the cup and duck (2.5 cm × 5 cm length × 0.6 cm) and a magnet under the stage, operated by the experimenter (7.6 cm diameter). The magnet was used to unobtrusively transfer the cup or duck from one box to the other underneath the stage.

Desire Task

Two food pairings were used during the desire task. The first pairing was lettuce and Honey Nut Cheerios, and the second pairing was broccoli and Pepperidge Farm goldfish crackers. All food items were placed in clear plastic containers.

Design and Procedure

Infants and their parents were invited to the laboratory for two testing sessions that each lasted approximately 45 min and were scheduled 1 to 2 weeks apart. Upon arrival, they were brought to a reception room where infants were familiarized with the experimenters and the environment, and parents completed a consent form, a demographic questionnaire, and the Receptive Vocabulary Checklist. The Receptive Vocabulary Checklist was administered to control for infants' verbal ability. Tasks were recorded to code infants' responses off-line. All families were offered \$20 in financial compensation per session for their participation in this study.

All infants participated in one belief task, one desire task, and the intention task during their first visit and the second belief task and desire task during their second visit. The order of the tasks was counterbalanced so that each task was presented first, second, or third. It is important to note that if an infant completed the belief task first during the first visit, then he/she did the second belief task first at the second visit. Counterbalancing created 12 different orders. No order effects were observed.

Intention Task

The intention task was adapted from Phillips and Wellman (2005). This task consisted of seven trials. The first three trials were familiarization trials, during which a black barrier separated the experimenter from a yellow duck. Each familiarization trial began with an attention-attracting noise and the curtain rising. During the ensuing 2-s demonstration phase, the experimenter reached over the barrier, grabbed the duck, and held it in front of her while gazing at it. Trials were coded live and began once the experimenter paused while holding the duck. The trial ended if the infant looked away from the display for more than 2 consecutive seconds after looking at the display for a minimum of 2 cumulative seconds or if he/she looked away for 10 consecutive seconds. A trial lasted a maximum of 30 s. The test trials were identical except that the barrier was removed from the stage. Of the four test trials, two trials were congruent, where the experimenter reached directly for the duck and then held it in front of her. In contrast, during the incongruent trial, the experimenter reached for the duck indirectly by following the same path as though the barrier was present. This reach was considered incongruent because the experimenter no longer needed to follow this path. The trials alternated between congruent and incongruent.

Belief Tasks

Infants participated in two belief tasks, a full-box belief task and an empty-box belief task, each completed on a separate day. The belief tasks were adapted from Onishi and Baillargeon (2005) to examine infants' understanding of true and false beliefs. These tasks are nonverbal and are based on the VOE paradigm. During each of the belief tasks, all infants completed three familiarization trials, followed by a false-belief induction trial and a false-belief test trial and then a true-belief induction trial followed by a true-belief test trial. An attention-attracting sound played at the beginning and end of each trial when the curtain was raised and then lowered.

During the 8-s familiarization trial, the experimenter raised the curtain, picked up the cup, and placed it inside one of the two boxes. Once the cup was hidden, the experimenter paused with her hand inside the box. The trial ended if the infant looked away from the display for more than 2 consecutive seconds after looking at the display for a minimum of 2 cumulative seconds. In addition, if infants looked away for 10 consecutive seconds before having looked at the screen for the minimum 2 s, the trial ended. A trial lasted a maximum of 30 s. During the second and third familiarization trials, the experimenter reached into the box where the cup was hidden and then paused with her hand inside the box until the trial ended.

During the false-belief induction trial, the cup moved from one box to the other through a magnet operated by the experimenter. Next, the infants observed a false-belief test trial during which the experimenter reached into the full box (the box with the cup). This search behavior was considered incongruent, because the experimenter's behavior (searching in the full box) was incongruent with her belief (no knowledge of the cup's new location). Next, infants observed a true-belief induction trial, where the experimenter moved the cup back to its original location. In contrast to the false-belief induction trial, the experimenter remained in sight and followed the cup's movement from one box to the other. Lastly, infants observed the true-belief test trial during which the experimenter again reached into the full box (the box with the cup). This time, the trial was considered to be congruent, because the experimenter's action (searching in the full box) was consistent with her belief

(knowledge of the cup location). This belief task is called the full-box task, as infants observed the experimenter search in the full box during each test trial. However, in the case of the false-belief test trial, the trial was incongruent, whereas during the true-belief test trial, it was congruent.

