Cognition 213 (2021) 104640

Contents lists available at ScienceDirect

Cognition

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

Can infants adopt underspecified contents into attributed beliefs? Representational prerequisites of theory of mind☆

Ágnes Melinda Kovács [a,*], Ernő Téglás [a], Gergely Csibra [a,b]

[a] Department of Cognitive Science, Cognitive Development Center, Central European University, Budapest, Hungary
[b] Department of Psychological Sciences, Birkbeck, University of London, UK

A R T I C L E I N F O

Keywords:
Theory of mind
Infants
Belief files
Underspecified beliefs
True and false beliefs

A B S T R A C T

Recent evidence suggests that young infants, as well as nonhuman apes, can anticipate others' behavior based on their false beliefs. While such behaviors have been proposed to be accounted by simple associations between agents, objects, and locations, human adults are undoubtedly endowed with sophisticated theory of mind abilities. For example, they can attribute mental contents about abstract or non-existing entities, or beliefs whose content is poorly specified. While such endeavors may be human specific, it is unclear whether the representational apparatus that allows for encoding such beliefs is present early in development. In four experiments we asked whether 15-month-old infants are able to attribute beliefs with underspecified content, update their content later, and maintain attributed beliefs that are unknown to be true or false. In Experiment 1, infants observed as an agent hid an object to an unspecified location. This location was later revealed in the absence or presence of the agent, and the object was then hidden again to an unspecified location. Then the infants could search for the object while the agent was away. Their search was biased to the revealed location (that could be represented as the potential content of the agent's belief when she had not witnessed the re-hiding), suggesting that they (1) first attributed an underspecified belief to the agent, (2) later updated the content of this belief, and (3) were primed by this content in their own action even though its validity was unknown. This priming effect was absent when the agent witnessed the re-hiding of the object, and thus her belief about the earlier location of the object did not have to be sustained. The same effect was observed when infants searched for a different toy (Experiment 2) or when an additional spatial transformation was introduced (Experiment 4), but not when the spatial transformation disrupted belief updating (Experiment 3). These data suggest that infants' representational apparatus is prepared to efficiently track other agents' beliefs online, encode underspecified beliefs and define their content later, possibly reflecting a crucial characteristic of mature theory of mind: using a meta-representational format for ascribed beliefs.

1. Introduction

In the last fifteen years, a wealth of studies have targeted the question whether preverbal infants show an understanding of others' mental states. Research has found that infants are surprised when a person searches for an object at a location that is incongruent with where she believes it to be (Onishi & Baillargeon, 2005; Scott, Richman, & Baillargeon, 2015; Träuble, Marinović, & Pauen, 2010), or when an outcome is incongruent with a protagonist's beliefs (Kovács, Téglás, & Endress, 2010), compared to belief-congruent outcomes. Infants may not simply react differently to such scenes after the fact, but they may also exploit their belief representations to anticipate that a protagonist will reach for an object at the location where she falsely believes the object to be (Southgate, Senju, & Csibra, 2007; Thoermer, Sodian, Vuori, Perst, & Kristen, 2012; but see Kampis, Karman, Csibra, Southgate, & Hernik, 2020; and Kulke, Reiß, Krist, & Rakoczy, 2018), or based on her false belief about the identity of the object (Buttelmann & Kovács, 2019). Infants also seem to consider others' beliefs when they engage in helping others (Buttelmann, Carpenter, & Tomasello, 2009), when they communicate through pointing (Knudsen & Liszkowski, 2012), and when they disambiguate referential communication (Király, Oláh, Csibra, & Kovács, 2018; Southgate, Chevallier, & Csibra, 2010). Such belief

☆ This paper is a part of special issue "Special Issue in Honour of Jacques Mehler, Cognition's founding editor".

* Corresponding author at: Central European University, Nádor u. 9., Budapest 1051, Hungary.
E-mail address: kovacsag@ceu.edu (Á.M. Kovács).

https://doi.org/10.1016/j.cognition.2021.104640
Received 9 September 2020; Received in revised form 16 February 2021; Accepted 17 February 2021
Available online 20 March 2021

attributions are not restricted to a specific modality, as infants can flexibly update false beliefs that have been formed based on information from one modality (e.g., vision), when receiving new information from a different domain (e.g., language: Song, Onishi, Baillargeon, & Fisher, 2008; communication: Tauzin & Gergely, 2018).

The accumulating evidence involving various situations and methodologies from more than 30 studies, including actual and conceptual replications, point to an understanding of false beliefs in infancy (see Scott & Baillargeon, 2017 for a review). Recently, however, some studies did not succeed in reproducing earlier findings, or resulted in partial replications (for overviews see Baillargeon, Buttelmann, & Southgate, 2018; Barone & Gomila, 2020; Poulin-Dubois et al., 2018). It will be a task for future research to determine how the accumulating positive and negative evidence should be interpreted; whether the subtle methodological differences may have contributed to the different findings (Baillargeon et al., 2018), and whether multi-lab efforts will result in identifying key preconditions that lead to reliable task performance. In any case, besides targeting the reliability of specific tasks, research should also aim at uncovering the processes that may actually drive infants' performance.

Experimental data revealing infants' sensitivity to others' mental states has been taken to support the idea that Theory of Mind (ToM) is an early emerging human ability (Carruthers, 2013). When evidence from infant research was complemented by findings from other species suggesting that other animals seem to fail on ToM tasks (Kaminski, Call, & Tomasello, 2008; Martin & Santos, 2014), initial commitments to human specificity of these processes became strengthened. Nevertheless, nonhuman animals were found to pass knowledge/ignorance tasks, for instance, relying on computations of seeing/not seeing (Flombaum & Santos, 2005; Hare, Call, & Tomasello, 2001). The successful performance on these tasks, however, were often interpreted in terms of non-mentalistic processes (e.g., the 'evil eye' hypothesis, or behavioral rules; Kaminski et al., 2008), as crucial evidence about nonhuman animals encoding others' reality incongruent states was missing. While until recently evidence seemed to support the claim that ToM is a capacity that is unique to humans, recent research provided new data suggesting that apes and monkeys also take into account others' mental states (specifically, their false beliefs) when predicting their actions (Hayashi et al., 2020; Kano, Krupenye, Hirata, Tomonaga, & Call, 2019; Krupenye, Kano, Hirata, Call, & Tomasello, 2016).

These recent findings, which reveal comparable performance in infants and nonhuman apes, lead to a set of questions. Do they imply that both infants and apes possess fully-fledged ToM abilities? Could the results obtained be explained by lower level mechanism, such as encoding associations or behavioral rules (Perner & Ruffman, 2005), or simply exploiting the relations between agents, objects, and locations (Apperly & Butterfill, 2009)? A further possibility is that both infants and other animals may possess some ToM abilities, however, human ToM may still be qualitatively different because it relies on representational and conceptual abilities that may only be found in the human cognitive repertoire.

