Are there signature limits in early theory of mind?

Abstract

Current theory of mind research faces the challenge of reconciling two sets of seemingly incompatible findings: While children come to solve explicit verbal false belief (FB) tasks from around age 4, recent studies with various less explicit measures such as looking time, anticipatory looking and spontaneous behavior suggest that even infants can succeed on some FB tasks. In response to this tension, two-systems theories propose to distinguish between an early-developing system, tracking simple forms of mental states, and a later-developing system, based on fully-developed concepts of belief and other propositional attitudes. One prediction of such theories is that the early-developing system has signature limits concerning aspectuality. We tested this prediction in two experiments. The first experiment showed (in line with previous findings) that 2- to 3-year-olds take into account a protagonist's true or false belief about the location of an object in their active helping behavior. In contrast, toddlers' helping behavior did not differentiate between true and false belief conditions when the protagonist's belief essentially involved aspectuality. Experiment 2 replicated these findings with a more stringent method designed to rule out more parsimonious explanations. Taken together, the present findings are compatible with the possibility that early theory of mind reasoning is subject to signature limits as predicted by the two-systems account.

Keywords: theory of mind, two-systems theory, signature limits

Are there signature limits in early theory of mind?

Current Theory of Mind (ToM) research faces what has been termed the "puzzle of belief reasoning" (Helming, Strickland, & Jacob, 2014). On the one hand, decades of research with a variety of mostly verbal false belief (FB) tasks suggests that children acquire the competence to ascribe beliefs around age 4 (for review, see the meta-analysis of Wellman, Cross, & Watson, 2001). Numerous findings indicate that the emergence of this competence is not a local phenomenon, affecting performance on one or two isolated tasks. Instead the competence reveals itself in systematically inter-correlated ways on a range of superficially diverse tasks where the common denominator of these tasks is merely that they require an understanding of representation (for review, see Perner & Roessler, 2012). This development thus seems to be best described as a conceptual revolution: it is the novel acquisition of a comprehensive and unitary cognitive capacity.

But a rapidly growing body of new evidence suggests that infants and toddlers perform competently on implicit FB tasks well before age 4 (for reviews, see Baillargeon, Scott, & Bian, 2016; Baillargeon, Scott, & He, 2010; Carruthers, 2013). Violation-of-expectation tasks have found that infants look longer at events in which a protagonist acts in a way that does not fit her beliefs (e.g., Onishi & Baillargeon, 2005; Surian, Caldi, & Sperber, 2007). Anticipatory looking studies have shown that, from the second year of life or earlier, children, just like adults, look in anticipation to where an actor is going to act based on her beliefs (Clements & Perner, 1994; Rubio-Fernández, 2013; Schneider, Bayliss, Becker, & Dux, 2012; Southgate, Senju, & Csibra, 2007). Studies with interactive measures have shown that infants and toddlers can spontaneously help and inform others in ways that are sensitive to the recipient's beliefs (D. Buttelmann, Carpenter, & Tomasello, 2009; D.

Buttelmann, Over, Carpenter, & Tomasello, 2014; Knudsen & Liszkowski, 2012; Southgate, Chevallier, & Csibra, 2010).

Three competing accounts of early implicit ToM findings

How can these two seemingly incompatible sets of findings be reconciled? Three main theoretical responses to this puzzle of belief reasoning are currently under discussion. Late competence accounts claim that proper ToM capacities are only required for solving explicit tasks, whereas the new implicit tasks using looking time and interaction measures reflect much simpler cognitive capacities (Heyes, 2014; Ruffman & Perner, 2005; Sirois & Jackson, 2007). According to such accounts, many of the looking time studies can be explained by low-level processes such as a novelty preference (Heyes, 2014) or the use of simple behavior rules (Ruffman & Perner, 2005).

Early competence accounts argue the converse. According to these accounts, implicit tasks are the true indicator of ToM capacities. Younger children's failures on explicit FB tasks do not reflect a deficit in ToM but merely extraneous demands imposed by these tasks. These demands are extraneous in the sense that they have nothing to do with ToM per se, but only with linguistic and other aspects of the explicit task structure (Baillargeon et al., 2010; Carruthers, 2013; Leslie, 2005).

Two-systems accounts oppose both late and early competence accounts. Instead, they claim, implicit tasks do tap ToM abilities of some kind, but these precocious abilities are distinct from the later-developing conceptual capacities measured in explicit tasks (e.g. (Apperly & Butterfill, 2009; Low, Apperly, Butterfill, & Rakoczy, 2016; Perner & Roessler, 2012; Rakoczy, 2012). On such views, younger children's failures on explicit FB tasks is not merely a consequence of extraneous demands but reflects a true conceptual deficit.

On a particularly promising two-systems account (Apperly & Butterfill, 2009; Butterfill & Apperly, 2013; Low et al., 2016), there are at least two systems for tracking beliefs and other mental states, which we shall label S1 and S2. Relative to S2, system S1 trades flexibility for gains in efficiency by relying on a simpler model of mental states. S1 is therefore limited in ways that S2 is not. For our purposes, the crucial limit is that S1 does not enable tracking false beliefs essentially involving *aspectuality*. To illustrate aspectuality, consider a popular film. Lois Lane is yet to discover that Clark Kent is Superman. She simultaneously believes that Superman is with her and that Clark Kent is elsewhere. She has incompatible beliefs about one and the same person under two different aspects. Only S2 is capable of tracking Lois' beliefs, which essentially involve aspectuality. Suppose that success on implicit FB tasks is a consequence of S1 only, whereas success on explicit FB tasks requires S2. In that case, infants' performance should exhibit the limits of S1. They may succeed on many implicit FB tasks which do not essentially involve aspectuality, such as the many tasks which involve simple mistakes about location only. But where FB tasks essentially involve aspectuality, infants should not succeed. Or so this two-systems account implies.

Which of the three accounts is right?

Late competence accounts become less and less plausible as more and more findings from various implicit tasks converge. But the empirical situation so far remains inconclusive between early competence and two-systems accounts. This is because most existing research has considered tasks for which the early competence and two-systems accounts make the same predictions. To illustrate, consider change-of-location FB tasks: the

protagonist is present when an object O is put into container 1, absent when it is then transferred to container 2, and the question is where subjects will expect the protagonist finally to search for O. Early competence accounts claim that infants operate with fully-fledged propositional attitude concepts and should thus succeed on such FB tasks. The two-systems account claims that infants operate with relational attitudes only. This is sufficient for tracking false beliefs about location and infants should succeed on such FB tasks. To distinguish early competence from two-systems accounts we must therefore look to other tests.

The two-systems account makes clear and testable predictions concerning implicit and explicit FB tasks. Concerning explicit tasks, the two-systems account predicts that whether or not a task requires dealing with aspectuality should make little difference: competence in different FB tasks should emerge in synchrony, and performance should be highly inter-correlated. This prediction is shared, naturally, with early competence accounts. But early competence accounts predict the same pattern of unity and convergence for implicit tasks as well. By contrast, the two-systems account predicts disunity and dissociation. Infants' (and others') performance on implicit tasks should depend on whether the task requires dealing with aspectuality. Young children may master change-of-location and related implicit FB tasks, but should not succeed on tasks that necessitate grasp of aspectuality.

The empirical state of the art

What is the state of the existing evidence vis-à-vis these predictions? Concerning unity of performance in various explicit FB tasks in older children, earlier studies directly

investigating young children's understanding of aspectuality found evidence for disunity and dissociation: children mastered tasks of aspectuality understanding only some years after passing standard FB tasks (Apperly & Robinson, 1998; Kamawar & Olson, 1999, 2011; Russell, 1987; Sprung, Perner, & Mitchell, 2007). But more recent research suggests that these findings masked children's competence due to extraneous task (Rakoczy, 2017; Rakoczy, Fizke, Bergfeld, & Schwarz, 2015). These new studies have developed simplified aspectual tasks. For example, one used a bunny that became a carrot when turned inside out. This object was put, as a carrot, into box 1 in the presence of the protagonist. In the second step, the protagonist was absent while the object was taken out of the box, turned inside out so that it became a bunny, and put back into Box 1 as a bunny. In the final step, the protagonist returns and observes the object being moved, as a bunny, from Box 1 to Box 2. The test question was, "Where will the protagonist look for the carrot?" To answer this question correctly ("in Box 1"), the child has to take into account aspectuality. This is because the protagonist, being unaware that the bunny is the carrot, believes that the bunny is in Box 2 and that the carrot is in Box 1. Four-year-olds succeeded on this aspectuality false belief task. Further, there were strong correlations between aspectuality and standard FB tasks (Rakoczy et al., 2015). This recent evidence thus speaks strongly in favor of convergence in performance on explicit FB tasks and unity in the underlying competence.

What about unity of performance in implicit FB tasks? A number of recent studies do initially appear to suggest that infants and toddlers can solve some implicit FB tasks requiring an understanding of aspectuality (F. Buttelmann, Suhrke, & Buttelmann, 2015; Scott & Baillargeon, 2009; Scott, Richman, & Baillargeon, 2015). However, all of these studies have used a single isolated situation or vignette, each of which leaves room for

alternative, more parsimonious explanations.[1] Even more fundamentally, it hard to interpret these findings given that the scenarios they rely on have neither been used nor validated in explicit tasks with older children or adults. (For preliminary evidence that adults in fact do not view the scenarios of Scott & Baillargeon (2009) as the task analysis assumes they should, see Low et al., unpublished data.) It thus remains unclear whether these implicit tasks really do tap the same representational processes as explicit tasks. Such doubt receives support from recent findings that performance on different types of implicit ToM tasks does not show the unity and convergence characteristic of performance on explicit tasks (Yott & Poulin-Dubois, 2016).