During the empty-box belief task, infants observed the same experimenter and trials, except that the yellow duck replaced the toy cup and the experimenter searched in the empty box instead of the full box during the true- and false-belief test trials. Like the full-box belief task, infants observed both an incongruent trial and a congruent trial, but this time, the congruent trial occurred during the false-belief test trial, because the experimenter searched in congruence with her belief (the toy's original location). In contrast, the incongruent trial corresponded to the true-belief test trial because the experimenter searched in the empty box after having observed the cup change locations.

By the end of the second visit, all infants had observed an incongruent trial and a congruent trial following both the true- and false-belief induction trials. By using a within-subjects design, infants' individual looking times during the incongruent and congruent trials could be compared for both the true-belief and false-belief scenarios. The order in which the infants completed the belief tasks was counterbalanced. Additionally, the design allowed for a congruent trial to be presented first and an incongruent trial to be presented second in one belief task and vice versa in the second belief task.

Desire Task

This task was adapted from Repacholi and Gopnik's (1997) interactive desire task. Infants participated in the congruent desire task during one visit and the incongruent desire task during the second visit. The congruent task was composed of three familiarization trials followed by one test trial. The familiarization trials began with an attention-attracting sound and the curtain rising. Placed in front of Experimenter 2 (E2) were two plastic containers, one filled with crackers and the other with broccoli. Placed in front of the primary experimenter (Experimenter 1 [E1]) were three pieces of broccoli and three crackers. The trial began with E1 picking up a cracker, showing it to the infant, and then eating it. After she placed the cracker in her mouth, she said with a look of content, "Mmm, crackers, mmm." She then picked up a piece of broccoli, placed it in her mouth, and said, "Eww, broccoli, eww," with a look of disgust. She ate all food items in the same manner while E2 watched with a neutral facial expression. These familiarization trials lasted approximately 20 s.

The test trial began when E1 turned to E2, looked up at the full containers in front of her, and said, "Can I have some?" with her hands open in front of her, palms up. E2 then looked at both containers of food, reached for the crackers, and placed some in her hand. E1 turned toward the infant and then looked down at the food with a neutral facial expression and paused. Both experimenters remained still for 10 s. The incongruent desire task followed the same procedure, except that the crackers and broccoli were replaced with Cheerios and lettuce. In addition, during the familiarization trials, E1 demonstrated a preference for the Cheerios; however, during the test trial, E2 gave her lettuce. In this way, the incongruent desire task differed from the congruent desire task, because E1 received the food for which she did not demonstrate a preference.

Coding and Reliability

Infants' looking times at the scene during each task were coded offline using INTERACT 8.0 (Mangold, 2010). To be included in the analyses, infants were required to observe at least 70%

of the test demonstration phase(s) for all tasks. Based on this criterion, 7 infants were excluded from the intention task (five 14-month-olds and two 18-month-olds), 11 infants (eight 14-month-olds and three 18-month-olds) were excluded from the true-belief task, 4 infants were excluded from the false-belief task (four 14-month-olds and no 18-month-olds), and one 18-month-old infant was excluded from the desire task. Additionally, any infant who did not watch the still phase of any test trial (score of 0) was excluded from the final analyses. As such, 6 infants were excluded from the true-belief task, 1 from the intention task, and 1 from the desire task.

To establish interrater reliability, an independent observer coded a minimum of 25% of the data. Using Pearson product-moment correlations to compare overall looking time at the scene, the mean interobserver reliability was calculated. Scores greater than $r = .9$ were considered to reflect high agreement. For all tasks, the entire trial was coded for looking time on and off the scene; however, looking time on screen was the variable of interest. In all cases, reliability was greater than $r = .90$ ($ps < .001$).