Here we argue that efficient belief attribution in humans may benefit from dedicated representational structures that capture the information integrated in a belief. This proposal has at its core a set of cognitive mechanisms that track others' beliefs online and entails forming prospective belief attributions as the events unfold (Király et al., 2018; Kovács, 2016; Leslie, 1994). Such online belief tracking may require identifying potential belief holders (agents), setting up 'belief files', and allocating content to them. These belief files enable ToM processes to store information about other agents' beliefs in a specific format that would support efficient encoding and updating of belief related information (Kovács, 2016).

Research on belief attribution has so far focused on contrasting true and false beliefs, where the distinction is made on the basis of the external validity of the belief content. This makes sense from a philosophical point of view, since a crucial property of beliefs (as a kind of representation) is that, although they aim at truth (Velleman, 2000), they can misrepresent reality. However, it is a mistake to take this contrast as the defining feature of, and as the crucial test of the ability for, theory of mind (Bloom & German, 2000), because the cognitive mechanisms for tracking online others' beliefs may not be concerned with the external validity of belief contents. On the one hand, one can predict someone's future actions on the basis of the beliefs attributed to her even if it is unknown whether those beliefs are true or false. If John believes that Mary is at home (as he left her there), we can predict that he will try to find her there, even if we do not know whether Mary is at home or not (for example, because we saw her in the front of the house and we are not sure whether she was leaving or returning). In this case, John has an outdated belief (its original justification no longer holds), which can be either true or false, but nevertheless its attribution to John supports specific action predictions.

On the other hand, even explicitly knowing that someone has a true belief about some state of affairs may not allow us to predict her actions if the attributed beliefs are underspecified, i.e., if their contents do not carry sufficient information for action prediction. If John hides an object in a room to an unknown location and then leaves, we may represent that 'John believes [x (is at location y)]', where 'x' stands for an unspecified object and 'y' stands for an unspecified location at the time of belief attribution. The belief files dedicated to track these contents should nevertheless capture the epistemically indeterminate nature of such attributions. As such, the content of this attributed belief may be characterized by some referential uncertainty regarding the identity, location, or other properties of entities. Importantly, in such cases there is some information that is known to the actor, this information may even be thought to be correct, but it is currently not available to the observer. While such belief attributions may invoke a set of hypotheses about the feasible contents (e.g., the possible hiding locations used by John), one may sustain these representations as informationally incomplete, which may not allow certain inferences to be drawn from it (for example, predicting where in the room would John go to find the object). Nevertheless, such underspecified belief attributions may have well specified uses. For example, if we find John's object in a box, we may complete the content of the belief file post hoc and use it for generating predictions about John's subsequent actions. This is made possible by the representational skeleton of belief files that may contain placeholders for the belief-content and the belief holder, and may allow for the rapid individual updating of the different functionally separate subcomponents (Kovács, 2016). After forming such an underspecified belief file, when information related to its exact content is revealed, only the content-placeholder will be updated, and one will not have to re-initiate the whole belief computation process.

Evidence pointing to the ability to attribute underspecified beliefs in infancy (in which one or more arguments of the belief are not specified at the time of attribution) would constitute a strong case against proposals according to which infants succeed in ToM tasks by forming associations or encountering/registering relations among agents, objects, or locations. Consider, for instance, a situation, in which an infant is watching her mother putting a toy into one of two opaque boxes without the infant being able to see in which one. Would the infant be able to represent that the mother knows where the object is, although the infant herself does not know? Now assume that while the mother is away the baby's brother comes in, finds the object in a box, and takes it away. Would the infant infer that, since the mother was not present when the brother found the toy, she must still believe the object to be in the box? The precondition for such an inference is a premise that the infant has ascribed a belief about the location of the toy to the mother to start with (i.e., a belief with underspecified content). In the absence of this attribution, updating the underspecified content captured in the attributed belief to a well-defined one when evidence becomes available cannot be performed without detailed recollection of previous events. Since the mother is not present when the location of the toy is revealed, unless the infant secured some belief ascription and sustained a belief file about her representation of the location of the toy, he would have no reason to think about her or her beliefs. Furthermore, since the mother has not come back yet, there is no pressing need to perform retrospective inferences either (Király et al., 2018). Nevertheless, it may be adaptive for infants to form an underspecified belief file and update it later, as such attributions may provide a connection between the past and present social events, allowing them to prepare for prospective interactions.

To investigate whether young infants possess the necessary cognitive architecture for attributing and updating belief files with an underspecified content and for using them even if their external validity is uncertain, we developed a task that simulates the situation described above. In the task, infants have the opportunity to attribute an unspecified belief to an agent, to update it when information becomes available, and to rely on this attributed belief in their own actions. Our dependent measure assumes that, when infants have no sufficient information to plan their own goal-directed action, they would be biased by the information contained in the beliefs they attribute to others either directly, or via an action prediction drawn from it (cf. Kovács et al., 2010; Southgate, 2019; Southgate & Vernetti, 2014). If we find such a bias, i.e., if infants' own choice of actions is primed by the updated belief content attributed to someone else in such cases, it would provide experimental support for the hypothesis that, just like adults, they can attribute and update belief files with underspecified content, and hence use representational structures that may rely on placeholders in belief attribution. Furthermore, such a result would also provide an argument against proposals that account for evidence of early ToM by appealing to the formation of simple associations or relations between the agent, the object and the location – which are never present simultaneously in our scenarios.

2. Experiment 1

We developed a new invisible hiding task for 15-month-olds (based on Call & Tomasello, 1999). The task has two conditions, differing in the kind of belief attributable to an interactive partner at the time of taking the dependent measure. In the Agent absent condition, infants watch an experimenter (E1) placing a toy into one of two opaque boxes such that they are unable to see in which one, then she leaves the scene. In the absence of E1, another experimenter (E2) reveals the location of the toy, and then E2 performs a second invisible hiding. Here we ask whether at the invisible hiding performed by E1 infants would be able to ascribe an underspecified belief to E1 about the location of the toy when they themselves have no knowledge about this. This event introduces a typical case of asymmetric knowledge: E1 has private knowledge about the location of the object, but the observer (here, the infant) does not have access to it. However, this lack of first person knowledge may not prevent infants from attributing a belief to E1, but the content of this attributed belief would remain underspecified. If infants attribute such a belief, would they be able to specify or update the content of such belief post hoc, when the location of the object is revealed by E2, and would they sustain such attributed representations even though a second invisible hiding takes place? Although the second hiding makes the attributed belief outdated with respect to the actual state of the world, and its truth value becomes unspecified, it may be useful to sustain it as it could support making predictions about E1's behavior, in case she comes back. Here we test whether infants (i) attribute underspecified beliefs to others, (ii) update the content of such beliefs post hoc when crucial information becomes available, and (iii) sustain such representations even if their validity is unknown. If infants resort to the content of such beliefs when they have no better information, they would be primed to search for the toy in the location where E1 believes the object to be.