A more direct test of two-systems accounts therefore requires direct comparison of implicit and explicit tasks, ideally using a single scenario. Low and colleagues have done just this in a series of studies. For the explicit tasks, their results indicate that competence in standard change-of-location scenarios converges with competence in modified scenarios that involve aspectuality (Low, Drummond, Walmsley, & Wang, 2014; Low & Watts, 2013): 3-year-olds consistently failed both non-aspectual and aspectual versions of explicit FB tasks, whereas 4-year-olds and adults consistently mastered both types of task. By contrast, in implicit versions of the tasks, all age groups showed signs of tracking the agent's belief in the standard (non-aspectual) FB task, but not in the aspectual FB task. This suggests there are signature limits in early implicit ToM, as predicted by the two-systems account.

This interpretation of the findings in terms of signature limits has recently been subject to some debate, however (Carruthers, 2013; Csibra, 2012; Jacob, 2012). Defenders

[1] For arguments that the tasks used by Scott & Baillargeon (2009) and Scott, Richman & Baillargeon (2015) do not clearly require understanding aspectuality, see Butterfill & Apperly (2013) and Low, Apperly, Butterfill & Rakoczy (2016). A specific worry concerning Buttelmann et al. (2015) is that it is not about children's understanding of beliefs regarding aspectuality at all, but rather about their understanding of reality vs. appearance, and it is not clear whether such tasks require an understanding of aspectuality at all. In addition, the effects in this study were small and basically restricted to one out of four sets of stimuli.

of early competence accounts have worried whether the different results in the aspectual FB and the standard FB tasks may have been due to different working memory demands or other performance factors (perhaps because only the aspectual FB task involves certain forms of mental rotation and simulation). Furthermore, these findings of disunity are so far restricted to looking behavior. It is thus unclear whether such performance patterns also translate into more active behavior.

Rationale of the present study

In the present study we intended to test whether there are signature limits in early ToM capacities. Going beyond mere looking behavior, we investigated children's spontaneous interactive behavior towards a protagonist. To directly investigate whether disunity in early ToM performance in various implicit tasks contrasts with unity in later explicit performance, we used the very same kind of materials and scenarios with which recent evidence for convergence and unity in explicit FB tasks was found (Rakoczy et al., 2015). In our non-aspectual, change-of-location conditions, children saw scenarios in which a protagonist had a true belief (TB), or false belief (FB), about an object's location (following D. Buttelmann et al., 2009). In our novel, aspectual conditions, the protagonist did (TB), or did not (FB), know about the dual aspects of an object. The crucial difference between the non-aspectual and aspectual conditions of our helping task is thus whether they require tracking beliefs essentially involving aspectuality. Accordingly, the distinctive predictions of the two-systems account are, first, that children's spontaneous interactive responses should differ appropriately between the TB and FB non-aspectual conditions; and, second, that there should be no such difference between TB and FB aspectual conditions.

Experiment 1

Method[2]

Participants. 67 children (26 females) were included in the final sample. The children's mean age was 31 months (*SD* = 2.66; *range* = 26-36). 9 additional children were not included in the analysis because they were uncooperative (*N* = 8) or due to experimental error (*N* = 1).

Design. Children were randomly assigned to one of four conditions: non-aspectual, change-of-location false belief (Loc_FB), non-aspectual, change-of-location true belief (Loc_TB), aspectual false belief (Asp_FB) and aspectual true belief (Asp_TB). Each child received two (in two exceptional cases three, as explained in the *Supplementary Materials*) trials in one of the conditions.

Materials. In the non-aspectual task, one toy object (a ball, a soft toy rabbit, a toy dog or toy donkey) per trial and two boxes were used. In the aspectual task, one reversible soft toy (a bear, bunny, tiger, pig), which could be turned inside out through a zipper and transformed (into a honey barrel, a carrot, a snail and a cake, respectively, see Figure 1), and one box were used per trial. Boxes were covered with a tissue on the inside to leave children ignorant about their (empty) contents. A puppet called Susi (animated by E2) was used as protagonist.

[2] For full details concerning the method of both studies, and for complementary results and analyses, see the Supplementary Material.

*Figure 1.*Examples of the stimuli used in the aspectual false belief task.

Procedure. Each child was tested in an interactive play setting with two experimenters. A parent was present during the session. After two warm-up trials (see *Supplementary Material* for details), children standardly received two trials of the condition they were randomly assigned to. In exceptional cases, a third trial was administered.

(1) Non-aspectual, change-of-location tasks. The basic sequence of events in these tasks, modeled on Buttelmann et al. (2009), was the following: There were two boxes. These were difficult to open, and the child learned how to open them. The protagonist, Susi, received a toy and expressed a liking for it. Susi then announced that she had forgotten something and had to leave, and that she would leave the toy in box 1 during her absence. She did so with the help of E1 (because she was unable to open the box on her own). The following sequence varied between conditions (see Fig. 2 and *Supplementary Material)*:

- Loc_FB condition: Susi left. In Susi's absence, E1 proposed playing a trick on Susi. E1 took the toy out of the box and sneakily placed it in the other box.
- Loc_TB condition: Susi left, and returned a few seconds later. In Susi's presence, E1 took the toy out of the box and placed it in the other box, saying: "Look, Susi and [child's name]!" and alternating gaze between the child and E2. Susi observed the location change closely, saying: "Ah, I see! Yes!"

After her return (in the Loc_FB condition), or after the observed location change (in the Loc_TB condition), Susi tried to open the box where her toy was formerly located. Susi failed to open the box, and showed disappointment. We call this event the *signal* in what follows. If the child did not react spontaneously, Susi asked "Can you help me?". If the child still showed no reaction, E1 prompted the child to help. If the child still refused, E1 asked "Should we help her together?", but didn't move and waited to see where the child would go. Children's reactions to the signal were coded. The task was repeated with different boxes and a different target toy.

(2) Aspectual task. Children were first familiarized with the dual aspectuality of the target objects. To this end, Susi (the protagonist) left the room and E1 showed the first soft toy to the child. Then she sneakily transformed it to its other aspect saying: "Look! The bunny is also a carrot!" (for example). E1 reminded the child, "Susi does not know that, right?". The child was then asked, "Can you make it so that it is a bunny again?". If the child was unable or unwilling, E1 helped. This was repeated with a second object, which was later used for the second test trial.

Following familiarization, Susi returned for the test trials. As in non-aspectual conditions, there were two boxes. These were difficult to open, but the child learned how to open them. Susi received the first soft toy and expressed her liking of it. Susi then announced that she had forgotten something, and that she would leave the toy in box 1 during her absence. She did so with the help of E1 (because she was unable to open the box on her own). The following sequence varied between conditions (see Fig. 2 and *Supplementary Material)*:

- Asp_FB condition: Susi left. In Susi's absence, E1 proposed playing a trick on Susi. E1 took the toy out of box 1. E1 transformed it (e.g. a carrot) while giggling, gesturing and whispering "Shh". E1 then returned it into the box.
- Asp_TB condition: Susi left, and returned a few seconds later. In Susi's presence, E1 took the toy out of box 1 and transformed it while saying: "Look Susi and [child's name]!" and alternating gaze between the child and E2. E1 then put it back into the box. Susi observed this event closely, saying: "Ah, I see! Yes!"

After her return (in the Asp_FB condition), or after the observed aspect change (in the Asp_TB condition), Susi moved back to her original position. E1 said, "Look Susi and [child's name]!" before taking the transformed toy (e.g. carrot) out of the box and placing it approximately 1m away from box 1 at a point equidistant between Susi and the child. After observing this closely, Susi then tried to open box 1. Susi failed to open the box, and showed disappointment. This signal and the following events were the same as in the non-aspectual, change-of-location tasks. The task was repeated with a different box and a different reversible toy.

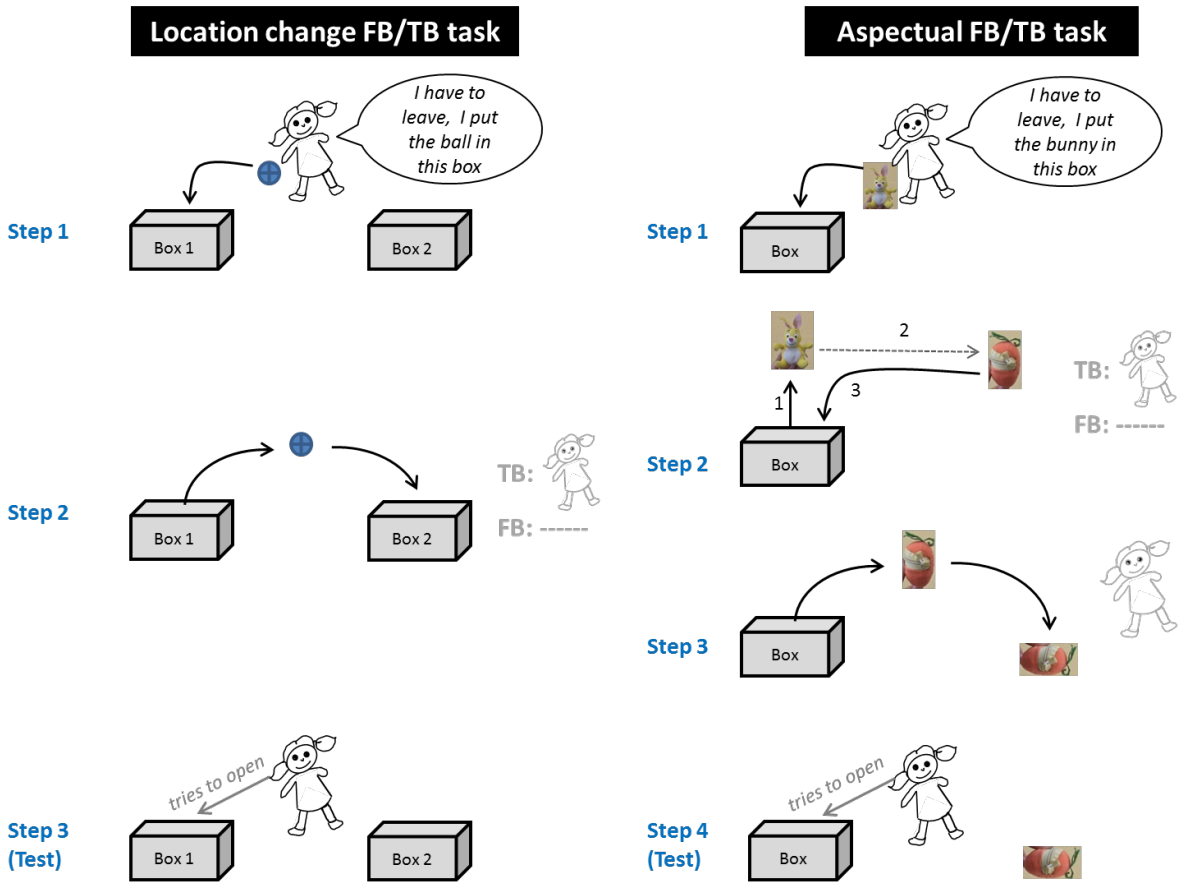


Figure 2. Schematic event sequences of the non-aspectual and aspectual tasks in Exp. 1.