RESULTS

Preliminary analyses revealed that the looking-time measures for the intention, desire, true-belief, and false-belief tasks were not normally distributed, and therefore, an additive (+1) log₁₀ transformation was applied. Following these adjustments, the data were normally distributed, thereby meeting the normality assumption for parametric statistical tests. As the results from analyses on both raw and transformed scores revealed the same findings, only those from the original raw scores are reported. To examine group performance on each task individually, only the scores from the infants who completed a given task were used, and therefore, the sample size varied depending on the task being examined ($n = 63$ to $n = 81$). Additionally, because infants were administered up to three VOE tasks during a testing session, infants' looking time during each demonstration phase was examined to control for potential fatigue effects. For all tasks, the average percentage of time infants watched the demonstration phase was greater than 97%, with a minimum percentage looking time greater than 70%. These results suggest high attention during the VOE tasks in the final sample. To ensure that infants were equally attentive during each of the true- and false-belief induction trials, infants' looking times were compared using a mixed-design analysis of variance (ANOVA). Results revealed no main effect of trial or interaction suggesting that infants looked equally long during both the true-belief induction trials (incongruent induction trial, $M = 8.67$ s, $SD = 6.40$ s; congruent induction trial, $M = 8.40$ s, $SD = 5.44$ s), as well as the false-belief induction trials (incongruent induction trial, $M = 6.44$ s, $SD = 4.61$ s; congruent induction trial, $M = 7.67$ s, $SD = 5.80$ s).

To examine infants' understanding of intentions, desires, and true and false beliefs, a mixed-design ANOVA with age group (14-month-olds, 18-month-olds) as the between-subjects factor and test trial (incongruent, congruent) as the within-subjects factor was used for each task. To examine infants' performance on the intention task, looking time at the first and second pair of congruent and incongruent test trials was averaged and then compared. Results revealed a main effect of trial, $F(1,66) = 5.92$, $p = .02$, $\eta_p^2 = .08$ (incongruent trial, $M = 6.30$ s, $SD = 4.08$ s; congruent trial, $M = 5.11$ s, $SD = 2.50$ s). No significant age effect or interaction effects were observed.

Results from the true-belief task revealed a main effect for trial, $F(1,61) = 3.88$, $p = .05$, $\eta_p^2 = .07$ (incongruent trial, $M = 7.45$, $SD = 5.94$; congruent trial, $M = 5.63$, $SD = 4.02$).

Additionally, there was a main effect of age, $F(1,61) = 5.67$, $p = .02$, $\eta_p^2 = .09$, such that 18-month-old infants looked longer at the scene compared to 14-month-olds. There was no significant interaction between age group and trial.

Results from the desire task revealed no main effect of trial; however, an interaction effect was observed, $F(1,79) = 6.57$, $p = .01$, $\eta_p^2 = .08$. As shown in Figure 1, planned follow-up pairwise comparisons revealed that 18-month-old infants looked significantly longer at the incongruent trial ($M = 8.99$ s, $SD = 1.36$ s) compared with the congruent trial ($M = 8.41$ s, $SD = 1.91$ s; $p = .04$, $\eta_p^2 = .05$). In contrast, younger infants looked equally long at the incongruent ($M = 8.57$, $SD = 1.37$) and congruent ($M = 9.07$, $SD = 1.90$; $p = .12$, $\eta_p^2 = .03$) test trials.

Results from the false-belief task revealed an effect of trial at the trend level, $F(1, 70) = 2.88$, $p = .09$, $\eta_p^2 = .04$ (incongruent trial, $M = 7.72$ s, $SD = 6.41$ s; congruent trial, $M = 6.19$ s, $SD = 5.83$ s). No age-group or interaction effects were observed.

Results from the belief tasks revealed that infants looked significantly longer at the incongruent trial compared with the congruent trial in the true-belief task but not in the false-belief task. Recall that during each testing session, infants completed an incongruent false-belief trial and a congruent true-belief trial or a congruent false-belief trial and an incongruent true-belief trial. Therefore, looking times at the incongruent trial and congruent trial within the same session can be compared as a measure of belief understanding. A mixed-design ANOVA was used to compare looking time, with day (first session, second session) and trial (incongruent, congruent) as within-subjects factors and age (14-month-olds, 18-month-olds) as the between-subjects factor. Results revealed only a main effect of trial, $F(1, 57) = 8.61$, $p = .01$, $\eta_p^2 = .13$. No main effect of age or testing day was found, nor were any significant interactions observed.