In contrast, such a priming effect would not be expected if E1 witnesses the second invisible hiding. In this case her belief about the initial location does not need to be sustained, rather, it has to be replaced with an underspecified belief about the new location of the toy. This new underspecified belief, although can be considered a true belief, does not support an exact behavioral prediction regarding the agent and cannot prime infants' search behavior. We implemented these events in the Agent Present condition that served as a contrast to the Agent Absent condition. Note that in both conditions infants are ignorant about the actual location of the toy at the moment of search, thus, in case no other factors influence their behavior, they should search randomly. Furthermore, the initial location of the object is revealed in both conditions, thus if infants are just influenced by where they have last seen the object, they should search similarly in the two conditions.

In order to make infants' search behavior unambiguous, we designed the task to include two target locations relatively far apart from each other. Since approaching these locations required locomotion, we chose to test 15-month-old infants, who would likely able to perform this task.

2.1. Methods

2.1.1. Participants

The participants were 24 15-month-old infants (mean age: 15.12 months, range: 15;00 to 15;27, 15 girls). We aimed for a sample size of 24 based on comparable studies (e.g., Buttelmann, Carpenter, & Tomasello, 2009). Additional infants were excluded from the analysis because of failing to choose a box in one or both test trials ($N = 3$), crying, fussiness or not finishing the task ($N = 4$), experimenter error ($N = 2$) or because they went behind the boxes to see where the object was before choosing (N = 3).

Parents in all experiments signed an informed consent before taking part in the study, and all participants received a small toy as a gift for their participation. This research complied with relevant ethical regulations and was approved by the Hungarian Ethical Review Committee for Research in Psychology (EPKEB).

2.1.2. Materials

In the familiarization phase, we used two small objects, which were hidden in one of two boxes in a counterbalanced order: a white plush chicken and a white and blue plush smurf figure. For test phase, we used two further objects in a counterbalanced order: a purple plush squeaking elephant and a green plush squeaking horse. The objects had similar size around 12–15 cm. The hiding locations were two identical white boxes (see Fig. 1).

2.1.3. Procedure

The infant and the parent met both experimenters in the waiting room, where they engaged in brief interactions, playing with toys before the start of the study. Both experimenters were present until the infant and the parent entered the experimental room accompanied only by E1. At the beginning of each trial, the infants were seated in their caregiver's lap on a beanbag approximately 2.5 m from the two boxes that served as hiding locations. Participants were exposed to 4 (or 6) familiarization trials and 2 test trials.

The purpose of the familiarization trials was to engage the infants in the repeated hiding-searching game, and to draw attention to the fact that E1 may hold information about the location of, and thus can help finding, the hidden toy. At the beginning of the familiarization trials the boxes were placed centrally close to each other with their opening oriented towards E1, who was facing the child from the other side of the boxes. E1 kneeled behind the boxes, lifted a toy at the midline of the two boxes, attracted infants' attention by saying "Look!" and performed an invisible hiding the following manner. She lowered the object at the midline of the boxes with two hands, separated the hands only when the object was occluded by the boxes, and placed simultaneously one hand in one box and the other hand in the other in a way that the child could not see in which box the object was hidden. Afterwards E1 moved the two boxes approximately 1.2 m apart (the length of her arms) symmetrically from the midline, stood up, stepped back, looked and pointed at the baited box, and said to the child "Come and find it! Where is the

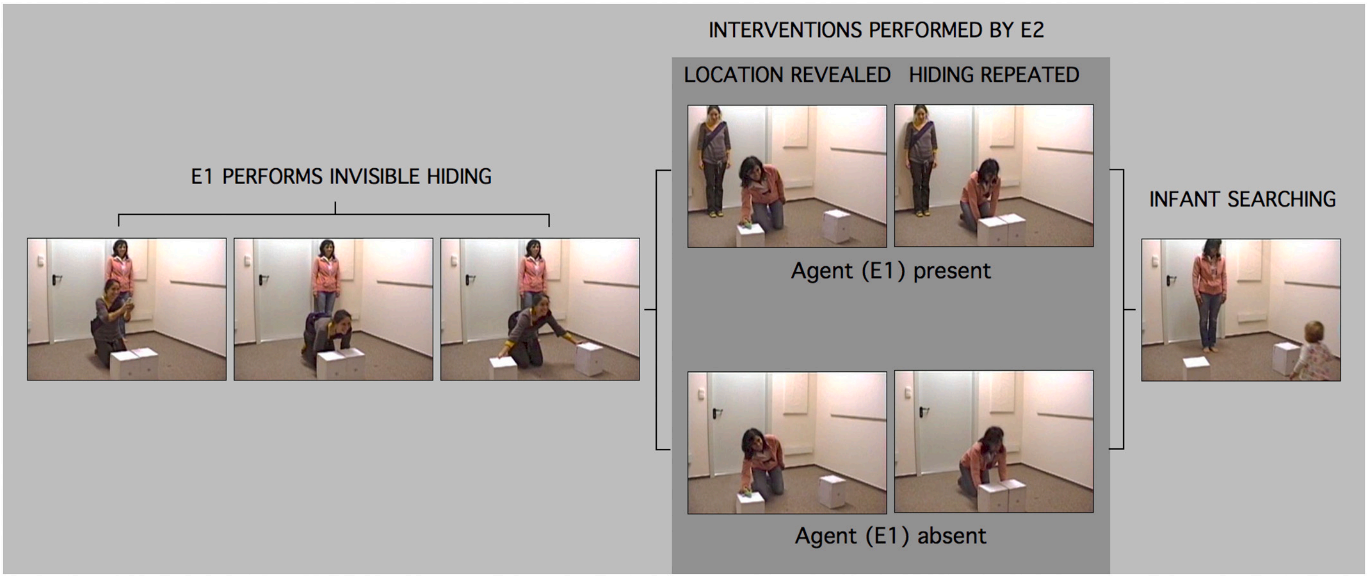

Fig. 1. Schematic illustration of the test events in Experiment 1. Experimenter 1 (E1) invisibly hides a toy, while Experimenter 2 (E2) is watching. Afterwards, in the Agent absent condition E1 leaves the scene, and E2 reveals the hiding location of the toy. Consequently, E2 performs a further invisible hiding after which infants are allowed to search. Events in the Agent present condition are similar, except that E1 witnesses the revealing and the hiding performed by E2. In the Agent absent condition infants may attribute an underspecified belief to E1 at the first hiding, update its content when the location was revealed by E2 in E1's absence, and sustain this belief, which may prime them to search for the object in this location. In contrast, in the Agent present condition infants do not have to sustain the belief of E1 that the toy is at the revealed location, as she was present at the rehiding and must have updated her belief.

[chicken/smurf]? Come and find it!," while alternating gaze between the child and the box. If the child successfully found the object in at least 3 out of 4 familiarization trials, we continued to the test phase. Otherwise the familiarization phase extended for another 2 trials. The object was hidden to the left or to the right box according to an ABBA order (starting with left or right side, counterbalanced across infants) followed by two BA trials if needed.

