frontiers in Developmental Psychology

TYPE Original Research
PUBLISHED 03 March 2026
DOI 10.3389/fdpys.2026.1727052

OPEN ACCESS

EDITED BY
Marco Lunghi,
University of Padova, Italy

REVIEWED BY
Elisa Di Giorgio,
University of Padua, Italy
Qiyu Huang,
Northeast Normal University, China

*CORRESPONDENCE
Ilaria Grazzani
ilaria.grazzani@unimib.it

RECEIVED 03 November 2025
REVISED 11 February 2026
ACCEPTED 11 February 2026
PUBLISHED 03 March 2026

CITATION
Grazzani I, Conte E and Datteri E (2026) False belief attribution in toddlers: an exploratory study with a novel unexpected-identity task.
Front. Dev. Psychol. 4:1727052.
doi: 10.3389/fdpys.2026.1727052

False belief attribution in toddlers: an exploratory study with a novel unexpected-identity task

Ilaria Grazzani[1]*, Elisabetta Conte[2] and Edoardo Datteri[3]

[1]Lab-PSE, Department of Human Sciences for Education 'R. Massa', University of Milano-Bicocca, Milan, Italy, [2]Department of Human and Social Sciences, University of Bergamo, Bergamo, Italy, [3]RobotiCSS Lab, Department of Human Sciences for Education 'R. Massa', University of Milano-Bicocca, Milan, Italy

Introduction: Several studies, in the context of the debate on early implicit theory of mind, have investigated whether infants and toddlers are able to attribute false beliefs concerning the identity of an object. As a result, there is a lack of consensus regarding whether young children are able to understand others' belief about an object's identity when it can be represented in different ways. In the present study, we address this issue by using for the first time a close adaptation of a test originally devised by Butterfill and Apperly to advance the theoretical debate on early theory of mind. Given that this novel identity task could not be completed based on a minimal theory of mind, its use can play a significant role in such a debate.

Method: Employing an eye tracking system and based on the violation-of-expectation-paradigm, we explored how the participants (50 toddlers aged 20–24 months) performed in the true belief condition and in the false belief condition on a new identity task with a dual-identity object.

Results: Statistical analyses showed that the looking times and number of visits were not significantly higher in the TB condition than in the FB condition, supporting the claim that toddlers of this age don't demonstrate an implicit understanding of false belief.

Discussion: We discuss these outcomes in relation to the need for new studies operationalizing Butterfill and Apperly's test in order to advance the theoretical debate on one-system vs. two-system accounts of early theory of mind.

KEYWORDS

eye tracking, identity task, implicit theory of mind, minimal theory of mind, one-system theory, toddlers, two-system theory

1 Introduction

A vast body of research, based on elicited-response or "explicit" tasks, has suggested that children only develop an explicit understanding of an agent's false belief between 3 and 5 years of age (Wellman et al., 2001). In typical theory-of-mind tasks designed for preschoolers, children are required to interpret situations where an agent's representation of a setting is not merely incomplete but false. Thus, in the famous change-of-location test (e.g., the "Max and the Chocolate Test"), an object is moved while a story character is absent (Wimmer and Perner, 1983), whereas in the unexpected-contents test (e.g., "The Smarties Test"), a box does not contain what it is supposed to contain (Hogrefe et al., 1986). To

pass the task, children must attribute a false belief to a character and predict his or her consequent behavior (e.g., looking for the object where it was put before the change of location; affirming that the box contains the Smarties as expected).

However, recent investigations using spontaneous-response tasks or implicit measures, based on inferences from children's behavior, suggest that ToM ability is already present in the second year of life. This suggestion has been made mainly in studies that analyzed looking times within the violation-of-expectation (VoE), anticipatory looking (AL) and helping paradigms respectively. According to VoE paradigm, when individuals' expectations are violated, they experience surprise, leading them to look longer at the location where the unexpected event is taking place as well as at the agent's face. Tests that exploit the VoE paradigm generally comprise two steps: initially, there is a habituation or familiarization phase when young children are systematically led to form context-dependent expectations via repeated exposure to a given event; subsequently, across a series of trials, they encounter events that are either anticipated or unforeseen with respect to the expectations induced during the previous phase. In a pioneering study that drew on this paradigm, Onishi and Baillargeon (2005) found that 15-month-olds are more surprised and display longer looking times when other actors' behavior is inconsistent with their beliefs than when it is belief-congruent. Similarly, using an anticipatory-looking (AL) paradigm and an eye tracking to record 2-year-olds' looking behaviors, Southgate et al. (2007) showed that infants, who watched actions on a computer monitor, made anticipatory saccades on the basis of an attribution of false belief to an actor (e.g., children looked toward the location where the agent believe an object to be, and not where the toy actually was). In addition, relying on an active helping paradigm, Buttelmann et al. (2009) designed a task requiring infants to help an adult achieve a goal. To determine what their adult interlocutor's goal was, infants had to consider what that person believed, and specifically whether he or she falsely believed that there was an object inside a box. The authors showed that by 18 months of age (and in the same direction but less clearly for 16-month-olds), infants successfully took the adult's belief into account and acted appropriately.

Among the scholars interested in infants' theory of mind, many propose a "one-system" (or early mindreading) account whereby infants are born with an innate psychological-reasoning system that provides them with framework for interpreting the actions of other agents (e.g., Baillargeon et al., 2016). This system undergoes gradual conceptual enrichment over time (see Carruthers, 2016; Sodian et al., 2020). In this vein, it has been proposed that the inability of children younger than 4 years to pass standard verbal false-belief tests could be due to the fact that such tests overwhelm their limited resources (Baillargeon et al., 2010; Carruthers, 2013). On this interpretation, children under 4 years can understand other people's beliefs, and this ability can be experimentally confirmed by reducing the processing demands associated with the test.

An alternative family of hypotheses suggests the existence of two systems underlying people's theory of mind abilities: one system that is cognitively efficient but inflexible and limited, plus another system that is flexible but slower and cognitively more demanding (Apperly and Butterfill, 2009). This view aims to explain how theory of mind abilities can be fast and efficient in rapidly changing circumstances, while remaining flexible enough to support general explanatory and predictive needs: it is not immediately clear how a single ToM system could achieve both cognitive efficiency and flexibility simultaneously, as implied by the one-system view (Apperly and Butterfill, 2009). Proponents of the two-system account argue that very young children cannot yet ascribe false beliefs to others, generally proposing that even success in implicit false-belief tasks does not reflect genuine attribution of mental states on the part of infants. Rather, according to the two-system account, younger children may pass implicit false-belief tasks merely by relying on behavioral rules or other lower-level processes (e.g., Low et al., 2016; Perner and Ruffman, 2005; Rakoczy, 2012; Ruffman, 2023); only a late-developing, more flexible system allows children to overcome the limitations posed by the early system, passing first-order explicit false-belief tasks thanks to the genuine attribution of salient mental states.