Lastly, to examine intertask relations, partial correlations between looking times on incongruent trials, while controlling for age and receptive vocabulary, revealed a statistically significant correlation only between the false-belief and intention scores (see Table 1). Additionally, as expected, language abilities were not correlated with performance on the VOE tasks, except for a marginally significant association with looking time at the incongruent scene on the true-belief task. No other significant correlations were observed when incongruent looking times or congruent looking times were compared across tasks.

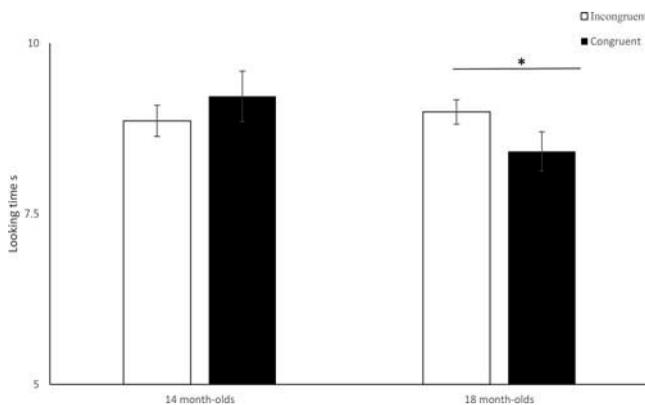


FIGURE 1. Mean looking time at the incongruent and congruent trials in the desire task by age.

TABLE 1

Partial correlations between looking time at the incongruent scenes in intention, true-belief, desire, and false-belief tasks, while controlling for age and receptive vocabulary.

<i>Task</i>	1	2	3	4
Intention	—			
True Belief	.06 <i>n</i> = 46	—		
Desire	.11 <i>n</i> = 55	-.07 <i>n</i> = 60	—	
False Belief	.41** <i>n</i> = 49	-.01 <i>n</i> = 59	-.01 <i>n</i> = 64	—

** $p < .01$.

DISCUSSION

The findings from the present study provide many contributions to the literature on ToM development in infancy. Firstly, the study addresses an important gap in the literature by examining the development of multiple ToM abilities at two time points during the 2nd year of life. Secondly, the inclusion of a wide range of ToM tasks combined with a within-subject design provides information about how ToM abilities are related to one another.

As expected, during the 2nd year, no developmental changes in intention understanding were observed as shown by longer looking times on average at the incongruent test trials compared with the congruent test trials in both age groups. These results are consistent with previous research showing that infants have developed an understanding of intentions by the end of the 1st year of life (Olineck & Poulin-Dubois, 2007; Phillips & Wellman, 2005; Woodward, 1998).

With regard to the true-belief task, our results revealed that both 14- and 18-month-olds looked longer at the incongruent test trial compared with the congruent test trial. Again, these results are consistent with previous findings indicating true-belief understanding in 15-month-old infants, with both within- and between-subject designs (Onishi & Baillargeon, 2005; Träuble et al., 2010).

Results from the desire task suggest that 18-month-old infants were sensitive to another person's desires and that they remembered the emotional expression associated with each type of food during the familiarization trials. However, 14-month-olds did not show this pattern. This is the first VOE adaptation of the Repacholi and Gopnik (1997) interactive task to measure desire understanding in infancy. It is an important addition to the literature because younger infants typically fail to respond to an object or food request. Thus, a VOE adaptation of the subjective desire task allows for testing desire understanding in younger samples with minimal task demands. These results suggest that infants' ability to reason about another person's desires develops during the 2nd year because even after removing the demands of the original desire task (food request, inhibition of own food preference), 14-month-olds do not appear to understand desires.

Finally, results from the false-belief task revealed that infants only tended to look longer at the incongruent scene compared with the congruent scene, which is inconsistent with studies conducted with both between- and within-subject designs demonstrating false-belief

understanding at 7 (Kovács et al., 2010), 13 (Surian et al., 2007), 15 (Onishi & Baillargeon, 2005; Träuble et al., 2010), 18 (Yott & Poulin-Dubois, 2012), and 24 months of age (Southgate et al., 2007). In contrast, these results do support the idea that false-belief understanding may be fragile and/or that it emerges later during the 2nd year of life, as reported in other recent studies (Poulin-Dubois et al., 2013; Sodian, 2011; Thoermer et al., 2012). For example, the present results replicate previous research by Thoermer and colleagues (2012), who demonstrated that only 55% of their 18-month-old sample looked significantly longer at the correct versus incorrect location with the anticipatory looking procedure. Taken together, these findings only partially support developmental theories that propose that true-belief understanding develops in the 1st year of life, followed by false-belief understanding by the middle of the 2nd year of life (Baillargeon et al., 2010). These results do, however, fully support the proposed developmental sequence of true belief developing before false belief—a hypothesis that has yet to be demonstrated empirically using a longitudinal design.