Coding. Sessions were coded from video by one coder. The child's first reaction to the signal (her failed attempt to open a box) was coded. The coder categorized this reaction as:

- open box 1: the child clearly referred to box 1, either by approaching the box and opening it successfully, or else by trying unsuccessfully to open it (e.g. pulling the elastic band).
- open box 2 (in the non-aspectual task only): the child clearly referred to box 2, either by approaching the box and opening it successfully, or else by trying unsuccessfully to open it
- give the object to Susi (in the aspectual task only): the child handed the toy Susi.

- ambiguous: the child showed a behavior that was clearly a reaction to the signal, but which did not fit any the above categories.
- invalid: a parent interfered; the child did not show any reaction at all; the child left the scene during the trial or did not pay attention to the event sequence; or the child could not be held back from reacting too early.

A second coder, who was blind for the experimental condition and hypotheses, coded 14 randomly selected videos (a 20% sample). The second coder agreed with the original coder on all trials ($\kappa = 1.00$).

Predictions. In the Loc_FB condition, a child tracking Susi's false belief may recognize that Susi is opening box 1 because she believes, falsely, that her toy is in box 1. This would give the child a reason to respond by opening box 2, which actually contains Susi's toy. The child has no comparable reason for opening box 2 in the Loc_TB condition. Accordingly, children who track Susi's false belief should open box 1 more often in Loc_TB than in Loc_FB. (Of course, it is not wrong for the child to respond by opening box 1 in Loc_FB. After all, the child may reason that the best way to help Susi is to assist with her proximal goal, which is to open box 1.) Similar reasoning suggests that children who track Susi's belief in the aspectual conditions should open box 1 more often in Asp_TB than in Asp_FB.

Recall that the two-systems account predicts that two-year-olds can track some false beliefs about location but cannot track false beliefs essentially involving aspectuality. The pattern of results predicted by the two-systems account is therefore the following: children should open box 1 more often in Loc_TB than Loc_FB, but should not open box 1 more often in Asp_TB than Asp_FB.

Results

All in all, 64 children (32 each in the non-aspectual and aspectual conditions) received at least one valid (including ambiguous) trial. For each of these children, at least one trial was also unambiguous. (For details concerning more comprehensive analyses involving invalid and ambiguous trials, see *Supplementary Material*.) For the main analysis, the first valid and unambiguous trial was used. When possible, one-tailed tests were conducted whenever directed a priori hypotheses were tested (such that children perform more often "open box 1" in contrast to "open box 2/give object" behavior in the TB than in the FB conditions). Figure 3 depicts children's helping behavior in the first valid and unambiguous trial. In non-aspectual conditions, children's helping behavior differed significantly between Loc_FB and Loc_TB conditions (p = .037, one-tailed Fisher's exact test). In the aspectual tasks, in contrast, helping behavior did not differ significantly between Asp_FB and Asp_TB conditions (p = .166, one-tailed Fisher's exact test).

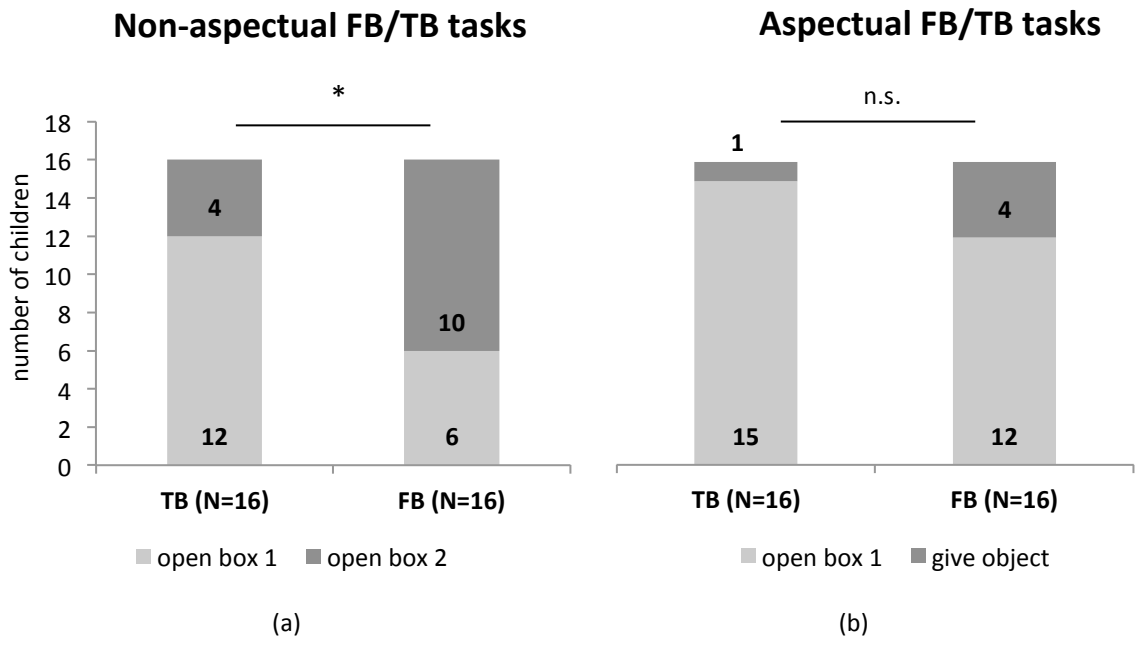


Figure 3. Number of children showing the different kinds of helping behavior in the first valid and unambiguous trial of (a) the non-aspectual tasks, (b) the aspectual tasks. [$^{*}p < .05$]

Complementary analysis. When comparing the non-aspectual and aspectual tasks for FB and TB, respectively, results revealed that children tended to perform differently in the non-aspectual than in the aspectual FB tasks (opening box 1 less often in Loc_FB than in Asp_FB, p = .037, one-tailed Fisher's exact test). No differences were found for the two TB tasks, Loc_TB and Asp_TB (p = .166).

Discussion

The results of the non-aspectual task in the present experiment replicate the general finding by Buttelmann et al. (2009) in the sense that toddlers responded differentially in Loc_FB and Loc_TB tasks. Performance differed on the novel, aspectual tasks which we developed to test the predictions of the two-systems theory. Children's helping behavior did not differ significantly between Asp_TB and Asp_FB conditions.

What does this pattern show? One possibility is that our results do indeed reflect the characteristic signature limits in children's early ToM abilities predicted by the two-systems account (Butterfill & Apperly, 2013). We will turn to that explanation, and to alternatives, in the General Discussion. Another possibility, however, is that the difference in performance reflects the influence of different performance factors and tasks demands. Our aspectual task might have imposed higher working memory demands since children had to keep in mind the two aspects of an object and Susi's visual access these. By contrast, the non-aspectual task involved just one object and just one aspect. In light of this extraneous difference, it may be argued that our novel aspectual tasks simply failed to uncover children's existing competence. This objection is particularly pressing given our relatively small sample size and thus power. A further potential objection is that we gave children a

warm-up trial in which Susi went from being knowledgeable to ignorant concerning the location of a toy (as described in the *Supplementary Material*). This might have facilitated performance in the non-aspectual task change task only.

Differences between our aspectual and non-aspectual tasks thus reveal the need for replication of our findings. The replication should use a more stringent, minimal contrast design in which all but the crucial aspects (the content of the belief to be ascribed, and whether it is true or false) are kept as constant as possible across tasks. Experiment 2 therefore investigated possible signature limits in early belief understanding with thoroughly matched non-aspectual and aspectual tasks, and with larger sample size intended to ensure sufficient power.

Experiment 2

In order to match the aspectuality and non-aspectual tasks more closely, new versions of these tasks with comparable performance demands were devised. The warm-up trials were modified to minimize the risk of priming children for a change-of-location false belief task. To equate working memory load as far as possible, the new version of the non-aspectual task involved two objects rather than one. To reduce the number of trials coded "invalid" and "ambiguous", the experimental setting was changed. Children sat on the lap of a parent at the table. A younger age group was investigated, which we hoped would mean fewer children would do things like run away or turn around during testing.

Method

Participants. 137 children (61 females) were included in the final sample[3]. Children's mean age was 26 months (*SD* = 1.92; *range* = 24-30). Five additional children were tested but not included in the analysis because they were uncooperative.