Butterfill and Apperly (2013) proposed that children may succeed in change-of-location tasks by relying on a "minimal theory of mind", whereby they do not fully represent others' false beliefs. Furthermore, they proposed a novel identity task that cannot be completed based on a minimal theory of mind, and which can therefore play a significant role in resolving the debate between the one-system and two-system views as clarified in the following section. To the best of our knowledge, no operationalisation of Butterfill and Apperly's identity task has yet been published in the scientific literature. This paper presents an implementation of the task and discusses the results of an exploratory analysis conducted with children aged 20–24 months. Before offering a brief review of other identity tasks published in the literature (section 3) and describing the details of our study (section 4), it is helpful to review Butterfill and Apperly's minimal theory of mind, and the reasons they gave for claiming that the identity task could not be passed by relying on it.

2 The minimal theory of mind

Butterfill and Apperly (2013) have distinguished between a system with the capability to represent others' beliefs as such, and a "minimal" mentalization system with the capacity to track others' beliefs without representing them as such. What does it mean to represent a belief *as such*? Beliefs are traditionally construed as propositional attitudes that link an agent (who holds the belief) and a mental representation (for a general defense of this classical theory, see Fodor, 1978). For example, if Anna believes that the granny puppet is behind a panel, a particular relationship—belief—obtains between Anna and a mental representation whose content is "the granny puppet is behind the panel". One of the most striking characteristics of beliefs is *referential opacity*: the truth of the belief-ascribing statement ("Anna believes that the granny puppet is behind the panel") is not necessarily preserved when co-referential terms are substituted. For example, suppose that a granny puppet is identical to a wolf puppet, as they are actually the same object which can appear to be either a granny or a wolf, depending on one's point of view (in this case, "granny puppet" and "wolf puppet" are the co-referential terms). If the granny puppet is behind the panel, then the wolf puppet is behind the panel (and vice versa). But if Anna *believes* that the granny puppet is behind the panel, this does not

imply that Anna *believes* that the wolf puppet is behind the panel (and vice versa).

If an agent A can understand that Anna believes that the granny puppet is behind the panel—thus, if they can understand that a certain relationship (belief) obtains between Anna and a mental representation—then they will not infer from that belief attribution that Anna also believes that the wolf puppet is behind the panel. Thus, experimental observation that the agent *does not* make that inference can be regarded as bearing out the theory that the agent understands beliefs as such. The referential opacity of beliefs is closely related to the *aspectual* side of beliefs. If A can represent beliefs as such, they can understand that Anna can have a belief about how the puppet appears to her—for example, as a granny puppet. Anna's belief is about the *granny* puppet, and the fact that she holds this belief will not imply—in A's mind—that Anna will have a similar belief about the wolf puppet, if she does not know that they are the same object.

How do these considerations relate to the work of Butterfill and Apperly (2013)? These authors hypothesized that children under 4 years old possess a minimal mentalization system that does not enable them to understand beliefs as such. Indeed, this minimal system enables an agent to represent spatial-temporal relations among other people and *objects* (but not between people and their mental representations, as in the belief-understanding case). For instance, such a system might enable Agent A (e.g., a child) to represent the fact that Agent B has most recently encountered Object O in a particular location (in the authors' terminology, that B has registered O at that location). Consider, for example, a standard change-of-location task.

- In the true-belief condition, Child A watches while Actor B sees that Object O is moved from Box 1 to Box 2. Under this condition, the minimal system allows A to represent the fact that the actor most recently encountered the object in Box 2.
- In the false-belief condition, B is not there when the object is moved from 1 to 2. The minimal system, therefore, enables Child A to represent the fact that B last encountered O in Box 1 (even though it had been moved to Box 2 in the absence of B).

Now, suppose that the child expects B to act as if the object were at the location where B last encountered it. Accordingly, in the true belief condition, the child should expect the actor—for example—to search for the object in Box 2 (because they last encountered the object there). In the false-belief condition, conversely, the child would expect B to look for the object in Box 1, for the same reason. Nothing in this account implies that A has tracked B's *beliefs* about the location of Object O: A has only represented where B last encountered the object. Yet, in Butterfill and Apperly's view, such representational machinery might enable the child to pass at least some non-verbal false-belief tests. On the other hand, the minimal system displays some signature limits. Let us recall that beliefs are inherently aspectual. Yet, by definition, the minimal system only tracks spatio-temporal relations between B, O, and a spatial location, and is totally insensitive to how object O *appears* to B. Thus, suppose that O is the granny-wolf puppet mentioned earlier. The minimal system would only enable Child A to track where B last encountered O, regardless of whether O appeared to B as a granny or as a wolf. It cannot enable A to track how the puppet appears to B. As such, it involves no mechanism whereby A could track B's inherently aspectual beliefs.

These considerations underscore the theoretical interest of identity tasks. Identity tasks involve an object that appears differently from different perspectives, such as the granny-wolf puppet example outlined above. Thus, Butterfill and Apperly (2013) came up with the following task. A is a child; B is an adult actor; and O is a double-faced puppet. A and B sit opposite one another at a table with a panel serving as an occluder placed between them, such that neither can see the other's side of the table. The task begins when O enters the scene from one side, so that A sees, say, the granny and B the wolf. O moves behind the occluder at A's side (implying that the adult actor cannot see it anymore).

In the true belief (TB) condition, the object then re-emerges at the other side of the occluder, allowing both adult and child to see it again. Then, both A and B see the object rotating, so that A ends up seeing the wolf, and B the granny. O then is moved away from the scene. B reaches around the occluder in an attempt to grasp something—although he or she has no reason to believe that there is something there. The false-belief (FB) condition differs from the TB condition in that the object rotates when it is behind the occluder so that only the child sees it rotating. As in the TB condition, the object will eventually emerge from the other side of the occluder before departing the scene totally. At this point, A can see the wolf and B the granny. B then reaches around the occluder in an effort to grab something.