There are several possible explanations for lack of replication of previous false-belief findings based on the VOE or anticipatory looking paradigms. First, in the current design, true- and false-belief tasks were combined with the incongruent and congruent trials administered on separate days. Administering the tasks in this way provides a very conservative test of infants' belief understanding because infants' looking times are a sensitive measure that can be affected by extraneous variables. Additional analyses revealed that if looking time during the incongruent and congruent trials within the same day were compared, infants looked significantly longer at the incongruent trials on both days. These results suggest infants were more surprised by the incongruent test trials in both the false-belief and true-belief scenarios. However, these results should be interpreted with caution, as the looking times are being compared across two different types of belief tasks.

The current study provides unique information about intertask relations between ToM tasks based on the VOE paradigm. A significant correlation was observed between false-belief incongruent looking time and intention incongruent looking time. This finding is particularly interesting, as it replicates previous research demonstrating a relation between performance on the same false-belief task and intention measured with the behavioral re-enactment task (Yott & Poulin-Dubois, 2012). It seems possible that to interpret an actor's false belief, infants also need an understanding of her intention to find the object, regardless of the object's location. This requirement may be less important in true-belief understanding, where both the actor and the child observe the same events. It is important to highlight, however, that no other significant correlations were observed. The fact that infants' looking times at the scenes were largely unrelated across tasks is consistent with previous research on ToM development in toddlers or young children using interactive tasks (Carlson, Mandell, et al., 2004; Chiarella et al., 2013; Hughes & Ensor, 2005). Taken together, these results suggest that ToM concepts may develop independently in infancy and may only integrate during the preschool years. Moreover, it is possible that the lack of intertask correlations is due to the fact that looking-time measures are sensitive to extraneous variables and therefore may have low test-retest reliability. However, even in the preschool years, research on intertask associations reveals mixed results ranging from no association to strong associations depending on the abilities being measured (Carlson, Mandell, et al., 2004; Hughes & Ensor, 2005; Rakoczy et al., 2015).

We believe that the present findings have implications for recent proposals about the nature of the precocious ToM abilities reported in infancy during the last decade. Supporters of a rich interpretation of implicit ToM, understanding in infancy have proposed that infants possess a

representational ToM but that this understanding is masked by the demands of standard tasks, including advanced language and executive-functioning abilities (Baillargeon et al., 2010). Such competence is revealed when tasks based on spontaneous responses are used to test ToM abilities, such as those based on the VOE paradigm. At first glance, results from the current study fit with such a rich proposal in the sense that like the mature form of ToM observed years later, intention, true-belief, and desire concepts develop before false-belief understanding (Wellman & Liu, 2004). However, our results are inconsistent with previous findings demonstrating false-belief understanding using implicit designs at 18 months of age or earlier (Onishi & Baillargeon, 2005; Scott et al., 2010; Träuble et al., 2010).

In contrast, supporters of lean or minimalist interpretations of implicit ToM in infancy have argued that a conceptual shift occurs in development and that performance on implicit ToM tasks can simply be accounted for by low-level perceptual processing or detection of statistical regularities (Heyes, 2014; Ruffman, 2014). Although the current study was not designed to directly address such proposals, if infants' responses on the tasks were explained by these low-level accounts, specifically in the belief tasks, one would expect such variables to influence performance consistently across tasks. Given that we observed variability in performance across tasks and little or no coherence among task performance at either age, it seems unlikely that infants were simply reacting to low-level changes between the familiarization and test scenes or to violation of statistical irregularities. In addition, the observed developmental progression from intention and true-belief understanding to desire and false-belief understanding is difficult to explain using these minimalist interpretations. Furthermore, we believe that the pattern of results observed for the desire task alone, which shows developmental changes during the 2nd year, cannot be easily explained in terms of perceptual novelty or processing of statistical irregularities.