Two test trials, an Agent present trial and an Agent absent trial were administered with order counterbalanced across infants. At the beginning of the test trials, E2 entered the room, greeted the child, saying 'Hello, [child's name]!' and positioned herself approximately 1 m behind E1 near the door (so that she could see the hiding events, see Fig. 1). Then E1 performed the invisible hiding as in familiarization, put the two boxes apart and stood up. From this moment on the two conditions diverged.

In the Agent absent trial, E1 left the scene through the door at this point, saying 'Bye, [child's name]! I am leaving now". Then E2 revealed the location of the hidden object: she kneeled down, showed the object by placing it on the top of the respective box, and said "See? It's here!" Afterwards she placed the object at the center, reunited the boxes in the middle and performed the second invisible hiding the same way as E1 had done before. Then she stood up, stepped back, looked to the child, and said "Come and find it! Where is the [elephant/horse]? Come and find it!", without gazing or pointing to any of the boxes.

The events of the Agent present trial were identical to those in the Agent absent trial except that E1 left the scene later. After she performed the first invisible hiding, she stood up moved back 1 m near the door, watched the revealing of the object by E2 as well as the second invisible hiding, and only then left the room. Afterwards the protocol continued as in the Agent absent trial: E2 looked to the child and said "Come and find it! Where is the [elephant/horse]? Come and find it!", without gazing or pointing to any of the boxes.

After the encouragement by E2, the infants could search for the toy in both conditions. The dependent measure was the box they approached for finding the object. Note that E1 was not present during the infants' search in either of the two test conditions. In case infants did not find the object at their first choice, they were allowed to retrieve it from the other box. If they did not search in the other box, the location of the object was shown to them and the next trial began.

2.1.4. Counterbalancing during the test phase

In both test trials the object was always first hidden to the side opposite to where the object had been hidden and found in the last familiarization trial. By fixing the belief prime location (the location where E1 hid the object and believed the object to be) contralateral to the location where the object was last hidden before we aimed to act against potential carry over effects from familiarization (i.e., that infants would simply search in the location where they last found the object). The second invisible hiding, which determined the final location of the object, was always opposite to the location where the object had been placed during the first invisible hiding within the same trial. We used the same hiding location in both test trials because in the second test trial a possible carry over effect from where the object was found in the first test trial makes an opposite prediction compared to our main hypothesis (according to which infants in the Agent absent condition will search where the object was hidden at the first hiding).

2.1.5. Response coding and analyses

Searching in a particular box was defined as approaching a box within reaching distance, and touching it or looking into it. If the infant searched in the location where E1 had originally left the object (and where she must have still believed the object to be in the Agent absent condition), the behavior was coded as being primed (score 1), if they searched in the other location it was coded as unprimed (score 0).

We compared the number of infants searching at the primed location to chance level (0.5) by binomial tests, and the search frequency between the two conditions by McNemar tests (both two-tailed). In addition, we calculated binomial Bayes Factors (BF) contrasting the null hypothesis (equal probability of searching at the two locations) to the alternative hypothesis of higher probability for the primed location (using the default hyperparameters in JASP).

2.2. Results and discussion

The proportion of infants searching at the primed location is depicted in Fig. 2. In the Agent absent condition, infants searched more often in the primed location than what would have been expected by chance (18 of 24 infants, binomial test, $p = .023$). The estimated Bayes Factor suggested 9.9:1 in favor of the alternative hypothesis in this condition. In contrast, searching behavior in the Agent present condition was not different from chance (10 of 24 infants, binomial test, $p = .540$), and the Bayes Factor favored the null hypothesis (BF = 6.9). Infants' performance differed between the two conditions (McNemar test, $p = .007$, for the individual data and the contingency tables of all experiments see the Supplementary data). This suggests that the search behavior of infants in the Agent absent condition was influenced by the initial location of the object, i.e., the location where E1 should have still believed it to be. In contrast, in the Agent present condition no such effect was observed. These results cannot be explained by encoding associations or relations between the agent, the object and the location because these three elements were never simultaneously present in a given scene. In addition, E1 was not present at the moment of the search and thus could not have triggered the retrieval of possible stored relations or associations. The fact the initial location of the object did not bias the search behavior of infants in the Agent present condition shows that it was not this location itself, but its relation to E1's beliefs, which influenced the behavior of the infants.

While first order associations can be excluded on the basis of infants' performance in Experiment 1, this experiment did not clarify whether it was directly the updated belief content that biased the infants towards the primed location or a behavioral prediction that they may have generated from this belief for E1. Experiment 2 investigated this question.

3. Experiment 2

Experiment 2 was identical to Experiment 1, except that a different object was introduced for the second hiding event. Thus, after E2 revealed where E1 hid the first object, this object was exchanged to a new object, which was then invisibly hidden, and then infants were allowed to search for this second object (that E1 never encountered). If the infants in the Agent absent condition of Experiment 1 were biased towards a specific location by E1's belief content (regarding where the first object was), this bias might disappear here as infants were now searching for a different object. If, however upon updating E1's initial underspecified belief with a well defined content they have also made a prediction about where would E1 search for (or point to) in case she returns, this prediction could still bias infants towards the primed location – but only in the Agent absent condition.

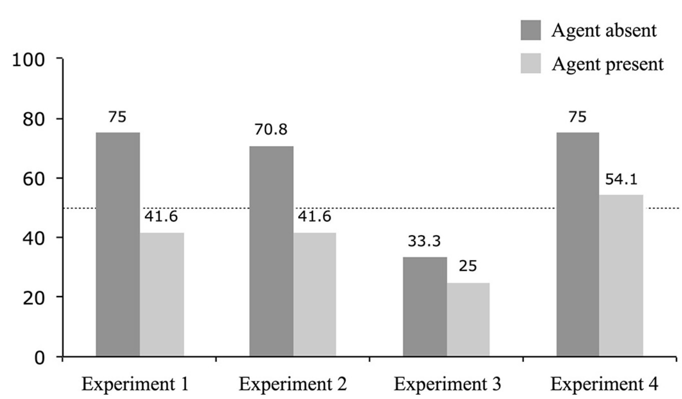

Fig. 2. Proportion of infants searching for the object in the primed location as a function of experimental conditions (Agent absent and Agent present). In Experiment 1, while searching for the toy hid by E1, infants' behavior was primed towards the location where E1 should have believed the object to be in the Agent Absent condition, which was not the case for the Agent present condition. These effects were also observed in Experiment 2, where infants were searching for a different object, and in Experiment 4 where the boxes were swapped after the location was revealed, but not in Experiment 3, where the swap was performed first and disrupted belief update. The dotted line marks chance level.

3.1. Methods

3.1.1. Participants

The participants were 24 15-month-old infants (mean age: 15;19, range: 15;00 to 15;29, 12 girls). Additional infants were excluded from the analysis because of failing to choose a box in one or both test trials ($N = 7$), parental interference ($N = 2$), fussiness and insisting keeping another object in hand during test ($N = 1$), or because they went behind the boxes to see where the object was before choosing ($N = 3$).