Design. As in Exp. 1, children were randomly assigned to one of four conditions: non-aspectual, change-of-location false belief (Loc_FB), non-aspectual, change-of-location true belief (Loc_TB), aspectual false belief (Asp_FB) and aspectual true belief (Asp_TB). Each child received two (in two exceptional cases three, as explained in the *Supplementary Materials*) trials in one of the conditions. (Two short additional tasks were administered during warm-up and at the end of the session. As these focused on another research question, they are not reported here.)

Materials. We used materials similar to those used in Exp. 1. For Exp. 2, we used a larger box was with more openings to allow more ways of searching in it (see Fig. 4). Also, in Exp. 2 we did not use the puppet, Susi. Instead E2 herself played the role of the protagonist. This because we were concerned that Susi might frighten younger children.

Figure 4. Example of the boxes used in Exp. 2.

[3] We reasoned that a significantly bigger sample in Experiment 1 would be appropriate because the tasks used in Experiment 2 may be more demanding, potentially making competence harder to detect and effect sizes smaller.

Procedure. We followed the same procedure used in Exp. 1 except for the warm-up trials (which are detailed in the *Supplementary Material*).

(1) Non-aspectual, change-of-location tasks. The basic sequence of events in these tasks, adapted from Buttelmann et al. (2009), was the following: E2 found two toys, and expressed a liking for them. She then announced that she had forgotten something and had to leave, and that she would leave the toys in the box during her absence. The following sequence varied between conditions (see Fig. 5 and *Supplementary Material*):

- Loc_FB condition: E2 left, and E1 reappeared from behind some curtains. E1 greeted the child and proposed playing a trick on E2. To this end, E1 took one of the toys out of the box and sneakily hid it under a tissue.
- Loc_TB condition: E1 appeared from behind the curtains, telling E2 she wanted to show her something before E2 left. As in Loc_FB, E1 then took one of the toys out of the box and hit it under a tissue. The only difference was that E1's actions were manifestly witnessed by E2, who did not leave the room. After E1 had hid the toy, E2 told the child she now really had to leave and left the room for several seconds.

Upon her return (in both conditions), E2 approached the table and reached into the box. She took out the remaining object and put it beside the box. She then began to search in the box again, saying "Hmm, Eh? I don't understand… but where is….". As in Exp. 1, this event is referred to as the *signal*. If the child did not react to the signal spontaneously, E2 sat down and expressed disappointment, saying "Hm. Oh no!". If the child did not react, E2 asked "Can you help me?". The task was repeated with a different box and two different toys.

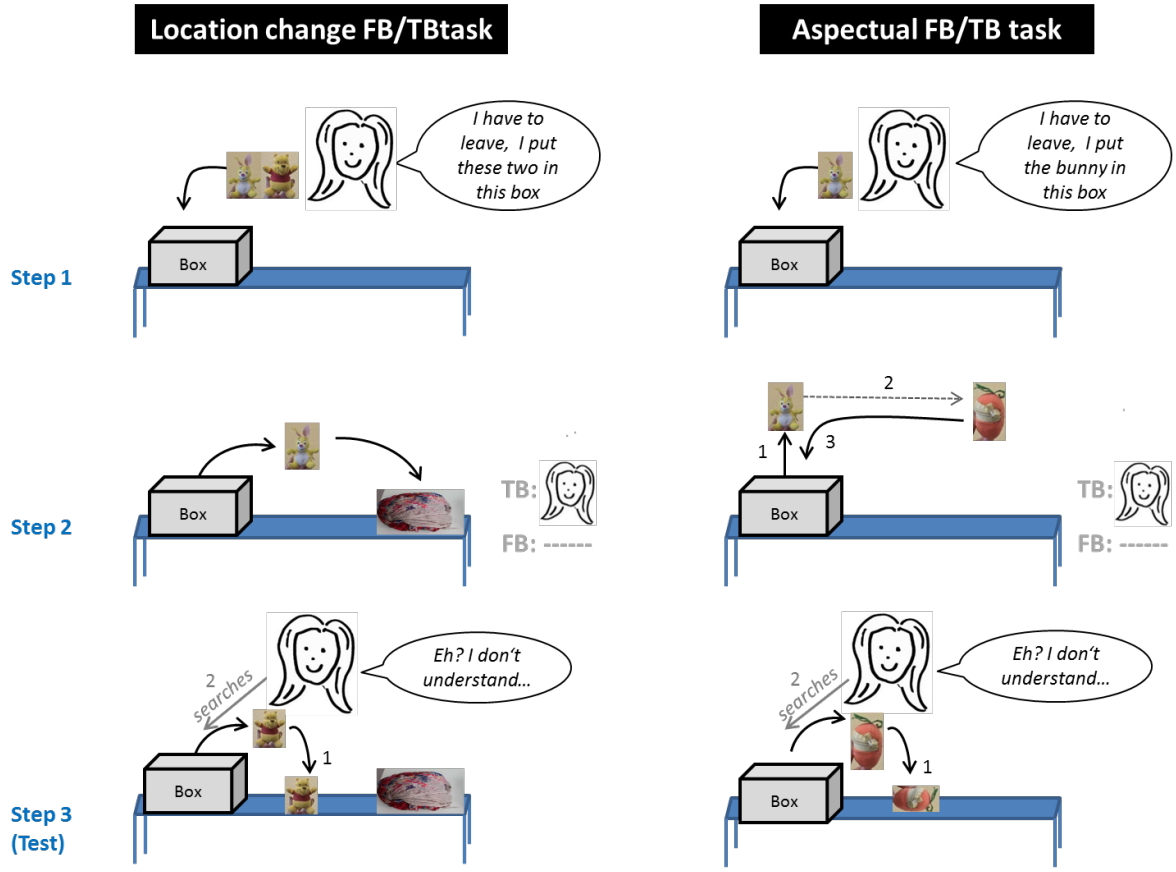


Figure 5. Schematic event sequences of the location and aspectual FB/TB tasks in Exp.. 2.

(2) Aspectual task. The procedure for the aspectual task was exactly as for the non-aspectual task, except for the following differences: When E1 appeared from behind the curtains for the first time, she introduced only one reversible toy to the child. She then transformed it into its other aspect, saying "Look! The bunny is also an carrot!" (for example). In Asp_FB, she did this in a sneaky way, whispering "Shh" and telling the child "'E2 does not know that, right?". The child was then asked, "Can you make it so that it is a bunny again?" (and was helped by E1 if unwilling or unable). E1 then placed the toy in its original aspect (e.g. bunny) on the table and returned to her place behind the curtains. After this, E2 came, found the toy, expressed her liking for it and put it into the box before leaving again.

Then, either in the absence (Asp_FB), or in the presence (Asp_TB), of E2, and in exactly the same sneaky or not sneaky way as in the non-aspectual tasks, E1 appeared from behind the curtains, took the toy out of the box, transformed it into its second aspect (e.g. carrot), and put it back into the box. When E2 reentered the room she reached into the box, took out the toy in its second, carrot aspect, put it aside and began to search again. E2 did all in this in the same manner she in which she acted in the non-aspectual tasks.

Coding. *A*s in Exp. 1, the child's first reaction to the signal was coded. The coder categorized this reaction as:

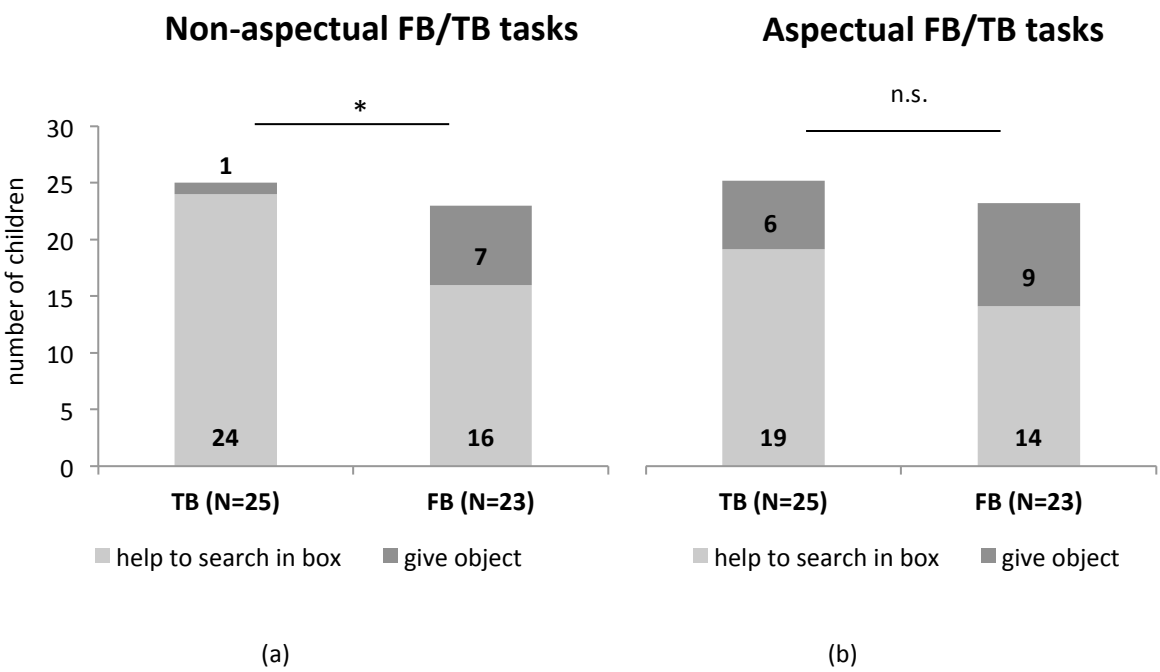
- help to search in the box: the child touched the box, pulled a grip, opened one of the doors, or tried to look inside.
- give the object: the child took the object, either from under the tissue (in the non-aspectual tasks) or else from on the table (in the aspectual tasks); or the child pointed to the object.
- ambiguous: the child showed a behavior that was clearly a reaction to the signal, but which did not fit any the above categories.
- invalid: a parent interfered; the child did not show any reaction at all; the child left the scene during the trial or did not pay attention to the event sequence; or the child could not be held back from reacting too early.