Results from the current study are consistent with the idea that infants' belief understanding is initially rigid and inflexible (Apperly & Butterfill, 2009). According to this view, infants can solve some ToM tasks but cannot solve more difficult ToM tasks. The current findings support this theory by demonstrating infants' understanding of intentions and true beliefs but only a nascent understanding of desires and false beliefs. Apperly and Butterfill (2009) proposed that children begin to reason about desires and beliefs gradually as they develop language, executive functioning, and an understanding of psychology concepts. The development of these related abilities allows for the emergence of a more flexible ToM system, which allows for the attribution of desires and false beliefs. Results from the current study do in part support this notion, as language abilities were not related to task performance.

The current study suggests that implicit ToM abilities develop in much the same way as later, more explicit abilities emerge, starting with intention understanding, followed by more complex ToM abilities such as desire and false-belief understanding. Moreover, while implicit ToM understanding may not reflect the same explicit understanding observed in preschool children, the current study combined with longitudinal studies suggests that these earlier ToM abilities are preconceptual abilities that provide the foundations for later explicit ToM understanding.

One limitation of the current study is that approximately half of the original sample of infants did not complete all four tasks. One reason for the high attrition rate was that infants had to complete tasks across two testing sessions. Measuring looking times across days opens up this measure to extraneous variables, such as mood, fatigue, or interest level. Therefore, if infants did not complete the belief task on 1 of the 2 days, they were excluded from the analyses. Future research should examine different ways of measuring multiple ToM abilities, using implicit designs in a single

session (e.g., anticipatory looking paradigm), to reduce attrition rates, but to also replicate and extend our findings. Another limitation to the current study is that individual patterns of ToM development could not be examined. That is, we were unable to examine if infants were likely to pass tasks in a particular sequence. This was due to several factors, including the fact that looking times were used to examine ToM ability, that only two age groups were examined, and that the number of infants who completed all four tasks was low. Nevertheless, in an effort to address this question, infants were given a pass or fail score on each task. On all tasks, the percentage of infants passing each task was not significantly above chance. In addition, we observed no significant difference in task performance or difficulty level. Taken together, these results suggest that changing the variables from continuous to dichotomous likely removed the individual variability observed in the looking times on incongruent and congruent trials. Moreover, the lenient pass/fail score was likely not sensitive enough to observe individual or group differences, if any. Future studies should examine individual differences in ToM development using interactive implicit tasks that are scored as success or failure. Additionally, it would be very interesting to examine how ToM abilities assessed with looking-time measures relate concurrently to ToM abilities assessed with interactive measures. We are currently investigating this issue with a range of implicit measures and hope to contribute to this literature on ToM development in infancy (Yott & Poulin-Dubois, 2015).

In conclusion, the present study confirms and extends previous research on infants' ToM abilities, expands our knowledge of intertask relations among ToM tasks, and makes an important contribution to the literature with regard to ToM development in infancy. Moreover, this study is the first to use a within-subjects design to examine ToM development in infancy with a wide range of tasks, including for the first time, a VOE adaptation of the desire task based on Repacholi and Gopnik (1997). Taken together, the results from the present study add to the mounting evidence supporting the notion that there is limited coherence with respect to infants' ToM understanding. These findings, in turn, suggest that mental-state reasoning involves significant developmental changes and that the data on infants' ToM "may not reflect the working of an innate, well-formed theory of mind" (Perner, 2014, p. 295). However, how these skills develop from infancy to childhood and what important cognitive changes take place during this period remain open to debate.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the contributions of Josée-Anne Bécotte, Cristina Crivello, Janice La Giorgia, Katherine Gittins, Sara Phillips, Monyka Rodrigues, and Ioanna Solomatina in data collection and coding. Finally, the authors would like to express their gratitude to the research participants whose contribution made this project possible.

FUNDING

This research was supported by grants from the Social Sciences and Humanities Research Council of Canada and the National Institute of Child Health and Human Development. In addition, this research was indirectly supported by the Centre for Research in Human Development.