3.1.2. Materials

The materials were identical to Experiment 1, except that in each test trial a pair of test objects were used. In addition to the ones used in Experiment 1, a yellow plush dog and plush leopard of similar size were added.

3.1.3. Procedure

The familiarization trials were the same as in Experiment 1. The difference in the test trials was that, in both conditions, after E2 revealed where E1 hid the object, she took a new object from her bag, attracted infants' attention to it, placed it in the middle between the two boxes, and put away the initially hidden object in her bag in the full view of the infant. Afterwards she continued with the second invisible hiding as in Experiment 1, using the new object.

3.2. Results and discussion

The proportion of infants who searched at the primed location is depicted in Fig. 2. More infants searched at the primed location in the Agent absent condition (17 of 24 infants, binomial test, $p = .064$; BF = 3.8 in favor of the alternative hypothesis), but slightly less than half of the infants did so in the Agent present condition (10 of 24 infants, binomial test, $p = .540$; BF = 6.9 in favor of the null hypothesis). Infants' performance differed between the two conditions (McNemar test, $p = .015$). These results replicate those of Experiment 1 showing a belief priming effect only in the Agent absent condition, and indicate that the identity of the searched object did not seem to affect infants' search behavior.

Beyond replicating Experiment 1, the present results also make it likely that the belief that the infants attributed to E1 influenced their searching behavior not directly but via some predictions they made about her potential behavior. Alternatively, it is also possible that the attributed belief had a more direct effect if the object in it were specified at a level (e.g., 'a toy') that would not discriminate between the two hidden objects.

Either way, the difference between the conditions suggests that what mediated between the primed location and the search behavior of infants was the belief attributed to E1. However, both accounts raise a question about the representational content of the updated belief attributed to E1. How did the infants encode the hiding location of the toy, or an action prediction they drew from it? Did they specify it as the absolute location of the object (left or right) or as the box, which was at this location when E1 left?

4. Experiment 3

To investigate this issue, in Experiment 3 we introduced an extra step in the procedure used in Experiment 1. Specifically, after E1 left the scene, E2 swapped the location of the two boxes *before* revealing the location of the object. This manipulation informs infants directly about the box where E1 left the object, but the absolute location where this happened should be inferred by reversing the box-swapping event. If the belief attributed to E1 refers to absolute location, infants now should search opposite to where the object is revealed. If this belief refers to the specific box in which the object was hidden, infants should be biased towards this box. Note that both kinds of inferences demand an extra inferential step compared those that were required in Experiments 1 & 2: either the box swapping event should be reversed, or infants should adopt the extra assumption that E1 could discriminate between the boxes (which looked alike).

4.1. Methods

4.1.1. Participants

The participants were 24 15-month-old infants (mean age: 15;15, age range: 15;7 to 15;29, 12 girls). Additional infants were excluded from the analysis because of failing to choose a box in one or both test trials ($N = 8$), always choosing the same box in familiarization ($N = 2$), fussiness or crying (N = 2), or experimenter error/equipment failure (N = 2).

4.1.2. Materials and procedure

The same objects were used as in Experiment 1. The procedure was also almost identical to Experiment 1, except a manipulation that was introduced in both test trials. In each test trial, *before* E2 revealed the location of the object hidden by E1, she swapped the location of the two boxes. More precisely, she first attracted the infant's attention to herself, then slowly lifted the box on her right, brought it to the midline between the boxes, and kept it there in the air. Then with her other hand she slowly slid the box on the left to the initial location of the other box, and then placed the box in her hand to where the left box was. Importantly, E2 moved only one box at a time making sure that infants were attentive to all the events. Afterwards, she revealed the location of the toy by taking it out from the box where it was hidden.

4.2. Results and discussion

In both conditions, fewer infants searched in the primed than in the unprimed location: 8 of the 24 infants in the Agent absent condition (binomial $p = .152$; BF = 10.2 in favor of the null hypothesis) and 6 of the 24 infants in the Agent present condition (binomial $p = .023$; BF = 13.7 in favor of the null hypothesis). The patterns of search were not different in the two conditions (McNemar test $p = .625$; see Fig. 2). The infants thus tended to search in the location where object had been previously revealed, i.e., at the opposite side from the location where E1 must have hidden it, and where she must have believed it to be in the Agent absent condition. However, we find it unlikely that this tendency was due to the fact that infants filled in the box as the hiding location to the belief file, as this effect was not strong enough to be different from chance in the relevant Agent absent condition and infants produced the same, or even stronger tendency in the Agent present condition, where E1's previous belief about object location was irrelevant. We tentatively conclude that our box-swapping manipulation disrupted belief-updating processes and resulted in random behavior, which might also have been additionally biased by the last seen location of the object. However, if this disruption is related to the belief updating process, it should have an effect only if it happens *before*, but not if it happens *after* the belief is updated by the newly available information. Experiment 4 tested this possibility.

5. Experiment 4

In Experiment 4 we used the same manipulation as in Experiment 3, but at different time point in the scenario. Specifically, E2 swapped the two boxes, but did so *after* revealing where the object was hidden by E1 (and where she must have believed the object to be). As the under-specified belief computed at hiding could be updated with the revealed content before the swapping took place (and thus was irrelevant with regard to the computed belief), the swapping manipulation may not have an effect on infants' search behavior. If the belief content attributed to E1 refers to the absolute location of the toy, this procedure should replicate the belief priming effects observed in Experiments 1 and 2. If the belief attribution or the behavioral prediction is couched in terms of the hiding box, it should only require infants to track the location of this box after it has already been identified. This should not be difficult to 15-month-olds. Thus, this experiment could reveal how infants encoded the attributed location of the object without demanding extra steps of inference.

5.1. Methods

5.1.1. Participants

The participants were 24 15-month-old infants (mean age: 15;12, age range: 15;0 to 15;25, 13 girls). Additional infants were excluded from the analysis because of failing to choose a box in one or both test trials ($N = 8$), always choosing the same box in familiarization ($N = 3$), fussiness or crying ($N = 2$), or experimenter error ($N = 1$).

5.1.2. Materials and procedure

The same objects were used as in Experiments 1 & 3. The procedure was also almost identical to Experiment 3, except that in the two test trials E2 swapped the two boxes *after* revealing where E1 hid the object. Thus, after retrieving the toy, E2 placed it on the ground at the midline, then performed the box swapping procedure the same way as in Experiment 3, she reunited the boxes in the middle and performed the second invisible hiding.