If the child can represent B's beliefs about the appearance of the object—more specifically, if the child can represent B's belief that O is an object with two faces—then *she will be surprised* when, in the TB condition, B reaches around the occluder. Correspondingly, if the child can represent B's beliefs about the appearance of the object she *will not be surprised* when B reaches around the occluder in the FB condition: the child will have enough representational abilities to infer *that B believes* that there is a wolf-like object behind the occluder (while the granny has gone away). On the contrary, if the child cannot represent B's belief about the appearance of the object, but—as in the minimal theory of mind—can only track O's position relative to both herself and B and remember when B last encountered O, then she will not be more surprised in the TB condition, relative to the FB condition, when B reaches around the occluder.

In summary, the relevance of this kind of identity task to the debate between the "one-system" and "two-system" accounts of Theory of Mind (ToM) abilities is as follows. If children under 4 years old show signs of surprise when the actor searches for the object in the TB condition but not in the FB condition, it is reasonable to hypothesize that they understand how the object appeared to the actor. Such an outcome would therefore support the hypothesis that children's ToM system is not minimal, as described above, and that they are able to understand beliefs. More specifically, in line with the argument put forward by Butterfill and Apperly (2013), this result would suggest that the participants could represent the actor's belief as such, i.e., its aspectual nature. A different result, showing that children are not more surprised in the TB than in the FB conditions, would not enable one to draw such a conclusion. Although other tasks could be devised

to contribute to the debate between the "one-system" and "two-system" accounts of ToM abilities, the identity task proposed by the authors—and developed in the study reported here—can be used to determine whether participants have a non-minimal understanding of other people's beliefs. While this would not conclusively resolve the debate, it would strongly support the view that, like adults, children under 4 years of age are able to represent the aspectual nature of beliefs, consistently with the "one-system" view.

3 Identity tasks in the literature

Flavell et al. (1983) were pioneers in using an explicit-response task to investigate preschoolers' ability to distinguish between the appearance and true identity of objects. In this task, children are presented with a deceptive object such as a fake rock made of sponge. The experimenter then reveals the object's true nature, for example by squeezing the imitation rock and allowing the child to examine it. The authors found that even at 4 and 5 years of age, children still struggle to consistently distinguish between an object's appearance, its reality, and their own mental representations of it. Gopnik and Astington (1988) conducted two studies with preschoolers to investigate the age at which children develop an understanding of representational change and how this relates to their false belief understanding. Their findings aligned with those of Flavell et al., indicating that children only begin to understand others' beliefs about the identity of objects and begin to grasp the distinction between appearance and reality around the age of 4 years.

Drawing on spontaneous response tasks, Scott and Baillargeon (2009) were the first to investigate infants' attributions to others of false beliefs about an object's identity, with participants observing events involving an agent and two toy penguins: one penguin could be disassembled (2-piece penguin) and one could not (1-piece penguin). These authors suggested that toddlers can attribute an agent with false beliefs concerning the identity of an object, providing new evidence of an early psychological-reasoning system. Since then, other studies based on identity tasks have confirmed these outcomes with infants and toddlers (Buttelmann and Kovács, 2019; Buttelmann et al., 2015; Kampis and Kovács, 2021; Scott et al., 2015).

Nevertheless, some authors, including Scott et al. (2015), have suggested that these findings could reflect infants' ability to recognize *types of objects* (categories of objects that possess the same properties, for example balls roll) rather than *object identities* (the fact that a certain object remains the same even if, for example, it changes position). It is one thing to expect an object of a certain *type*—for example, a divisible penguin—to be present at a certain location or to have specific functions; it is another thing to expect *that particular* divisible penguin to be there. Reasoning about object identity differs from reasoning about object types. While Butterfill and Apperly (2013) acknowledge the interest of studies involving reasoning about object types, they argue that "their implementation does not provide evidence that infants ascribe beliefs about identity". In other words, knowing something about certain types or categories of objects is different from reasoning about the identity of a specific object, that is the purpose of the present study.

Indeed, in keeping with the two-system account, other studies based on identity tasks did not find that infants and toddlers differentiate between true beliefs and false beliefs about the identity of an object (Fizke et al., 2017; Low and Watts, 2013). However, Fizke et al. (2017), for example, using the helping paradigm and comparing toddlers' performance on tasks with opaque objects, that require understanding of aspectuality, with toddlers' performance on non-aspectual tasks that require only change of location, recognized that their aspectual task could have placed high demands on participants' working memory, attentional load, difficulty in understanding the scene, thus precluding successful performance.

Overall, the outcomes of explicit-response tasks suggest that children only begin to distinguish between the appearance and the true identity of objects (e.g., a fake rock made of sponge) from around 4 years. Toddlers' performance on non-traditional tasks, on the other hand, suggest that they could already attribute false beliefs about the identity of an object. Indeed, the identity tasks adopted in the studies carried out so far have not provided unambiguous support for either the one-system or the two-system view of false belief understanding in young children. It is worth noting that none of these studies accurately reflected the structure of the task described by Butterfill and Apperly (2013) and discussed in the previous section. The next sections present an operationalisation of this task and the results of an exploratory study conducted with children aged 20–24 months.

4 The current study: an operationalization of Butterfill and Apperly's identity task

To the best of our knowledge, this is the first study to operationalize the experimental paradigm proposed by Butterfill and Apperly (2013). Following the authors' argumentation, the minimal system enables a subject to represent spatial-temporal relations among other people and objects but not between people and their mental representations, as in the case of understanding false beliefs. Given that the novel identity task applied in the current study cannot be completed based on a minimal theory of mind, we believe that its use can play a significant role in the debate on early theory of mind. We utilized two videos as stimuli, one for the TB condition and the other for the FB condition, and eye tracking technology to monitor the behavior of the toddlers involved. Our aim was to investigate whether toddlers would exhibit greater surprise in the TB condition than in the FB condition, as indicated by their looking-time patterns (with surprise signaled by the number of fixations and visits to the regions of interest). In line with Butterfill and Apperly's reasoning and the VoE paradigm, such a result would support the hypothesis that toddlers can represent the actor's beliefs about the object's appearance. This could not easily be explained by attributing only a minimal theory of mind to them. Different results would not enable one to draw such a conclusion, leaving open the possibility that they possess a minimal theory of

mind. The next subsection describes the methods and rationale of the study.

5 Method

5.1 Participants

To determine an appropriate sample size for the study, we performed a power analysis with G*Power (Faul et al., 2007). Drawing on prior research in the field, we anticipated a medium effect size (Cohen's $d = 0.50$) and set the desired power level *a priori* at 0.95. Our analysis showed that we need at least 45 participants to achieve 95% statistical power for the subsequent analyses.