A third coder, who was blind to the experimental condition and hypotheses, coded 27 randomly selected tapes (36 trials, a 20% sample). This coder agreed with the original coder on nearly all trials ($\kappa = .96$).

Results

All in all, 109 children (51 in the non-aspectual and 58 in the aspectual conditions) received at least one valid (including ambiguous) trial. Of these children, 48 children had at least one valid and unambiguous trial per type of condition. (For details regarding the distribution of all trials, including invalid ones, see *Supplementary Material.*) As in Exp. 1, the first valid and unambiguous trial was used for the main analysis (see Figure 6; for an analysis including ambiguous trials, see Supplementary Material). In non-aspectual conditions, children's helping behavior differed significantly between Loc_FB and Loc_TB conditions (N = 48, p = .018, one-tailed Fisher's exact test). In the aspectual tasks, helping behavior did not differ significantly between Asp_FB and Asp_TB conditions (N = 48, p = .207, one-tailed Fisher's exact test).

Complementary analysis. Comparing the Loc_TB and Loc_FB conditions revealed that children tended open the box more often in Loc_TB than in the Asp_TB, (p = .049, one-tailed Fisher's exact test). No differences were between Loc_FB and Asp_FB (p = .379, one-tailed Fisher's exact test).



(a) (b)

Figure 6. Number of children showing the different kinds of helping behavior in the first valid and unambiguous trial (a) the non-aspectual tasks, (b) the aspectual tasks. [$^{*}p < .05$]

Discussion

In Exp. 2 we aimed to test for the patterns of results found in Exp. 1 with more stringent contrasts. To this end, the non-aspectual and aspectual tasks were matched as closely as possible in terms of irrelevant performance factors. The results largely converge with those of Exp. 1. Children's helping behaviour differed between the two non-aspectual conditions, Loc_FB and Loc_TB; but did not differ bewteen the two aspectual conditions, Asp_TB and Asp_FB.

General Discussion

Summary and limitations of the present studies

Two experiments tested for signature limits in the toddlers' early understanding of false belief as indicated in their active helping behavior. Both experiments found, in line with previous work, that children's helping behavior differed between a condition in which the protagonist had a true belief and a condition in which she had a false belief about the location of an object. However, in parallel aspectual tasks in which the protagonist had a true or false belief essentially involving aspectuality, toddlers' helping behavior did not differ significantly between true and false belief conditions. Experiment 2 controlled for differences in inferential complexity between the tasks in terms of working memory and other performance factors, replicating the basic results of Experiment 1.

It should be noted, however, that these findings are not particularly strong or unambiguous. It is true that a key prediction of the two-systems account was born out in the difference between performance on non-aspectuality, change-of-location tasks and performance on aspectual tasks, opposing accounts cannot be entirely ruled out by our findings. But when comparing performance on just the aspectuality and non-aspectuality FB tasks, we found a difference in Experiment 1 only. While the two-systems account does not directly generate a prediction concerning this comparison, we would be more confident in our findings if we had found a difference on just the aspectuality and non-aspectuality FB tasks in both experiments. Clearly, more comprehensive, systematic and sensitive tests for such potential patterns of contrast between aspectual and non-aspectual FB/TB tasks will be needed in future research.

Furthermore, while we attempted to match the non-aspectual and aspectual tasks as carefully matched as possible, some differences do remain. One such difference may complicate the interpretation of our findings. Consider the range of appropriate behaviours in each FB task. In neither the non-aspectuality FB task nor the aspectuality FB task is there a single correct response. When the protagonist asks for help, it is not incorrect to help her with what is certainly her proximal goal (opening this box) rather than with what is plausibly her more distal goal (getting her toy). But for methodological reasons, our tasks focus on trials where children help the protagonist with her more distal goal. And helping the protagonist with her more distal goal may be more demanding on our aspectuality FB task than on our non-aspectuality FB task. On the aspectuality FB task, helping the protagonist with getting her toy would ideally involve not only giving the toy to the protagonist but also pointing out, by communication or by turning it inside out, the hidden aspect under which the protagonist is actually looking for the object. If children appreciated that helping would ideally involve this, the complexity of the aspectual FB condition might in principle have simply overwhelmed and paralyzed children. While we cannot rule out this possibility altogether, we think it unlikely. This is in part because children did not need to show such complex responses: they could succeed merely by pointing to, or giving, the toy, without any further communicative attempt. Further, in another recent study with analogous design but in which the aspectuality was not realized by revertible objects with hidden aspects, children showed qualitatively the very same pattern of responses (Oktay-Gür, Schulz, & Rakoczy, 2017).

What do the results show?

Keeping in mind their limitations, what do the present findings show? On the one hand, children's performance might reflect the true limits of their ToM capacities. The limits

observed here are the very signature limits in early mindreading predicted by the two-system account (Apperly & Butterfill, 2009; Butterfill & Apperly, 2013). On this account, an early-developing system is assumed to enable a person to solve some ToM tasks but not others. Limits on which tasks such a system could solve follow from conjectures about the kind of mental state that they represent. In the case of non-aspectual, change-of-location FB tasks, operating with relational attitudes is sufficient for solving the task. In the case of tasks which require tracking beliefs or other mental states which essentially involve aspectuality, in contrast, operating with relational attitudes is not sufficient. Rather, fully-fledged propositional attitude concepts are needed. The pattern of findings in our studies might thus indeed reflect signature limits of an early-developing ToM capacity that operates by tracking relational attitudes rather than ascribing fully-fledged propositional attitudes. This possibility gains plausibility when considering the present findings in concert with a recent study which investigated the same structured contrast between aspectual and non-aspectual tasks but used explicit rather than implicit tasks (Rakoczy et al., 2015). In those experiments, children aged 4 to 6 years performed equally well on both aspecual and non-aspectual tasks, and both tasks were highly correlated with each other, and with a standard FB task. Taking these two sets of studies together highlights a contrast. There is a pattern of unity and convergence for performance on explicit FB tasks, and a pattern of disunity and dissociation for performance on implicit tasks. This contrast was recently found in another line of research (Low et al., 2014; Low & Watts, 2013). It is just what is to be expected according to the two-systems account.

Alternatively, however, the present findings might be explained in terms of performance (rather than competence) factors. For example, subjects in our experiments may have assumed, for some reason, that the protagonist was omniscient about aspects of

objects but not about their locations. Such an assumption might in principle have been triggered by subjects' difficulty in understanding and coordinating different aspects of objects (Perner, Mauer, & Hildenbrand, 2011). Crucially, according to this interpretation, children were led by extraneous task demands to make such a omniscience assumption in the aspectual but not in the non-aspectual conditions, even though in principle they would be capable of understanding that people can fail to be omniscient of objects' aspects as well as their locations.

How might future research decide between these two interpretations? The two-systems interpretation predicts that the difference in performance between non-aspectual and aspectual tasks will be observed in implicit but not explicit tasks, whereas the alternative ('omniscience'/extraneous task demands) interpretation predicts that, all other things being equal, this difference in performance will appear in explicit as well as implicit tasks. A fundamental problem for the omniscience interpretation, however, is that other studies have measured implicit and explicit responses to a single scenario. These studies have uniformly observed a difference in performance between aspectual and non-aspectual implicit tasks which disappears on explicit tasks (Low et al., 2014; Edwards & Low, 2017). Furthermore, as already mentioned, another study developed explicit tasks implementing the very same contrasts between aspectual and non-aspectual we used, and found no performance differences between aspectual and non-aspectual tasks (Rakoczy et al., 2015). This suggests that an interpretation in terms of omniscience about aspectuality and extraneous task demands is unlikely to explain the present findings.

Three potential methodological caveats with regard to the present findings should be mentioned. First, it might be wondered whether the stimuli in the present study were appropriate for testing children's understanding of false beliefs involving aspectuality. The

objects used here (e.g. bunny/rabbit) were selected because earlier research on children's object individuation with exactly these stimuli show that even 1-year-olds do understand the dual nature of these objects (Cacchione, Schaub, & Rakoczy, 2013). However, it might be objected that children could have represented the situation in terms that would undermine our claim to be testing an understanding of false beliefs involving aspectuality. After all, children might have represented the protagonist as having beliefs about one object (a bunny, for example) with another object (a carrot, for example) hidden inside it. If this were true, what we term 'aspectual' tasks would in fact have been non-aspectual tasks about location and containment. As this objection nicely illustrates, it is difficult or impossible to create situations that can only be understood as involving aspectuality. However, two points should be noted in response to this objection. First, if the children in the aspectuality conditions really had represent the protagonist as having beliefs about one object being hidden in another, we would expect their performance to differ between FB and TB just as it did on the non-aspectuality conditions. In fact this is not what we observed. Second, a recent study with older children (Rakoczy et al., 2015) compared performance on tasks involving the stimuli used here with new stimuli that also had dual aspects but could not be construed as involving one object hiding another (for example, a single object featured both as Susi and as the doctor). Children's performance with the two sets of stimuli was absolutely comparable. This suggests that the findings from the present experiments are unlikely to depend on irrelevant peculiarities of the stimuli.

A second methodological caveat concerns the dependent measure of our tasks. We asked, What will toddlers spontaneously do in response to a protagonist's attempt to open, or search in, a box? Using this dependent measure limits the strength of the findings compared to standard, explicit ToM tasks which use a two-alternative, forced-choice

measure with exactly one correct answer. Future studies might extend the current work to compensate for this weakness by, for example, combining multiple dependent measures. For example, our tasks could be combined with measures of proactive gaze and action trajectories.