5.2. Results and discussion

More infants searched at the primed the location in the Agent absent condition than what would have been expected by chance (18 of 24 infants, binomial test, $p = .023$; BF = 9.9 in favor of the alternative hypothesis), but this was not the case in the Agent present condition (13 of 24 infants, binomial test, $p = .839$; BF = 2.9 in favor of the null hypothesis; McNemar test between conditions: $p = .062$). The performance of infants in the Agent absent condition of this experiment was different from that of the corresponding condition from Experiment 3 (Fisher's exact $p = .008$). Thus, when the box-swapping manipulation was implemented after the content of belief file could be updated, infants performed as in Experiments 1 & 2. Note that we coded the 'primed location' in terms of absolute location of the first invisible hiding, so our results suggest that object locations in the attributed beliefs were encoded in absolute space (left/right) and not linked to a specific container. Whether this finding is due (i) to infants' inherent preference to encode object location in absolute, rather than context-relative terms, (ii) to tracking behavioral predictions in this way, or (iii) to the recognition that E1 would not have been able to discriminate between the boxes in case she came back, is a further question that our study cannot answer at this stage. But whenever tracking the relevant elements of others' beliefs is not disrupted by manipulation that require further retrospective inferences from infants, they seem to update under-specified beliefs that track location information with the absolute location of objects when this information becomes available.

6. General discussion

Results from four experiments point to the possibility that human infants may possess powerful representational apparatus to encode belief related information. We aimed to explore the cognitive architecture underlying infants' abilities to track other agents' beliefs and targeted characteristics that were not documented earlier. The results demonstrate that 15-month-old infants can track others' beliefs even when they have no sufficient evidence for specifying their exact contents. Indeed, it is often the case that the evidence from which one can draw inferences regarding others' beliefs is ambiguous or incomplete. Such lack of information prevents the allocation of specific content to the ascribed belief, however, this does not prevent belief attribution.

Data from the present study suggests that infants are prepared to deal with such uncertainties in belief ascription and can readily track underspecified beliefs. They possess representational structures that allow the specification of the contents of such belief files later, even in the absence of the agents to whom the beliefs have been attributed. Entertaining such attributions may then prime their behavior in specific ways (Experiments 1, 2, & 4). In Experiment 2 such an effect emerged even when the infants searched for an object that was different from the one involved in belief attribution. This suggests that the effect of the belief content attributed to another person may be mediated via the predictions infants make about the agent's potential subsequent behavior (that is, in case she came back she would search at, or (given the familiarization phase) would indicate by pointing, a specific location). Such a priming effect was not observed when belief updating was disrupted by swapping the boxes before the content of the belief could have been specified (Experiment 3) but reappeared when the specification of the content preceded the swapping of the boxes (Experiment 4). In sum, the data suggest that infants attributed underspecified beliefs to E1, updated the content of such beliefs later when relevant information became available, and sustained such representations even if their validity was unknown. Furthermore, they seemed to resort to the content of such beliefs when they had no better information, and their search was influenced by the content of these beliefs.

Across the familiarization trials of the present experiments E1 was always presented as a reliable, helpful protagonist: after each invisible hiding she pointed to the actual location of the object. Could infants rely on E1's reliability or helpfulness (rather than on her beliefs) to predict the location of the target in test?.[1] Infants are indeed sensitive to these characteristics of their social partners, however, it is unclear whether such sensitivity is sufficient to explain our findings. It is rather unlikely that during the familiarization trials infants could develop the assumption that the location where E1 hides the object in test at Time 1 will be predictive for where E2 will hide it in the future, that is, posit that E1's initial hiding will be potentially helpful for a subsequent hiding event. E1's behavior was indeed reliable and helpful within the context of single hiding events in familiarization (she signaled the location of the object after each hiding), but importantly, not across hiding events. Her hiding at Time 1 in familiarization *was not predictive* of the location of the object at a consecutive hiding event at Time 2, given the counterbalancing. Therefore, infants' potential assumptions of predictability between hiding events were not warranted by the familiarization. In addition, endorsing such assumptions would also require additional and rather complex processes, such as second-order goal attributions to E1 or E2 (e.g., E1 has the goal to hide the object, and this hiding aims at indicating the future hiding location of an object hidden by a different person; or E2 has the goal to reveal the object's earlier hiding location, which aims at signaling its location after the next hiding) and future directed thinking. Furthermore, and most importantly, such assumptions would predict identical performance in the two conditions, which is not what we found.

If infants uphold the assumption that E1 will helpfully or reliably hide in test where E2 will hide at a later time point, why would they rely on it only in the Agent absent condition, but not in the Agent present condition? One may speculate that the Agent present condition mapped onto different real-life experiences of infants than the Agent absent condition did. For instance, infants may rarely experience in real life that two people *knowingly hide objects* in the same location. Indeed, most often we do not know what task-relevant real-life experience infants bring to the lab. In any case, if infants assume that two people should *not hide knowingly* into the same location, one would expect that they reliably search in the *other location* in the Agent present condition, which is, however, not the pattern what we found. Given that we do not see how these explanations could provide an alternative account for the search patterns observed in our studies, we maintain that the most likely option is that infants attributed an underspecified belief to E1 in our experiments. What are the mechanisms underlying this feat?

The present experiments involved situations in which infants had to encode beliefs that were epistemically indeterminate. This was realized through an invisible hiding protocol that served to introduce uncertainty regarding the exact location of an object. When infants observe the hiding of an object to an unknown location, they may represent the agent's beliefs as 'She believes [the object (is in box x at location y)]', where 'x' and 'y' point to referents that are underspecified. The representations that grasp such contents should reflect the epistemic uncertainty characteristic to these cases. This can be achieved by forming underspecified ("empty") belief files (Kovács, 2016), in which placeholders stand for some aspects of the content. Such placeholders can be smoothly updated whenever relevant information becomes available. These belief contents can also be thought of grasping a set of hypotheses about feasible options, such as potential locations of an object. Sometimes, this set is not exhaustive ("the toy is either in Box 1 or Box 2 or somewhere else", where "somewhere" stands for unidentified locations that fall under the description). One crucial difference between beliefs that operate with exhaustive and open-ended disjunctions is that the former ones yield a reduced hypothesis space. Our experiments involved two boxes as possible hiding locations, and the hiding was performed such that it excluded other possibilities (also, the familiarization trials demonstrated this to the infants). Since some studies suggest that infants can exploit representations of multiple alternatives in probabilistic (Cesana-Arlotti, Téglás, & Bonatti, 2012; Téglás & Bonatti, 2016) or logical inferences (Cesana-Arlotti et al., 2018; Cesana-Arlotti, Kovács, & Téglás, 2020), it is possible that infants in our studies relied on specific disjunctive alternatives (e.g., "the toy is either in the box on the left or in the box on the right"). Whether such disjunctive relations can be integrated into infants' belief attributions is a question for further research.

Our results also suggest that representing the content of others' beliefs (and/or the behavioral predictions one can draw from those) affects infants' own searching behavior in a situation in which this is not necessarily valid information. Indeed, in our test trials infants were searching for a toy in the absence of the agent holding the relevant belief. She was not present to trigger action predictions, neither were infants required to react to her actions, which would have benefited form taking into account her beliefs. The infants' behavior was nevertheless influenced by her belief content, a phenomenon that was also documented in other studies that used looking patterns or search duration as a dependent measure (Kampis & Kovács, 2020; Kovács, Téglás, & Endress, 2010). Recent proposals have argued that infants' sensitivity to others' beliefs in different tasks may be explained by the fact that they are initially strongly altercentric (Kampis & Southgate, 2020; Southgate, 2019). Such a tendency would entail giving more weight to how others' represent the world compared to their own representations – a phenomenon that is thought to diminish as they grow older. While our study shows that attributed representations do influence infants' behavior, it does not speak to the question whether representations attributed to others would prevail when they compete with infants' own corresponding representations. A stronger test case for such accounts would be to test situations in which attributed belief contents (e.g., 'object in the left box') compete with comparable first person representations (e.g., 'object in the right box').