The initial sample comprised 65 toddlers aged between 20 and 24 months. Provision of parental informed consent was obtained for each participant. We adopted 2 inclusion criteria for the study: (1) participants in the target age range of 20–24 months; (2) communication-language abilities within the thresholds for typical development. Regarding the last-mentioned criterion, the participants' mothers were asked to complete the PVB (Caselli et al., 2018), which is a short, standardized Italian-language version of the *McArthur-Bates-CDI* (Fenson et al., 2000). This test allowed us to exclude participants with significant delays in communication and language development given that the experimental task required them to comprehend the verbal utterances of the adult in the video (Conte et al., 2029). Toddlers were excluded if they were not collaborative during the testing (e.g., they cried or were inattentive and did not engage with the tasks;), or if their language abilities were well below the normative level suggested by Caselli et al. (2018).

Regarding the eye tracking data acquisition, we used a conservative criterion for trial inclusion, namely there had to be recorded gaze data in more than 50% of the critical time period. This resulted in excluding 3 toddlers (2 in FB and 1 in TB condition) who had no trials left to analyze. One more child was excluded because she looked away during the critical time. In total, the eye tracking data of 50 participants reached the criteria to be included in data analysis. The final sample comprised participants between 20 and 24 months (23 females; $M_{age} = 22.10$ months, SD = 1.44). In the FB condition ($n = 25$) there were 11 females, $M_{age} = 22.0$ months; in the TB condition there were 12 females, $M_{age} = 22.2$ months.

The study design was approved by the Ethics Committee of the University of Milano-Bicocca (protocol n. 348, 2018). The research was conducted according to guidelines laid down in the Declaration of Helsinki. Written informed consent was obtained by toddlers' parents who were free to withdraw their consent at any time during the study. We adopted a between-subject research design, such that half of the children watched the video representing the FB condition and the other half of the children watched the video of the TB condition.

5.2 Material

Two videos lasting about 1 min each were created *ad hoc* for the experiment (FB condition and TB condition), based on the test devised by Butterfill and Apperly (2013). The development of the videos took about 1 year. Changes were made following a pilot phase designed to test toddlers' interest in and attentiveness to the proposed content. The actress in the final videos was sitting at a table with a plywood occluder in front of her. The occluder (83 x 50 cm) was set in the middle of the table (which measured 100 x 60 cm), such that an 8.5 cm gap remained between the occluder and the edge of the table on either side. Thus, both the right and left borders of the table remained in view.

5.3 Procedure

During an initial familiarization phase, a double-sided puppet measuring 20 x 25 x 10 cm per side, with one side representing a wolf and the other a granny (see Figure 1) was displayed for about 40 s to each participant. Note that the toddlers used to play with puppets in their nursery. The researcher showed that the puppet was a wolf on one side and a granny on the other side and rotated the puppet twice to illustrate the two identities to the child, saying: "now you see the wolf…, now you see the granny" or "now you see the granny…, now you see the wolf". The toddlers were allowed to touch the puppet and rotate it themselves. The researcher concluded the familiarization phase by explaining that the puppet would appear in a video that the child would now be shown.

The experiment was carried out at nursery in a quiet room set up *ad hoc* for the study. Children participated one by one in a testing session. Each child first entered the room with their early childhood teacher, who was instructed to remain silent and to sit behind the child and slightly to one side. The toddler was seated on a child's chair set approximately 30 cm from a 15.55-inch color DELL monitor that had been fitted with a Tobii Pro Nano (sampling rate of 60 Hz) eye-tracker system. We used Tobii's standard 5-point calibration procedure for Tobii Studio software with the default settings (white dots on a full screen black background, with randomized ordered target points). In five cases, it was not possible to obtain a reliable calibration due to the children's difficulty in remaining still during the procedure.

5.3.1 Testing

In the false-belief (FB) condition, (1) the agent introduced herself by saying, "Hi! I'm Anna. Today, we are going to play together." (2) The puppet, with the granny-side visible to the child, popped up from below the table and moved to the left side of the table where it was no longer concealed by the occluder, prompting the agent to lean around the occluder and state: "I see a wolf". (3) The puppet moved back behind the occluder and stayed there. (4) At that point, Anna said: "I'm going out for a while", stood up, and left the room. While the agent was absent, the puppet rotated so that the child could see the wolf-side of the puppet. After 5 s, the agent returned stating "I'm back!" and sat down at the table again. (5) The puppet moved to the right edge of the table, the agent leaned around the occluder and stated: "I see a granny". (6) At that point, the puppet disappeared out of sight, below the table, and the

FIGURE 1
Material used in the study: Double-sided puppet (wolf and granny).

agent said: "Oh! The granny's gone". (7) She straightened up and immediately reached around the left side of the occluder with her right hand/arm, saying: "Where is the wolf?". On failing to find it, she sat up straight again.

In the true-belief (TB) condition, the first two scenes of the video were the same as in the FB condition. However, after the agent stated: "I see a wolf", (3) the puppet rotated to make the granny side visible to her. (4) Hence, the agent said: "And it's also a granny". (5) The puppet moved to the right edge of the table and then disappeared under the table out of sight. (6) The agent leaned around the occluder and stated: "Oh! The granny's gone". (7) She straightened up and immediately reached around the left side of the occluder with her right arm and hand, saying: "Where is the wolf?". On failing to find it, she straightened up again. Figure 2 represents the sequence of events in both the FB and TB conditions.

5.3.2 Coding

According to the violation-of-expectation paradigm (Margoni et al., 2024; Yeung et al., 2016), infants spend more time looking at and visiting specific areas of interest (AOI) when reacting to surprising or unexpected events. In this explorative new study, the definition of the AOIs was supported by a preliminary exploratory analysis of fixation maps, which highlighted three main attention clusters. We, therefore, identified three well separated areas of interest (AOIs), defined by the pixels corresponding to specific scenes: the area to the bottom left of the occluder, where the agent's hand movements can be seen (width: 550, height: 479); the area to the bottom right of the occluder, where the object disappears (width: 571, height: 426), and the area encompassing the agent's face (width: 191, height: 204). The three areas corresponded to the three main attention patterns shown by children and directed, respectively, to the face, relevant for grasping mental processes (Graham and LaBar, 2012), and to the hands, relevant for understanding the meaning of the actions produced by the agent.

The use of eye tracking within the VoE paradigm allows not only to measure where children look during a trial but also where, how often, and how long they look at specific AOIs. The eye tracking data collected for each participant was evaluated for acceptability based on quality, accuracy, and precision indices, having as a robust point of reference the thresholds specified for toddlers by Dalrymple et al. (2018) and by Carter and Luke (2020). In this study, for each AOI we analyzed the time interval (5 s) from when the actress starts moving—while asking "Where is the wolf?" and searching for the object by stretching her hand behind the occluder—until the end of the video stimulus.