Third, and relatedly, null findings in the aspectual tasks from Experiments 1 and 2 are inherently difficult to interpret statistically. The absence of significant differences between FB and TB cases may simply be due to a lack of power. This is a particularly pressing concern given the fundamentally binary quality of our data. Future studies might overcome this weakness not only by combining multiple dependent measures but also by using continuous measures.

Conclusions and future directions

All in all, the two experiments reported here present preliminary evidence compatible with the claim that early-developing ToM capacities are subject to signature limits. Taken together with recent findings that on an explicit level older children do not show the same signature limits but perform uniformly across a wide range of FB tasks (Rakoczy et al., 2015), and taken together with converging evidence from similar studies with looking time measures (Low et al., 2014; Low & Watts, 2013), the present findings are in line with predictions of the two-systems account of mindreading.

But while being in line with the predictions of the two-systems account, the present findings by themselves cannot strictly decide between this and alternative accounts given the methodological caveats just identified. Further research is needed to test for patterns of performance in early theory of mind, and in particular for signature limits. What is

required if we are to decide between competing accounts is a broad range of implicit tasks concerning the generality and flexibility of young children's abilities to track beliefs and other mental states (Yott & Poulin-Dubois, 2012, 2016). Such tasks must involve a broad range of scenarios, belief contents and methods (such violation-of-expectation, anticipatory looking, communicative and interactive measures). Perhaps most importantly, more attention to when and why infants fail to track mental states is required, ideally in conjunction with further direct comparisons between performance on implicit and explicit tasks. This is, after all, where the competing accounts make clearly different predictions.

One particularly instructive strategy in this context may be to investigate children's readiness to *learn* to track beliefs about novel scenarios (see also Heyes, 2014). For example, when children are confronted repeatedly with scenarios in which novel objects change their (unusual) locations or their (unusual) aspects in novel ways, children's learning history might be particularly instructive. According to an early competence account, infants and young children have a full-blown concept of belief. It follows that after initial asymmetries regarding previous experience with unusual locations or unusual aspects have been evened out through training, young learners should be capable of solving aspectual and non-aspectual FB tasks with equal ease. In contrast, the two-systems theory would predict that an initial asymmetry in performance between aspectual and non-aspectual FB tasks should persist even despite learning about object's unusual locations and aspects.

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. doi: 10.1037/a0016923.

Apperly, I. A., & Robinson, E. J. (1998). Children's mental representation of referential relations. *Cognition, 67*(3), 287-309.

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

Baillargeon, R., Scott, R. M., & He, Z. (2010). False-belief understanding in infants. *Trends in Cognitive Sciences, 14*(3), 110-118. doi: 10.1016/j.tics.2009.12.006

Buttelmann, D., Carpenter, M., & Tomasello, M. (2009). Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition, 112*(2), 337-342.

Buttelmann, D., Over, H., Carpenter, M., & Tomasello, M. (2014). Eighteen-month-olds understand false beliefs in an unexpected-contents task. *Journal of experimental child psychology, 119*, 120-126.

Buttelmann, F., Suhrke, J., & Buttelmann, D. (2015). What you get is what you believe: Eighteen-month-olds demonstrate belief understanding in an unexpected-identity task. *Journal of Experimental Child Psychology, 131*, 94-103. doi: http://dx.doi.org/10.1016/j.jecp.2014.11.009

Butterfill, S. A., & Apperly, I. A. (2013). How to Construct a Minimal Theory of Mind. *Mind & Language, 28*(5), 606-637. doi: 10.1111/mila.12036

Cacchione, T., Schaub, S., & Rakoczy, H. (2013). Fourteen-month-old infants infer the continuous identity of objects on the basis of non-visible causal properties. *Developmental Psychology, 49*(7), 1325-1329. doi: 10.1037/a0029746

Carey, S. (2009). *The origin of concepts*. New York, NY: Oxford University Press; US.

Carruthers, P. (2013). Mindreading in Infancy. *Mind & Language, 28*(2), 141-172. doi: 10.1111/mila.12014

Clements, W. A., & Perner, J. (1994). Implicit understanding of belief. *Cognitive Development, 9*(4), 377-395.

Csibra, G. (2012). Revising the belief revision paradigm. *Culture and Cognition Blog (http://www.cognitionandculture.net/home/blog/44-pierre-jacobs-blog/2455-do-we-use-different-tools-to-mindread-a-defendant-and-a-goalkeeper)*.

Edwards, K. and Low, J. (2017). Reaction time profiles of adults' action prediction reveal two mindreading systems. *Cognition*, 160:1–16.

Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences, 8*(7), 307-314.

Helming, K. A., Strickland, B., & Jacob, P. (2014). Making sense of early false-belief understanding. *Trends in Cognitive Sciences, 18*(4), 167-170.

Heyes, C. (2014). False belief in infancy: a fresh look. *Developmental Science, 17*(5), 647-659. doi: 10.1111/desc.12148

Jacob, P. (2012). Do we use different tools to mindread a defendant and a goalkeeper? *Cognitin & Culture (Online science blog), http://www.cognitionandculture.net/home/blog/44-pierre-jacobs-blog/2455-do-we-use-different-tools-to-mindread-a-defendant-and-a-goalkeeper*.

Kamawar, D., & Olson, D. R. (1999). Children's representational theory of language: The problem of opaque contexts. *Cognitive Development, 14*(4), 531-548. doi: 10.1016/s0885-2014(99)00018-0

Kamawar, D., & Olson, D. R. (2011). Thinking about representations: The case of opaque contexts. *Journal of Experimental Child Psychology, 108*(4), 734-746. doi: 10.1016/j.jecp.2010.10.005

Knudsen, B., & Liszkowski, U. (2012). 18-month-olds predict specific action mistakes through attribution of false belief, not ignorance, and intervene accordingly. *Infancy, 17*(6), 672-691.

Leslie, A. M. (2005). Developmental parallels in understanding minds and bodies. [Review]. *Trends Cogn Sci, 9*(10), 459-462. doi: 10.1016/j.tics.2005.08.002

Low, J., Apperly, I. A., Butterfill, S. A., & Rakoczy, H. (2016). Cognitive Architecture of Belief Reasoning in Children and Adults: A Primer on the Two-Systems Account. *Child Development Perspectives, 10*(3), 184-189. doi: 10.1111/cdep.12183

Low, J., Drummond, W., Walmsley, A., & Wang, B. (2014). Representing How Rabbits Quack and Competitors Act: Limits on Preschoolers' Efficient Ability to Track Perspective. *Child Development, 85*(4), 1519-1534. doi: 10.1111/cdev.12224

Low, J., & Watts, J. (2013). Attributing False Beliefs About Object Identity Reveals a Signature Blind Spot in Humans' Efficient Mind-Reading System. *Psychological Science, 24*(3), 305-311. doi: 10.1177/0956797612451469

Oktay-Gür, N., Schulz, A., & Rakoczy, H. (2017). Children exhibit different performance patterns in explicit and implicit theory of mind tasks. *Manuscript submitted for publication*.

Onishi, K. H., & Baillargeon, R. (2005). Do 15-Month-Old Infants Understand False Beliefs? *Science, 308*(5719), 255-258. doi: 10.1126/science.1107621

Perner, J., Mauer, M. C., & Hildenbrand, M. (2011). Identity: Key to children's understanding of belief. *Science, 333*(6041), 474-477. doi: 10.1126/science.1201216

Perner, J., & Roessler, J. (2012). From infants' to children's appreciation of belief. *Trends in Cognitive Sciences, 16*(10), 519-525. doi: 10.1016/j.tics.2012.08.004

Rakoczy, H. (2012). Do infants have a theory of mind? *British Journal of Developmental Psychology, 30*(1), 59-74. doi: 10.1111/j.2044-835X.2011.02061.x

Rakoczy, H. (2017). In defense of a developmental dogma: children acquire propositional attitude folk psychology around age 4. [journal article]. *Synthese, 194*(3), 689-707. doi: 10.1007/s11229-015-0860-8

Rakoczy, H., Fizke, E., Bergfeld, D., & Schwarz, I. (2015). Explicit theory of mind is even more unified than previously assumed: belief ascription and understanding aspectuality emerge together in development. *Child Development, 86*(2), 486-502. doi: 10.1111/cdev.12311

Rhodes, M., & Brandone, A. C. (2014). Three-year-olds' theories of mind in actions and words. [Original Research]. *Frontiers in Psychology, 5*. doi: 10.3389/fpsyg.2014.00263

Rubio-Fernández, P. (2013). Perspective tracking in progress: Do not disturb. *Cognition, 129*(2), 264-272. doi: http://dx.doi.org/10.1016/j.cognition.2013.07.005

Rubio-Fernández, P., & Geurts, B. (2013). How to Pass the False-Belief Task Before Your Fourth Birthday. *Psychological Science, 24*(1), 27-33. doi: 10.1177/0956797612447819

Ruffman, T., & Perner, J. (2005). Do infants really understand false belief? *Trends in Cognitive Sciences, 9*(10), 462-463.

Russell, J. (1987). 'Can we say ...?' Children's understanding of intensionality. *Cognition, 25*, 289-308. doi: 10.1016/S0010-0277(87)80007-0

Schneider, D., Bayliss, A. P., Becker, S. I., & Dux, P. E. (2012). Eye movements reveal sustained implicit processing of others' mental states. *Journal of Experimental Psychology: General, 141*(3), 433-438.

Scott, R. M., & Baillargeon, R. (2009). Which penguin is this? Attributing false beliefs about object identity at 18 months. *Child Development, 80*(4), 1172-1196.

Scott, R. M., Richman, J., & Baillargeon, R. (2015). Infants understand deceptive intentions to implant false beliefs about identity: new evidence for early mentalistic reasoning. *Cognitive Psychology, 82*, 32-56.

Sirois, S., & Jackson, I. (2007). Social cognition in infancy: A critical review of research on higher order abilities. *European Journal of Developmental Psychology, 4*(1), 46-64. doi: Doi 10.1080/17405620601047053

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

Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science, 18*(7), 587-592.