Furthermore, it is not only infants who show such effects, but adults are also spontaneously influenced by the beliefs and the perspective they attribute to others even in situations where they are not required to pay

[1] We would like to thank two anonymous reviewers and the editor for raising alternative explanations.

attention to others (Buttelmann & Buttelmann, 2017; Elekes, Varga, & Király, 2016; Kovács et al., 2010; Samson, Apperly, Braithwaite, Andrews, & Bodley Scott, 2010; van der Wel, Sebanz, & Knoblich, 2014). Of particular interest to the present study are the findings by Kovács et al. (2010), who have found evidence that spontaneously attributed beliefs can prime adults' reaction times and infants' looking times. For the adult data, a low level explanation was proposed by Phillips et al. (2015), who argued that the findings might be explained by an artifact arising from the timing of participants' button presses. Importantly, however, a subsequent study from an independent lab has directly excluded this low level alternative explanation (El Kaddouri, Bardi, De Bremaeker, Brass, & Wiersema, 2019). Note also that, even if the alternative explanation for the adult results were valid, it could not explain Kovács et al. (2010) results with infants. In sum, evidence from adults and infants seem to converge suggesting that humans may be particularly prone to priming effects induced by the beliefs attributed to others.

Young learners might often encounter situations when it would be useful to track other agents' relevant knowledge or beliefs, even without having exhaustive evidence about their exact content. Accounts that propose that infants would solve ToM tasks by encoding relations or associations between agents, objects and locations (Apperly & Butterfill, 2009; Perner & Ruffman, 2005) predict that representing mental states without a well-specified content should pose difficulties for infants. In other words, these accounts would predict that infants should not be able to instantiate belief files without representing their exact content. The present data provide evidence against such accounts. While human infants most often make use of their own representational apparatus involved in representing the physical word to encode others' beliefs (Kampis, Parise, Csibra, & Kovács, 2015, Southgate & Vernetti, 2014), they can engage flexible and powerful belief attribution processes. They do not just encode underspecified beliefs, but also outdated beliefs that are independent from their first person representations about the state of the world, even if they cannot judge whether those beliefs are true or false. Thus, contrary to proposals suggesting that attributed belief representations must be linked to one's first person representations that are anchored to specific referents (Perner & Leahy, 2016), such attributions may not have to be dependent or tightly linked to one's own representations.

Our results thus suggest that that infants' representation of others' mental states shares formal properties with metarepresentations. Specifically, i) the content of these states is 'shielded' from first person representations (Cosmides & Tooby, 2000), and ii) the existence and the content of these states can be manipulated independently from each other (Leslie, 1987, 1994). Infants attribute an epistemic state with underspecified content to an actor, and fill in the content of this state in the absence of the actor, via updating their primary representations. Hence, infants can handle the *existence* and the *content* of epistemic states independently, which indicates the application of an abstract metarepresentational format.

In sum, our findings point to the presence of a representational apparatus subserving online belief tracking that allows infants to operate with underspecified ('empty') belief files. While it seems that other animals can also perform some kind of belief tracking based on specific contingencies between agents, objects, and locations (Krupenye et al., 2016), encoding belief files that rely on placeholders and have a metarepresentational format may be beyond their abilities. Future studies should target the question whether these abilities are specific to humans.

Acknowledgements

We thank parents and children participating in this study and B. Köllöd-Széplaki, M. Tóth and Á. Volein for their help with data collection. We would like to thank to Jacques Mehler for his continuous inspiration and friendship over many years, the extensive insightful discussions and for his invaluable support in promoting big ideas. This study was partly supported by a European Union's Seventh Framework Programme (FP7/2007-2013) ERC Starting Grant #284236 (REPCOLLAB), by a European Union's Horizon 2020 Research and Innovation Programme ERC Starting Grant #639840 (PreLog) and by James S McDonnell Foundation, Grant #220020449.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.cognition.2021.104640.

References

Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review, 116*(4), 953–970.
Baillargeon, R., Buttelmann, D., & Southgate, V. (2018). Invited commentary: Interpreting failed replications of early false-belief findings: Methodological and theoretical considerations. *Cognitive Development, 46*, 112–124. https://doi.org/10.1016/j.cogdev.2018.06.001.
Barone, P., & Gomila, A. (2020). Infants' performance in the indirect false belief tasks: A second-person interpretation. *Wiley Interdisciplinary Reviews: Cognitive Science*. https://doi.org/10.1002/wcs.1551. Dec 14:e1551. (Epub ahead of print. PMID: 33319503).
Bloom, P., & German, T. P. (2000). Two reasons to abandon the false belief task as a test of theory of mind. *Cognition, 77*, B25–B31.
Buttelmann, F., & Buttelmann, D. (2017). The influence of a bystander agent's beliefs on children's and adults' decision-making process. *Journal of Experimental Child Psychology, 153*, 126–139.
Buttelmann, F., & Kovács, Á. M. (2019). 14-month-olds anticipate others' actions based on their belief about an object's identity. *Infancy, 24*(5), 738–751. https://doi.org/10.1111/infa.12303.
Buttelmann, D., Carpenter, M., & Tomasello, M. (2009). Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition, 112*, 337–342.
Call, J., & Tomasello, M. (1999). A nonverbal false belief task: The performance of children and great apes. *Child Development, 70*, 381–395.
Carruthers, P. (2013). Mindreading in infancy. *Mind & Language, 28*, 141–172. https://doi.org/10.1111/mila.12014.
Cesana-Arlotti, N., Kovács, Á. M., & Téglás, E. (2020). Infants recruit logic to learn about the social world. *Nature Communications, 11*, 5999. https://doi.org/10.1038/s41467-020-19734-5.
Cesana-Arlotti, N., Martín, A., Téglás, E., Vorobyova, L., Cetnarski, R., & Bonatti, L. L. (2018). Precursors of logical reasoning in preverbal human infants. *Science, 359*, 1263–1266.
Cesana-Arlotti, N., Téglás, E., & Bonatti, L. L. (2012). The probable and the possible at 12 months: Intuitive reasoning about the uncertain future. In J. B. Benson (Serial Ed.), F. Xu, & T. Kushnir (Vol. Eds.) (Eds.), *Rational constructivism in cognitive development* (pp. 1–25). Academic Press: Elsevier Inc.
Cosmides, L., & Tooby, J. (2000). Consider the source: The evolution of adaptations for decoupling and Metarepresenations. In D. Sperber (Ed.), *Metarepresentations: A Multidisciplinary Perspective*. Oxford University Press.
El Kaddouri, R., Bardi, L., De Bremaeker, D., Brass, M., & Wiersema, J. R. (2019). Measuring spontaneous mentalizing with a ball detection task: Putting the attention-check hypothesis by Phillips and colleagues (2015) to the test. *Psychological Research*, 1–9.
Elekes, F., Varga, M., & Király, I. (2016). Evidence for spontaneous level-2 perspective taking in adults. *Consciousness and Cognition, 41*, 93–103.
Flombaum, J. I., & Santos, L. R. (2005). Rhesus monkeys attribute perceptions to others. *Current Biology, 15*, 447–452.
Hare, B., Call, J., & Tomasello, M. (2001). Do chimpanzees know what conspecifics know? *Animal Behaviour, 61*, 139–151.
Hayashi, T., Akikawa, R., Kawasaki, K., Egawa, J., Minamimoto, T., Kobayashi, K., … Hasegawa, I. (2020). Macaques exhibit implicit gaze bias anticipating others' false-belief-driven actions via medial prefrontal cortex. *Cell Reports, 30*(13), 4433–4444. e5.
Kaminski, J., Call, J., & Tomasello, M. (2008). Chimpanzees know what others know, but not what they believe. *Cognition, 109*(2), 224–234.
Kano, F., Krupenye, C., Hirata, S., Tomonaga, M., & Call, J. (2019). Great apes use self-experience to anticipate an agent's action in a false-belief test. *Proceedings of the National Academy of Sciences of the United States of America, 116*(42), 20904–20909. https://doi.org/10.1073/pnas.1910095116.
Kampis, D., Karman, P., Csibra, G., Southgate, V., & Hernik, M. (2020). *A two-lab direct replication attempt of Southgate, Senju, & Csibra (2007)*. PsyArXiv Preprints. https://doi.org/10.31234/osf.io/gzy26.
Kampis, D., & Kovács, Á. M. (2020). Seeing the world from others' perspective: 14-month-olds show altercentric modulation effects by others' beliefs. *PsyArXiv Preprints*. https://doi.org/10.31234/osf.io/an7h3.
Kampis, D., Parise, E., Csibra, G., & Kovács, Á. M. (2015). Neural signatures for sustaining object representations attributed to others in preverbal human infants. *Proceedings of the Royal Society B: Biological Sciences, 282*(1819). https://doi.org/10.1098/rspb.2015.1683.