Following Dalrymple et al. (2018) recommendations, we recorded the following eye tracking measures helpful to capture toddlers' signs of surprise during the critical time period: the number of fixations, the total duration of fixations, the number of visits, and the total duration of visits. In eye tracking research, a fixation occurs when the eye remains relatively still, focusing on a specific point within an AOI. A visit, on the other hand, is a sequence of fixations within an AOI. The visit duration metric includes both looking time and the time spent moving between fixations.

Based on the functioning of eye tracking, which records gaze data that—in turn—must be filtered to identify fixations, we used the canonical Tobii I-VT (Fixation) Preset filter. This is the default configuration of the I-VT filter for screen-based projects. For threshold, it uses a velocity threshold of 30°/s. In terms of parameters, it merges fixations if they are closer than 0.5 degrees and less than 75 ms apart, and it discards fixations shorter than 60 ms. The assignment of a fixation to an AOI was done through a spatial and temporal mapping process integrated into the software analysis module. The process involved data classification (Gaze Filter), and the identification of reference coordinates typical of screen-based projects, with data mapped directly onto the screen coordinates. The process of assigning a fixation to an AOI also involved spatial intersection to verify that the coordinates fell within the drawn rectangle.

We analyzed the eye tracking metrics (fixations and visits were our dependent measures) for each AOIs, comparing the outcomes across the FB and TB conditions. Under the FB condition (the agent is unaware of the dual identity of the object, i.e., wolf and granny), the agent's behavior in the video aligns with her presumable state of knowledge about the puppet identity. Conversely, in the TB condition (the agent is aware of the object's dual identity), the agent's behavior is inconsistent with her presumable state of knowledge. We explored whether the number of fixations and time spent looking at the identified AOIs would be greater in the TB condition—in which a violation of expectation occurred—than in the FB condition.

6 Results

First of all, we checked whether the data fulfill the criteria for parametric testing, that is normality of distribution, homogeneity of variances (homoscedasticity) and independence of observations. Then we conducted statistical analyses using IBM SPSS Version 29 software and applying the ANOVA test which requires that the

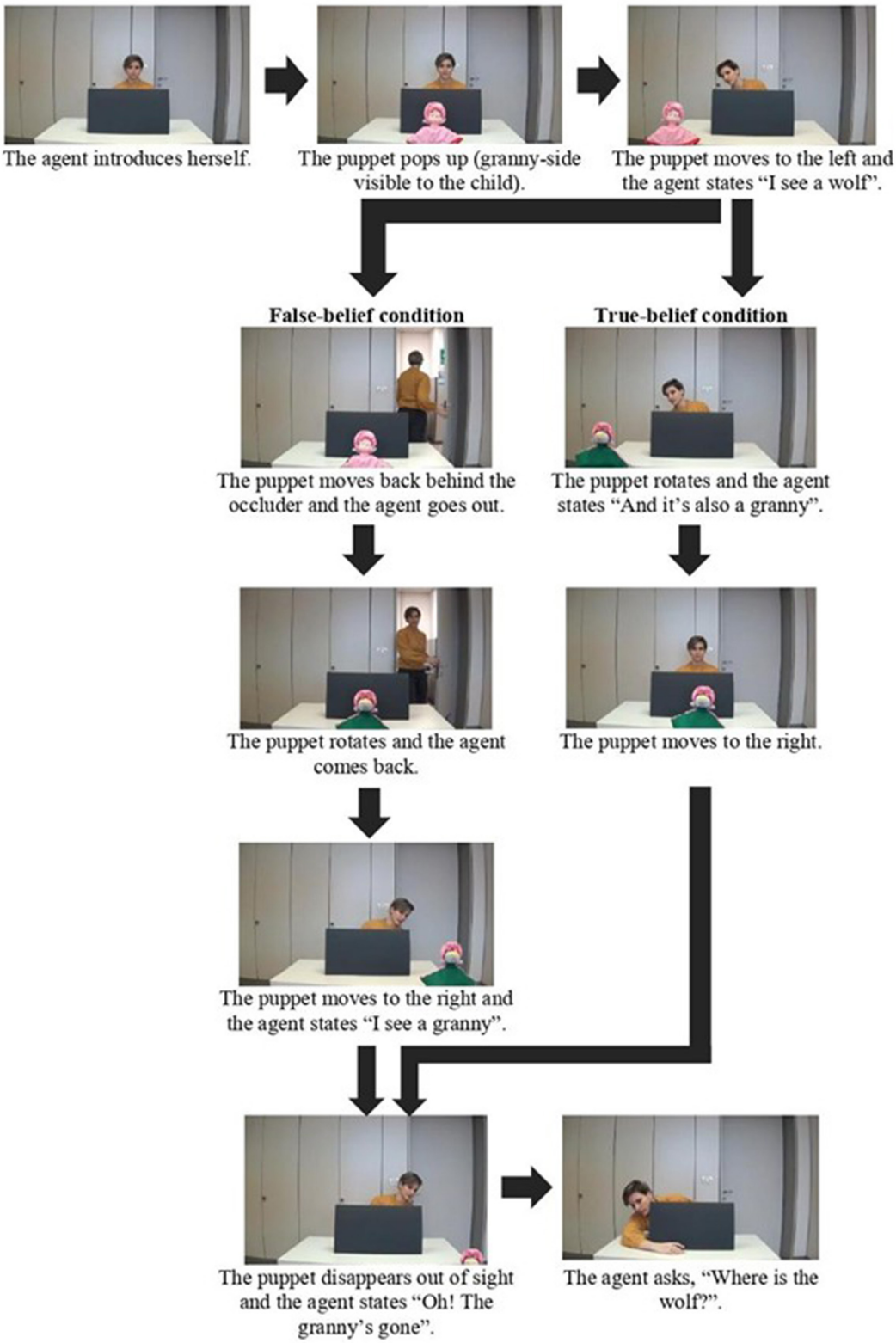

FIGURE 2
Experimental procedure: Schematic representation of the video sequences in the FB and TB conditions.

dependent variables are continuous, while the independent variable has at least two levels (Blaine, 2023). This section comprises two subsections outlining descriptive statistics for all study measures on one side and the outcomes of ANOVA test, exploring differences between the FB and TB conditions for any of the variables, on the other side.