REFERENCES

- Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review*, *116*, 953–970. doi:10.1037/a0016923
- Baillargeon, R., Scott, R. M., & He, Z. (2010). False-belief understanding in infants. *Trends in Cognitive Sciences*, *14*, 110–118. doi:10.1016/j.tics.2009.12.006
- Buttelmann, D., Carpenter, M., & Tomasello, M. (2009). Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition*, *112*, 337–342. doi:10.1016/j.cognition.2009.05.006
- Carlson, S. M., Mandell, D. J., & Williams, L. (2004). Executive function and theory of mind: Stability and prediction from ages 2 to 3. *Developmental Psychology*, *40*, 1105–1122. doi:10.1037/0012-1649.40.6.1105
- Carlson, S. M., Moses, L. J., & Claxton, L. J. (2004). Individual differences in executive functioning and theory of mind: An investigation of inhibitory control and planning ability. *Journal of Experimental Child Psychology*, *87*, 299–319. doi:10.1016/j.jecp.2004.01.002
- Chiarella, S. S., Kristen, S., Poulin-Dubois, D., & Sodian, B. (2013). Concurrent relations between perspective-taking skills, desire understanding, and internal-state vocabulary. *Journal of Cognition and Development*, *14*, 480–498. doi:10.1080/15248372.2012.689390
- Clements, W. A., & Perner, J. (1994). Implicit understanding of belief. *Cognitive Development*, *9*, 377–395. doi:10.1016/0885-2014(94)90012-4
- Gopnik, A., & Astington, J. W. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child Development*, *59*, 26–37.
- Henderson, A. M. E., & Woodward, A. L. (2012). Nine-month-old infants generalize object labels, but not object preferences across individuals. *Developmental Science*, *15*, 641–652. doi:10.1111/j.1467-7687.2012.01157.x
- Heyes, C. (2014). False belief in infancy: A fresh look. *Developmental Science*, *17*, 647–659. doi:10.1111/desc.12148
- Hughes, C., & Ensor, R. (2005). Executive function and theory of mind in 2 year olds: A family affair? *Developmental Neuropsychology*, *28*, 645–668. doi:10.1207/s15326942dn2802_5
- Hughes, C., & Ensor, R. (2007). Executive function and theory of mind: Predictive relations from ages 2 to 4. *Developmental Psychology*, *43*, 1447–1459. doi:10.1037/0012-1649.43.6.1447
- 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. doi:10.1126/science.1190792
- Leslie, A. M., Friedman, O., & German, T. P. (2004). Core mechanisms in 'theory of mind.' *TRENDS in Cognitive Sciences*, *8*, 528–533. doi:10.1016/j.tics.2004.10.001
- Low, J. (2010). Preschoolers' implicit and explicit false-belief understanding: Relations with complex syntactical mastery. *Child Development*, *81*, 597–615. doi:10.1111/j.1467-8624.2009.01418.x
- Low, J., Drummond, W., Walmsley, A., & Wang, B. (2014). Representing how rabbits quack and competitors act: Limits on preschoolers' efficient ability to track perspectives. *Child Development*, *85*, 1519–1534. doi:10.1111/cdev.12224
- Low, J., & Watts, J. (2013). Attributing false beliefs about object identity reveals a signature blind spot in humans' efficient mind-reading system. *Psychological Science*, *24*, 305–311. doi:10.1177/0956797612451469
- Mangold, P. (2010). *Interact user guide, V. 8.0. ff.* Arnstorf, Germany: Mangold International.
- Olineck, K. M., & Poulin-Dubois, D. (2007). Imitation of intentional actions and internal state language in infancy predict preschool theory of mind skills. *European Journal of Developmental Psychology*, *4*, 14–30. doi:10.1080/17405620601046931
- Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, *308*, 255–258. doi:10.1126/science.1107621
- Perner, J. (2014). Commentary on Ted Ruffman's 'Belief or not belief ...'. *Developmental Review*, *34*, 294–299. doi:10.1016/j.dr.2014.05.002
- Peterson, C. C., Wellman, H. M., & Slaughter, V. (2012). The mind behind the message: Advancing theory-of-mind scales for typically developing children, and those with deafness, autism, or Asperger syndrome. *Child Development*, *83*, 469–485. doi:10.1111/j.1467-8624.2011.01728.x
- Phillips, A. T., & Wellman, H. M. (2005). Infants' understanding of object-directed action. *Cognition*, *98*, 137–155. doi:10.1016/j.cognition.2004.11.005
- Poulin-Dubois, D., Polonia, A., & Yott, J. (2013). Is false belief skin-deep? The agent's eye status influences infants' reasoning in belief-inducing situations. *Journal of Cognition and Development*, *14*, 87–99. doi:10.1080/15248372.2011.608198

- Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, *1*, 515–526. doi:10.1017/S0140525X00076512
- Rakoczy, H., Bergfeld, D., Schwarz, I., & Fizke, E. (2015). Explicit theory of mind is even more unified than previously assumed: Belief ascription and understanding aspectuality emerge together in development. *Child Development*, *86*, 486–502. doi:10.1111/cdev.12311
- Repacholi, B. M., & Gopnik, A. (1997). Early reasoning about desires: Evidence from 14- and 18-month-olds. *Developmental Psychology*, *33*, 12–21. doi:10.1037/0012-1649.33.1.12
- Ruffman, T. (2014). To belief or not belief: Children's theory of mind. *Developmental Review*, *34*, 265–293. doi:10.1016/j.dr.2014.04.001
- Ruffman, T., & Perner, J. (2005). Do infants really understand false belief? Response to Leslie. *Trends in Cognitive Sciences*, *9*, 462–463. doi:10.1016/j.tics.2005.08.001
- Scott, R. M., Baillargeon, R., Song, H., & Leslie, A. M. (2010). Attributing false beliefs about non-obvious properties at 18 months. *Cognitive Psychology*, *61*, 366–395. doi:10.1016/j.cogpsych.2010.09.001
- 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. doi:10.1177/0956797611411584
- Sodian, B. (2011). Theory of mind in infancy. *Child Development Perspectives*, *5*, 39–43. doi:10.1111/j.1750-8606.2010.00152.x
- Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science*, *18*, 587–592. doi:10.1111/j.1467-9280.2007.01944.x
- Surian, L., Caldi, S., & Sperber, D. (2007). Attribution of beliefs by 13-month-old infants. *Psychological Science*, *18*, 580–586. doi:10.1111/j.1467-9280.2007.01943.x
- Surtees, A. D., Butterfill, S. A., & Apperly, I. A. (2012). Direct and indirect measures of level-2 perspective-taking in children and adults. *British Journal of Developmental Psychology*, *30*, 75–86. doi:10.1111/j.2044-835X.2011.02063.x
- Thoermer, C., Sodian, B., Vuori, M., Perst, H., & Kristen, S. (2012). Continuity from an implicit to an explicit understanding of false belief from infancy to preschool age. *British Journal of Developmental Psychology*, *30*, 172–187. doi:10.1111/j.2044-835X.2011.02067.x
- Träuble, B., Marinović, V., & Pauen, S. (2010). Early theory of mind competencies: Do infants understand others' beliefs? *Infancy*, *15*, 434–444. doi:10.1111/j.1532-7078.2009.00025.x
- Wellman, H. M. (2014). *Making minds: How theory of mind develops*. Oxford, UK: Oxford University Press.
- 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. doi:10.1111/cdev.2001.72.issue-3
- Wellman, H. M., Fang, F., & Peterson, C. C. (2011). Sequential progressions in a theory-of-mind scale: Longitudinal perspectives. *Child Development*, *82*, 780–792. doi:10.1111/j.1467-8624.2011.01583.x
- Wellman, H. M., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development*, *75*, 523–541. doi:10.1111/j.1467-8624.2004.00691.x
- 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. doi:10.1016/0010-0277(83)90004-5
- Woodward, A. (1998). Infants selectively encode the goal object of an actor's reach. *Cognition*, *69*, 1–34. doi:10.1016/S0010-0277(98)00058-4
- Wright, C. K., Cosetti, M., Jones, R., Kelton, E., Meier Rafal, V., Richman, L., & Stanhaus, H. (2005). Preschool children's understanding of conflicting desires. *Journal of Cognition and Development*, *6*, 427–454. doi:10.1207/s15327647jcd0603_6
- Yott, J., & Poulin-Dubois, D. (2012). Breaking the rules: Do infants have a true understanding of false belief? *British Journal of Developmental Psychology*, *30*, 156–171. doi:10.1111/j.2044-835X.2011.02060.x
- Yott, J., & Poulin-Dubois, D. (2015). *Scaling of theory of mind tasks in infancy*. Manuscript submitted for publication.