Sprung, M., Perner, J., & Mitchell, P. (2007). Opacity and discourse referents: Object identity and object properties. *Mind & Language, 22*(3), 215-245. doi: DOI 10.1111/j.1468-0017.2007.00307.x

Surian, L., Caldi, S., & Sperber, D. (2007). Attribution of beliefs by 13-month-old infants. *Psychological Science, 18*(7), 580-586. doi: 10.1111/j.1467-9280.2007.01943.x

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development, 72*(3), 655-684. doi: Doi 10.1111/1467-8624.00304

Yott, J., & Poulin-Dubois, D. (2012). Breaking the rules: Do infants have a true understanding of false belief? [Article]. *British Journal of Developmental Psychology, 30*(1), 156-171. doi: 10.1111/j.2044-835X.2011.02060.x

Yott, J., & Poulin-Dubois, D. (2016). Are Infants' Theory-of-Mind Abilities Well Integrated? Implicit Understanding of Intentions, Desires, and Beliefs. *Journal of Cognition and Development, 17*(5), 683-698. doi: 10.1080/15248372.2015.1086771

Are there signature limits in early theory of mind?

Supplementary Material

Experiment 1

Method

Participants. 67 children (26 females) were included in the final sample,. The children's mean age was 31 months (*SD* = 2.66; *range* = 26-36). 9 additional children were not included in the analysis because they were uncooperative (N=8) or due to experimental error (*N* = 1). Children in both experiments were recruited from a local database of parents who had volunteered to participate in child development studies and were tested in the lab. Children were all native German speakers, and came from mixed (mostly middle to high) socio-economic backgrounds.

Design. Children were randomly assigned to one of two groups who received either a change of location or an aspectual FB/TB task. Within each group, half of the children were randomly assigned to sub-groups receiving FB or TB versions of the task. There were thus 4 conditions: non-aspectual, change-of-location false belief (Loc_FB), non-aspectual, change-of-location TB (Loc_TB), aspectual false belief (Asp_FB) and aspectual true belief (Asp_TB). Each child received two (in two exceptional cases three; see below) trials of one of the four conditions.

Materials. In the non-aspectual tasks, one toy object (a ball, a soft toy rabbit, a toy dog or toy donkey) per trial and two boxes were used. In the aspectual task, one reversible mini soft toy (a bear, bunny, tiger, pig), which could be turned inside out through a zipper and transformed (into a honey barrel, a carrot, a snail and a cake, respectively, see Figure 1), and one box were used per trial. Boxes were covered with a tissue on the inside to leave children ignorant about the (empty) content. A puppet called Susi (animated by E2) was used as protagonist.

Procedure. Each child was tested in an interactive play setting with two experimenters, with a parent present during the session. Each session began by the child's playing one or two warm-up games with E1 and the puppet Susi. Susi was animated by E2. The warm-up games were followed by a standardized warm-up trial. For the warm-up trial and the subsequent test trials, children were asked to sit on a cushion beside E1, and 2m from E2 and the puppet. The parent was allowed to sit behind the child, and was instructed not to interfere by saying or showing anything to the child.

*Figure 1.*Examples of the stimuli used in the aspectual false belief task.

Standardly, children received two trials of the condition they were randomly assigned to. In exceptional cases a third trial was administered; this was done if the child's responses in both trials were coded as 'ambiguous' (see below for details on coding).

Warm-up trial. E1 introduced Susi and the child to a soft toy lion. E1 then gave the lion to Susi, who expressed a liking for it. Susi then announced that she had forgotten something, and left the room through a door behind her. In her absence, E1 asked the child to play a trick on Susi by hiding the lion (e.g. in a drawer). After this, E1 and the child called Susi. On her return, Susi looked around in the room searching for the lion. Susi then asked the child for help. If the child hesitated to help, E1 prompted the child to help Susi. If the

child still refused, E1 said: “Come on, we’ll do it together”. When Susi was shown the lion, she said, “Ah, there is the lion. Thank you!”

Table 1. Detailed event sequences of the spontaneous helping tasks in Experiment 1

	Non-aspectual, change-of-location tasks	Aspectual tasks
1	**Warm-up trial** E2 with puppet Susi sits approximately 2m vis á vis the child and E1, all on their particular cushion. E1 shows a plush lion to both: "Look, what I've got here!" Susi says "Wow, a lion, it's great, I want it!" and reaches for it. E1 says "Do you want it, Susi?" and gives it to her. Susi says "Oh no, I forgot sth.! I have to leave for short. I'll be right back." She leaves the room through a door behind her. E1 whispers to the child "You know what? Now we'll play a trick on Susi. We hide the lion. Where should we hide it?" E1 and the child choose a location and hide the lion. E1 behaves in a sneaky way, giggling and gesturing and whispering "Shh". Then they sit down at their cushions. They call Susi into the room. Susi re-enters, looks around and says "Oh, but where is the lion?" and expresses disappointment: "Hmm. Oh no." She asks the child for help ("Can you help me?"). If the child hesitates to help, E1 prompts "You can help her". If the child still refuses, E1 proposes to help together ("Should we help her together?"). When Susi is shown the lion she says "Ah, there it is, thank you".	
2	--	**Introduction of the dual aspect objects** Susi announces "Oh no, I forgot something! I have to leave for short" and leaves. E1 shows a reversible mini soft toy to the child and its dual aspectuality: "Look what I got here! A bunny! And now I'll show you something: the bunny is also a carrot. Where is the bunny now? Can you make it so that it's a bunny again?" The child transforms the object, if necessary with the help of E1. E1 introduces the 2nd object in the same way. Then she emphasizes: "And Susi doesn't know that, right?" She asks the child to play sth. and to call Susi back.
3	**Introduction of the boxes/box** E1 shows one box to the child: "Look I have a box here. And you can open it like this. You want to try?" She waits until the child has opened the box, eventually with her help. She pays attention that the child does not grasp inside and closes the box again. (She shows the 2nd box in the same way in the Non-aspectual tasks). During all time, Susi and E2 turn their heads away from the scene.	
4	**Introduction of the target object** E1 shows the first target object to Susi and the child ("Look what I've got here!"). Susi reaches for the object saying "Wow, a 'bunny', it's great!" E1 gives it to her ("Do you want it?"). Now Susi announces: "Oh no, I forgot sth.! I have to leave for short. I put the bunny in this box. 'E1' can you help me? Okay. Because now I have to leave for short". She puts the bunny into the one box and leaves.	
5	**Location-change** **False belief:** E1 says "You, [name of child], shall we play a trick on Susi? Look, what I do now. Okay. Hihi. Shh!" She changes the location of the target in a sneaky way emphasizing at the end "And Susi can't see that, right?" **True Belief:** Susi comes back a few seconds after she left. E1 says "Look, Susi and [name of child]! Like this and like this!" She changes the location of the target, thereby alternating gaze between the child and the puppet. Susi observes this event closely saying "Ah I see! Yes!" Then she takes her seat on her cushion between the boxes again.	**Aspect-change** **False belief:** E1 says "You, [name of child], shall we play a trick on Susi? Look, what I do now. Okay. Hihi. Shh!" and transforms the object into its 2nd aspect in a sneaky way and puts it back, emphasizing at the end "And Susi can't see that, right?" **True Belief:** Susi comes back a few seconds after she left. E1 says "Look, Susi and [name of child]! Like this and like this!" She changes the aspect of the target, thereby alternating gaze between the child and the puppet. Susi observes this event closely saying "Ah I see! Yes!" Then she takes her seat on her cushion again.
6	--	**Location-change** E1 says "And now, look, Susi and [name of child], what I do now! Like this and like this!" She takes the target out of the box and places it in the same distance as the box from the Susi and the child on the carpet. Susi observes this event closely saying "Ah I see! Yes!" Then she takes her seat on her cushion again.
7	**The signal** Susi looks 2 times from one box to the other/from the box to the object, saying "Hmm" in a thoughtful way. Then she says "Okay", approaches one box, pulls two times on the top/lid/cover, expressing effort ("Uff") and then sits down at her cushion, looking down, expressing disappointment ("Hmm. Oh no!"). If the child had not reacted until now, she says "Hmm", looks at the child and down again. If the child did not react she asks "Can you help me?" If the child hesitates, E1 prompts: "You can help her". If the child still refuses, E1 proposes "Should we help her together?" but waits where the child moves. If the child helps to open one box, E1 disrupts quickly and says: "Come on we play another game and removes box(es) and object to begin a second trial with different boxes and a different target object. If the child in the aspectuality condition gives the object to Susi, she neutrally says "Hmm" and waits a few seconds (to see if the child also transforms the object). Then E1 disrupts and removes box and object.	