Descriptive statistics. In Table 1, means and standard deviations (in parentheses) of the target variables for the AOI surrounding the agent's face, observed from the moment she asks: "Where is the wolf?" until the end of the video stimulus, are reported. Table 2 presents the means and standard deviations for the AOI where the agent's hand was visible, observed in the same time interval; Table 3, finally, refers to the AOI where the puppet disappears on the right side of occluder. More precisely, Tables 1, 2 and 3 report, for each AOIs and for each research condition (FB vs. TB), the mean and standard deviation of the following metrics: the number of fixations, the total duration of fixations, as well as the number of visits, and the total duration of visits in milliseconds.

Inferential statistics. In order to explore whether there were statistically significant differences between the FB and TB conditions we applied the ANOVA test for any of the study variables. The level of statistical significance was set at $p = 0.01$ to ensure robust and reliable statistical inferences given the exploratory nature of the study (Blaine, 2023). Regarding the AOI of the agent's face, we found no statistically significant differences for the number of fixations ($F = 0.290$, df $= 1$, $p = 0.593$), the total duration of fixations ($F = 0.749$, df $= 1$, $p = 0.391$), the number of visits ($F = 1.80$, df $= 1$, $p = 0.186$), and the total duration of visits ($F = 0.544$, df $= 1$, $p = 0.464$).

With respect to the AOI of the agent's hand (left side of the occluder), we found no statistically significant difference between the FB and the TB conditions for the number of fixations ($F = 0.687$, df $= 1$, $p = 0.411$), the total duration of fixations ($F = 1.75$, df $= 1$, $p = 0.192$). Moreover, no significant difference for the number of visits ($F = 4.19$, df $= 1$, $p = 0.046$; $\eta^2 = 0.014$) and the total duration of visits ($F = 3.70$, df $= 1$, $p = 0.060$) emerged.

Finally, with respect to the AOI of the right side of the occluder no significant differences emerged for any of the variables under study: the number of fixations ($F = 0.780$, df $= 1$, $p = 0.381$), the total duration of fixations ($F = 0.005$, df $= 1$, $p = 0.947$), the number of visits ($F = 0.073$, df $= 1$, $p = 0.789$), and the total duration of visits ($F = 0.041$, df $= 1$, $p = 0.841$).

TABLE 1 Means and standard deviations of the target variables (agent's face).

Variables	FB condition (mean and DS)	TB condition (mean and DS)
Number of fixations	8.16 (2.77)	7.64 (3.94)
Total duration of fixations	4,003.20 (1,488.46)	3,559.00 (2,090.64)
Number of visits	3.60 (1.00)	3.12 (1.48)
Total duration of visits	4,149.72 (1,525.26)	3,770.28 (2,071.39)

TABLE 2 Means and standard deviations of the target variables (left side of occluder).

Variables	FB condition (mean and DS)	TB condition (mean and DS)
Number of fixations	5.52 (2.85)	4.84 (2.93)
Total duration of fixations	2,426.64 (1,179.09)	1,830.20 (1,060.56)
Number of visits	2.40 (0.95)	1.88 (0.83)
Total duration of visits	2,555.16 (1,221.07)	1,918.76 (1,113.48)

TABLE 3 Means and standard deviations of the target variables (right side of occluder).

Variables	FB condition (mean and DS)	TB condition (mean and DS)
Number of fixations	0.32 (0.47)	0.48 (0.77)
Total duration of fixations	142.56 (270.24)	147.72 (271.82)
Number of visits	0.32 (0.47)	0.36 (0.56)
Total duration of visits	142.56 (270.24)	158.40 (283.67)

7 Discussion and conclusion

The aim of this study was to try to adapt a test outlined by Butterfill and Apperly (2013) to explore whether toddlers, in their 2nd year of life, can understand another person's false belief about the identity of an object, when this object has a dual identity (a granny-wolf puppet). To the best of our knowledge, this is the first study operationalizing an identity task that cannot be overcome merely by relying on behavioral rules (as said before, simple rules that directly connect an observable situation to a behavior, without going through mental states) and, therefore, can play a significant role in the debate on one-system vs. two-system accounts of early theory of mind.

In this study, we set up a novel unexpected identity task to be assessed via eye tracking. In this test, a puppet looked different from different perspectives (appearing as either a wolf or a granny), thus bearing a dual identity. Each toddler in our sample individually viewed a video and observed the behavior of an agent (actress). In the FB condition video, in which only the watching child (and not the agent) could see the double-sided puppet rotating, the agent reached around the occluder in an attempt to find the wolf where she had last seen it, being ostensibly unaware of the object's dual identity. In the TB condition video, the agent saw a granny-wolf puppet rotating and moving away, and then—even though she had no ostensible reason to believe that the wolf was still there—still reached around the occluder in an attempt to grasp the wolf (as shown in the last fragment of Figure 2).

Following Butterfill and Apperly (2013), if toddlers showed signs of surprise when the agent looked for an object under the TB condition, but not when she did the same under the FB condition, it could be assumed that they had understood how the object appeared to the actress. This might suggest that they were able to understand the agent's false belief as such. A different experimental result would not support that theoretical hypothesis.

In the present exploratory study, we found that looking times and number of visits did not differ significantly between the TB and FB conditions, with respect to the three AOIs identified, that is the agent's face, and both the left and the right side of the occluder, respectively. In other words, the present findings did not indicate that the children in our sample were more surprised by the agent's behavior in the TB condition when, despite her knowledge of the object's dual identity, the actress continued to search for the wolf-object after the granny-object had left the scene. These results, therefore, do not show that children were able to attribute false beliefs to the agent. Indeed, these findings are compatible, with

the suggestion that toddlers, from 20 to 24 months of age, could possess only a minimal theory of mind, enabling them to track the puppet's position relative to the actress and recall when the actress last encountered the puppet.

The findings of this exploratory study are not in line with those of other investigations based on different identity tasks which reached the conclusion that infants and toddlers could attribute false belief to others (Buttelmann and Kovács, 2019; Buttelmann et al., 2015; Scott and Baillargeon, 2009). Note that, compared to these studies, the use of a single object—the double-sided puppet (granny and wolf)—in our own research required participants to have a genuine appreciation of dual identity in order to successfully attribute false belief. Indeed, the present findings are in line with those of other studies which, investigating the developmental trajectory of false belief reasoning, reported negative findings and failure to replicate body of data on early implicit theory of mind. The failure to replicate those studies that, starting with Onishi and Baillargeon (2005), have supported the presence of implicit ToM in infants has led, in our opinion, to a certain slowdown in studies on implicit ToM. This slowdown has also affected the implementation of studies based on the use of identity task. With the present exploratory study, we aimed at drawing researchers' attention to the use of this type of task in order to contribute to the comprehension of children's early mentalizing abilities.