Kampis, D., & Southgate, V. (2020). Altercentric Cognition: How Others Influence Our Cognitive Processing. *Trends in Cognitive Sciences, 24*(11), 945–959. https://doi.org/10.1016/j.tics.2020.09.003.

Király, I., Oláh, K., Csibra, G., & Kovács, Á. M. (2018). Retrospective attribution of false beliefs in 3-year-old children. *Proceedings of the National Academy of Sciences, 115* (45), 11477–11482.

Knudsen, B., & Liszkowski, U. (2012). 18-month-olds predict specific action mistakes through attribution of false belief, not ignorance, and intervene accordingly. *Infancy, 17*, 672–691.

Kovács, Á. M. (2016). Belief files in theory of mind reasoning. *Review of Philosophy and Psychology, 7*(2), 509–527.

Kovács, Á. M., Téglás, E., & Endress, A. D. (2010). The social sense: Susceptibility to others' beliefs in human infants and adults. *Science, 330*(6012), 1830–1834.

Krupenye, C., Kano, F., Hirata, S., Call, J., & Tomasello, M. (2016). Great apes anticipate that other individuals will act according to false beliefs. *Science, 354*(6308), 110–114.

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

Leslie, A. (1987). Pretense and representation: The origins of "theory of mind". *Psychological Review, 94*(4), 412–426.

Leslie, A. M. (1994). *Pretending* and *believing:* Issues in the theory of ToMM. *Cognition, 50*, 211–238.

Martin, A., & Santos, L. R. (2014). The origins of belief representation: Monkeys fail to automatically represent others' beliefs. *Cognition, 130*, 300–308.

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

Perner, J., & Leahy, B. (2016). Mental files in development: Dual naming, false belief, identity and Intensionality. *Review of Philosophy and Psychology, 7*, 491–508.

Perner, J., & Ruffman, T. (2005). Infant's insight into the mind: How deep? *Science, 308*, 214–216. https://doi.org/10.1126/science.1111656.

Phillips, J., Ong, D. C., Surtees, A. D. R., Xin, Y., Williams, S., Saxe, R., & Frank, M. C. (2015). A second look at automatic theory of mind: Reconsidering Kovacs, Teglas, and Endress (2010). *Psychological Science*, 1–15. https://doi.org/10.1177/0956797614558717.

Poulin-Dubois, D., Rakoczy, H., Burnside, K., Crivello, C., Dörrenberg, S., Edwards, K., … Perner, J. (2018). Do infants understand false beliefs? We don't know yet–a commentary on Baillargeon, Buttelmann and Southgate's commentary. *Cognitive Development, 48*, 302–315.

Samson, D., Apperly, I. A., Braithwaite, J. J., Andrews, B. J., & Bodley Scott, S. E. (2010). Seeing it their way: Evidence for rapid and involuntary computation of what other people see. *Journal of Experimental Psychology: Human Perception and Performance, 36* (5), 1255–1266.

Scott, R. M., & Baillargeon, R. (2017). Early false-belief understanding. *Trends in Cognitive Sciences, 21*, 237–249. https://doi.org/10.1016/j.tics.2017.01.012.

Scott, R. M., Richman, J. C., & Baillargeon, R. (2015). Infants understand deceptive intentions to implant false beliefs about identity: New evidence for early mentalistic reasoning. *Cognitive Psychology, 82*, 32–56. https://doi.org/10.1016/j.cogpsych.2015.08.003.

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

Southgate, V. (2019). Are infants altercentric? The other and the self in early social cognition. *Psychological Review*. https://doi.org/10.1037/rev0000182.

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

Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science, 18*(7), 587–592. https://doi.org/10.1111/j.1467-9280.2007.01944.x.

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

Tauzin, T., & Gergely, G. (2018). Communicative mind-reading in preverbal infants. *Scientific Reports, 8*, 9534. https://doi.org/10.1038/s41598-018-27804-4.

Téglás, E., & Bonatti, L. L. (2016). Infants anticipate probabilistic but not deterministic outcomes. *Cognition, 157*, 227–236.

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. https://doi.org/10.1111/j.2044-835X.2011.02067.

Träuble, B., Marinović, V., & Pauen, S. (2010). Early theory of mind competencies: Do infants understand others' beliefs? *Infancy, 15*, 434–444. https://doi.org/10.1111/j.1532-7078.2009.00025.x.

Velleman, J. D. (2000). On the aim of belief. In *The possibility of practical reason* (pp. 244–281).

van der Wel, R. P., Sebanz, N., & Knoblich, G. (2014). Do people automatically track others' beliefs? Evidence from a continuous measure. *Cognition, 130*(1), 128–133.