Nevertheless, our findings should be taken with extreme caution and appreciated as the evidence of the effort to operationalize a complex research apparatus using eye tracking technology. On one hand, the lack of statistically significant differences between FB condition and TB condition does not allow us to reject the null hypothesis. On the other hand, this conclusion should be considered with caution and interpreted as a stimulus for future studies. In fact, the lack of significant differences in a violation-of-expectation paradigm could be linked to various factors that affected the validity of the experiment. As well described by Dalrymple et al. (2018), data collection with toddlers can be, in general, less accurate than with older children and adults (e.g., moving excessively, not looking at the screen, and poor calibration quality) and require greater interpretative caution.

Discussing the results in more detail, one possible explanation for null findings lies in the nature of the task. We are aware that the unexpected identity task may have been too difficult considering the age of the children. It has required attention for about a minute during which numerous actions occur involving both the protagonist (she observes, speaks, in one case she moves away, returns) and the object itself (it appears in different guises, it moves, and it rotates).

Second, even if 20-month-old children begin to show the first forms of perceptual perspective taking (Moll and Meltzoff, 2011), the ability to understand that another person may see an object different from the one they see or that another person does not see what is hidden from their view may still be present in a very rudimentary form. Therefore, this cognitive limitation may have played an important role in the child's visual behavior during the experiment in both conditions.

Third, a key concern is the mismatch between FB and TB conditions. In fact, in the FB condition the protagonist witnesses the wolf puppet move from the right to the central position behind the occluder, whereas in the TB condition the protagonist observes the granny puppet move from the right to the center. Consequently, it could be unclear whether these differences stem from variations in the protagonist's belief state or from uncontrolled scenario-specific confounds (e.g., spatial location of identity change).

Fourth, in the current design during the familiarization phase the experimenter (an adult) explains the puppet's features to the toddler, implying the experimenter's knowledge of the dual-sided identity. If the toddlers infer that the actress shares this knowledge from the start, the core manipulation of initial belief state would be invalidated. Moreover, if the children do not hold the assumption that the actress believes that a single object is behind the screen (e.g., the puppet present at the end is the same one observed during the rotation phase), then the actress' expectation of the wolf puppet's presence in the true belief condition could be perceived as reasonable.

Despite these limitations which have an impact on the interpretation of the results of this study, we hope to have contributed to the ongoing scientific debate on theory of mind in early childhood, particularly adopting of a new identity tasks and measuring participants' fixations and visits of specific areas of interests as dependent variables. This debate, lately, encompasses several critical issues, including the problem of non-replication (Kulke et al., 2018; Sabbagh and Paulus, 2018), the question of whether certain evidence for early false belief attribution might be more appropriately explained by "low-level novelty" for instance their colors, and movements, (see Heyes, 2014, for a more in-depth treatment of the hypothesis that looking behaviors may be driven by lower-level properties of test stimuli), and the failure to identify a clear developmental trajectory in young children's theory of mind (Barone et al., 2019; Poulin-Dubois, 2025; Scott et al., 2022). Our study, whose exploratory nature we once again emphasize, aimed to reinvigorate interest in the use of identity tasks to investigate young children's' early ability to attribute false beliefs.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Ethics Committee of the University of Milano-Bicocca. The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation in this study was provided by the participants' legal guardians/next of kin. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

IG: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. EC: Conceptualization, Data curation, Investigation, Methodology, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. ED: Conceptualization, Data curation, Methodology, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declared that financial support was received for this work and/or its publication. This study was funded by a grant from the University of Milano Bicocca (2023-ATE-0213) awarded to IG.

Acknowledgments

We would like to thank the Lab-PSE collaborators (https://www.labpse.it/en/) for their initial help in preparing the experiment and SR LABS Company for their fundamental consultancy in data coding. We wish to thank the research participants (the toddlers and their family) whose contribution made this project possible as well as the early childhood education services involved in this study, and Clare O'Sullivan for proofreading an initial version of this paper.

Conflict of interest

The author(s) declared that this work was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author IG declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Generative AI statement

The author(s) declared that generative AI was not used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

References

Apperly, I. A., and Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychol. Rev.* 116, 953–970. doi: 10.1037/a0016923

Baillargeon, R., Scott, R. M., and Bian, L. (2016). Psychological reasoning in infancy. *Annu. Rev. Psychol.* 67, 159–186. doi: 10.1146/annurev-psych-010213-115033

Baillargeon, R., Scott, R. M., and He, Z. (2010). False-belief understanding in infants. *Trends Cogn. Sci.* 14, 110–118. doi: 10.1016/j.tics.2009.12.006

Barone, P., Corradi, G., and Gomila, A. (2019). Infants' performance in spontaneous-response false belief tasks: a review and meta-analysis. *Infant Behav. Dev.* 57:101350. doi: 10.1016/j.infbeh.2019.101350

Blaine, B. (2023). *Introductory Applied Statistics*. Cham, Switzerland: Springer. doi: 10.1007/978-3-031-27741-2

Buttelmann, D., Carpenter, M., and 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

Buttelmann, F., and Kovács, Á. M. (2019). 14-month-olds anticipate others' actions based on their belief about an object's identity. *Infancy* 24, 738–751. doi: 10.1111/infa.12303

Buttelmann, F., Suhrke, J., and Buttelmann, D. (2015). What you get is what you believe: eighteen-month-olds demonstrate belief understanding in an unexpected-identity task. *J. Exp. Child Psychol.* 131, 94–103. doi: 10.1016/j.jecp.2014.11.009

Butterfill, S. A., and Apperly, I. A. (2013). How to construct a minimal theory of mind. *Mind Lang.* 28, 606–637. doi: 10.1111/mila.12036

Carruthers, P. (2013). Mindreading in infancy. *Mind Lang.* 28, 141–172. doi: 10.1111/mila.12014

Carruthers, P. (2016). Two systems for mindreading? *Rev. Phil. Psych.* 7, 141–162. doi: 10.1007/s13164-015-0259-y

Carter, B. T., and Luke, S. G. (2020). Best practices in eye tracking research. *Int. J. Psychophysiol.* 155, 49–62. doi: 10.1016/j.ijpsycho.2020.05.010

Caselli, M. C., Bello, A., Rinaldi, P., Stefanini, S., and Pasqualetti, P. (2018). *Il primo vocabolario del bambino: gesti, parole e frasi: valori di riferimento fra 8 e 36 mesi delle forme complete e delle forme brevi del questionario MacArthur-Bates CDI., 2.* Milano: Franco Angeli.

Conte, E., Ornaghi, V., Grazzani, I., Pepe A. and Cavioni, V. (2029). Emotion knowledhe, theory of mind, and language in young children: Testing a comprehensive conceptual model. Front. Psycol. 10:2144. doi: 10.3389/fpsyg.2029.02144

Dalrymple, K. A., Manner, M. D., Harmelink, K. A., Teska, E. P., and Elison, J. T. (2018). An examination of recording accuracy and precision from eye tracking data from toddlerhood to adulthood. *Front. Psychol.* 9:803. doi: 10.3389/fpsyg.2018.00803

Faul, F., Erdfelder, E., Lang, A.-G., and Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav. Res. Methods* 39, 175–191. doi: 10.3758/bf03193146

Fenson, L., Pethick, S., Renda, C., Cox, J. L., Dale, P. S., and Reznick, J. S. (2000). Short-form versions of the MacArthur communicative development inventories. *Appl. Psycholing.* 21, 95–116. doi: 10.1017/S0142716400001053

Fizke, E., Butterfill, S., Van De Loo, L., Reindl, E., and Rakoczy, H. (2017). Are there signature limits in early theory of mind? *J. Exp. Child Psychol.* 162, 209–224. doi: 10.1016/j.jecp.2017.05.005

Flavell, J. H., Flavell, E. R., and Green, F. L. (1983). Development of the appearance-reality distinction. *Cogn. Psychol.* 15, 95–120. doi: 10.1016/0010-0285(83)90005-1

Fodor, J. A. (1978). Propositional attitudes. *Monist* 61, 501–523. doi: 10.5840/monist197861444

Gopnik, A., and Astington, J. T. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child Dev.* 59, 26–37. doi: 10.2307/1130386

Graham, R., and LaBar, K. S. (2012). Neurocognitive mechanisms of gaze-expression interactions in face processing and social attention. *Neuropsychologia* 50, 553–566. doi: 10.1016/j.neuropsychologia.2012.01.019

Heyes, C. (2014). False belief in infancy: a fresh look. *Dev. Sci.* 17, 647–659. doi: 10.1111/desc.12148

Hogrefe, G.-J., Wimmer, H., and Perner, J. (1986). Ignorance versus false belief: a developmental lag in attribution of epistemic states. *Child Dev.* 57:567. doi: 10.2307/1130337

Kampis, D., and Kovács, Á. M. (2021). Seeing the world from others' perspective: 14-month-olds show altercentric modulation effects by others' beliefs. *Open Mind* 5, 189–207. doi: 10.1162/opmi_a_00050

Kulke, L., Reiß, M., Krist, H., and Rakoczy, H. (2018). How robust are anticipatory looking measures of theory of mind? Replication attempts across the life span. *Cogn. Dev.* 46, 97–111. doi: 10.1016/j.cogdev.2017.09.001

Low, J., Apperly, I. A., Butterfill, S. A., and Rakoczy, H. (2016). Cognitive architecture of belief reasoning in children and adults: a primer on the two-systems account. *Child Dev. Perspectives* 10, 184–189. doi: 10.1111/cdep.12183

Low, J., and Watts, J. (2013). Attributing false beliefs about object identity reveals a signature blind spot in humans' efficient mind-reading system. *Psychol. Sci.* 24, 305–311. doi: 10.1177/0956797612451469

Margoni, F., Surian, L., and Baillargeon, R. (2024). The violation-of-expectation paradigm: a conceptual overview. *Psychol. Rev.* 131, 716–748. doi: 10.1037/rev0000450

Moll, H. and Meltzoff, A.N. (2011). How does it look? level 2 perspective-taking at 36 months of age. *Child Dev.* 82, 661–673. doi: 10.1111/j.1467-8624.2010.01571.x

Onishi, K. H., and Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science* 308, 255–258. doi: 10.1126/science.1107621

Perner, J., and Ruffman, T. (2005). Infants' insight into the mind: how deep? *Science* 308, 214–216. doi: 10.1126/science.1111656

Poulin-Dubois, D. (2025). *Theory of Mind in Childhood*. Cambridge, MA: Cambridge University Press. doi: 10.1017/9781009413503

Rakoczy, H. (2012). Do infants have a theory of mind? *British J. Dev. Psycho.* 30, 59–74. doi: 10.1111/j.2044-835X.2011.02061.x

Ruffman, T. (2023). Belief it or not: how children construct a theory of mind. *Child Dev. Perspectives* 17, 106–112. doi: 10.1111/cdep.12483

Sabbagh, M. A., and Paulus, M. (2018). Replication studies of implicit false belief with infants and toddlers. *Cogn. Dev.* 46, 1–3. doi: 10.1016/j.cogdev.2018.07.003

Scott, R.M., Roby, E., and Baillargeon, R. (2022). "How sophisticated is infants' theory of mind?," in *The Cambridge Handbook of Cognitive Development*, eds. O. Houdè and G. Borst (Cambridge, MA: Cambridge University Press), 242–268. doi: 10.1017/9781108399838.015

Scott, R. M., and Baillargeon, R. (2009). Which penguin is this? Attributing false beliefs about object identity at 18 months. *Child Dev.* 80, 1172–1196. doi: 10.1111/j.1467-8624.2009.01324.x

Scott, R. M., Richman, J. C., and Baillargeon, R. (2015). Infants understand deceptive intentions to implant false beliefs about identity: new evidence for early mentalistic reasoning. *Cogn. Psychol.* 82, 32–56. doi: 10.1016/j.cogpsych.2015.08.003

Sodian, B., Kristen-Antonow, S., and Kloo, D. (2020). How does children's theory of mind become explicit? A review of longitudinal findings. *Child Dev. Perspectives* 14, 171–177. doi: 10.1111/cdep.12381

Southgate, V., Senju, A., and Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychol. Sci.* 18, 587–592. doi: 10.1111/j.1467-9280.2007.01944.x

Wellman, H. M., Cross, D., and Watson, J. (2001). Meta-analysis of theory-of-mind development: the truth about false belief. *Child Dev.* 72, 655–684. doi: 10.1111/1467-8624.00304

Wimmer, H., and 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

Yeung, H. H., Denison, S., and Johnson, S. P. (2016). Infants' looking to surprising events: when eye-tracking reveals more than looking time. *PLoS ONE* 11:e0164277. doi: 10.1371/journal.pone.0164277