Test trials

(1) Non-aspectual tasks (following Buttelmann et al., 2009). E1 introduced two boxes (green & blue or yellow & red in counterbalanced order) to the child (Susi turned away so she couldn't see), taught the child how to open and close them (pulling an elastic band from a hook) and prompted the child to open the box herself. The child opened and closed both boxes successfully (with the help of E1 if the child was not able to do so on her own) but was prevented by E1 from grasping inside under the tissue. E1 placed the two boxes equidistantly between Susi and the child (approximately 1,5m apart, color and side counterbalanced) and showed one of three toys (which one was fully counterbalanced) to them. Susi excitedly expressed her desire/liking for the toy and E1 gave it to her. Susi then announced again that she forgot something and would put the toy into box 1 for the time of her absence. She did so with the help of E1 (because she wasn't able to open the box on her own). The following sequence varied between conditions (*see Fig. 2)*:

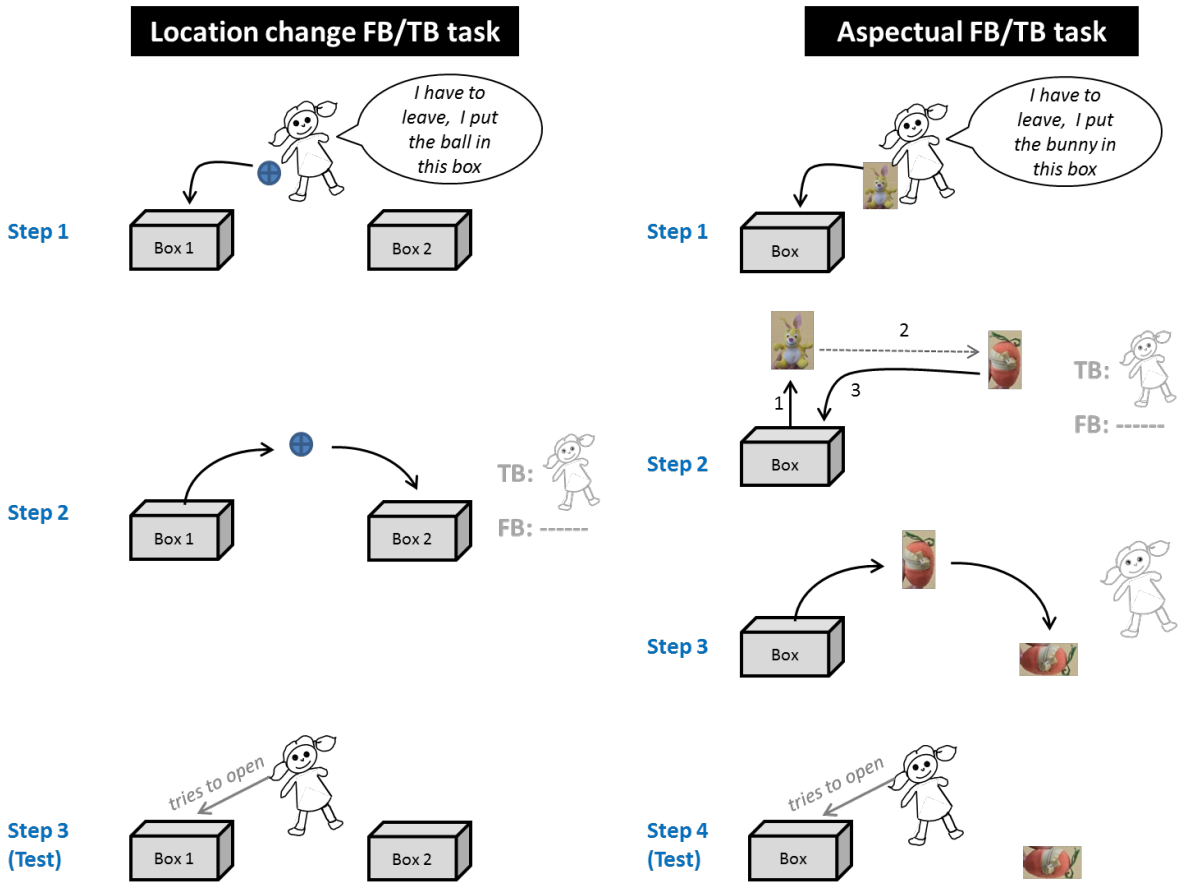


Figure 2. Schematic event sequences of the location and aspectual tasks in Exp. 1.

- FB condition: Susi left and in her absence, E1 proposed to 'play a trick' on Susi. She took the toy out of the box and placed it in the other box, giggling and gesturing and whispering "Shh".
- TB condition: Susi left and returned a few seconds later. In her presence, E1 took the toy out of the box and placed it in the other box, saying: "Look, Susi and [child's name]!" and alternating gaze between the child and E2. Susi observed the location change closely, saying: "Ah I see! Yes!"

After her return (FB)/the observed location change (TB) Susi moved back to her position between the boxes, looked from one box to the other saying first "Hmm" and then "Okay". She tried to open the box where the object formerly was located by pulling on the lid shortly

(this event is referred to as the 'signal' in what follows). She failed to open the box and sat down again showing disappointment: "Hmm. Oh no!" If the child did not react, she asked "Can you help me?" If the child still showed no reaction, E1 prompted the child to help and if the child still refused E1 asked to help together ("Should we help her together?") but didn't move and waited where the child would go. Children's reactions to the signal were coded. The task was repeated with different boxes and a different target toy.

(2) Aspectual task. After the warm-up trial children were first familiarized with the dual aspectuality of the target objects. To this end, the puppet left the room and E1 showed the first mini soft toy (which one was fully counterbalanced) to the child. Then she transformed it to its 2nd aspect saying: "Look! The X is also a Y!" (e.g. bunny and carrot). She did that in a sneaky way, whispering "Shh" and told the child: "Susi does not know that, right?" to ensure that the child would not assume that Susi knew all aspects of the object. The child was then asked to transform the object into its original state: "Can you make it so that it is a bunny again?" If the child was not able or not willing, E1 helped and both transformed the object together. This was repeated with a second object which was later used for the second test trial.

For the test trials Susi returned. E1 now showed one box (colors counterbalanced across children and trials) to the child (Susi turned away so she couldn't see this), taught the child how to open and close it (pulling an elastic band from a hook) and prompted the child to open the box herself. The child opened and closed the box successfully (with the help of E1 in case the child was not able to do so on her own) but was prevented by E1 from grasping inside under the tissue. E1 placed the box at one side (approximately 0,8m from the center, side counterbalanced) between Susi and the child and showed the first mini soft toy to them. Susi excitedly expressed her desire/liking for the toy and E1 gave it to her. Susi

then announced again that she forgot something and would put the toy in one of the boxes for the time of her absence. She did so with the help of E1. The following sequence varied between conditions:

- FB condition: Susi left and E1 proposed to 'play a trick' on Susi. She took the toy out of the box, transformed it into the other aspect (e.g. a carrot), giggling and gesturing and whispering "Shh", and returned it into the box.
- TB condition: Susi left and returned a few seconds later. In her presence, E1 took the toy out of the box, transformed it - while saying: "Look Susi and [child's name]!" and alternating gaze between the child and E2 - and put it back into the box. Susi observed this event closely, saying: "Ah I see! Yes!"

After her return (FB)/ the observed aspect change (TB) Susi moved back to her original position. Now, E1 stated again: "Look Susi and [child's name]!" and took the transformed toy (e.g. carrot) out of the box and placed it approximately 1m away from the box on the carpet equidistant between Susi and the child. Susi observed this location-change closely saying "Ah I see! Yes!", then she looked from the box to the object saying first "Hmm" and then "Okay". She then tried to open the box. This signal and the following events were the same as in the location-change task. The task was repeated with a different box and a different target reversible toy.

Coding. For both studies sessions were coded from videotape by one coder. The child's first reaction to the signal of the protagonist (the attempt to open the box) was coded and it was determined whether it fell into one of the two central categories:

- open box 1 (in both tasks): if the child clearly referred to box 1 (by the puppet) by approaching the box and opening it successfully or by trying to open it (e.g. pulling the elastic band) even if not successfully.
- open box 2 (in the non-aspectual tasks)/give the object to the puppet (in the aspectual tasks): if the child clearly referred to box 2 by approaching the box and opening it successfully or by trying to open it, or handed the object to the puppet.

There were two ways in which a given trial could fail to fall into one of these two categories:

- ambiguous: This coding applied if the child showed a behavior that clearly was a reaction to the signal (i.e. sat still before and began to respond after the puppet's attempt to open box 1) but which did not fit either of the two above categories. This could be because the child did perform both above behaviors (e.g. taking the object and putting it in box 1) or because she did something totally irrelevant (e.g. shaking the hand of the puppet when asked for help).
- invalid: This coding was given if a parent interfered, if the child did not show any reaction at all, if the child left the scene during the trial or did not pay attention to the event sequence, or if the child could not be held back from reacting too early, that is, producing a behavior (such as opening box 1 or fetching the object) before the signal (and thus not in response to the implicit test question – which makes it impossible to interpret what this behavior means regarding their ascription of beliefs to the protagonist.

A second observer who was blind for the experimental condition and hypotheses coded a 20% sample of 14 randomly selected tapes for reliability which was $\kappa = 1.00$ for the four-category-coding (open box 1, open box 2/give object, invalid, ambiguous).

Results

For 3 children both trials were coded as invalid because the parent interfered in both trials (*N* = 1 in the Asp_FB condition), because the child did not pay attention to the event sequences (*N* = 1 in the Loc_TB condition) or showed no reaction to the signal (*N* = 1 in the Asp_FB condition). Data from 64 children were thus used in the final analysis. Mean age was *M* = 31;3 (*SD* = 3.09, 6 girls, 10 boys) in the non-aspectual FB condition, *M* = 31;3 (*SD* = 2.14, 6 girls, 10 boys) in the non-aspectual TB condition, *M* = 31;2 (*SD* = 2.73, 6 girls, 10 boys) in the aspectual FB and *M* = 31;1 (*SD* = 2.48, 7 girls, 9 boys) in the aspectual TB condition. There was no significant age difference between conditions ($F(3) = .11$, $p = .96$).

Trials coded as invalid or ambiguous. Of the 64 children, the first trial was coded as invalid in 8 cases (4 in the Loc_FB, 3 in the Loc_TB and 1 in the Asp_FB condition) because of experimental error (*N* = 1), because the child was inattentive (*N* = 3), did not react to the signal (*N* = 1) or reacted before the signal (*N* = 3). For 2 children the first trial was coded as ambiguous reaction (both in the Asp_TB condition) because the child took the not-referred-to object and tried to put it in the box (*N* = 1) or gave another object to Susi (N = 1). 2 children showed an ambiguous reaction in both first trials (gave a car to Susi/shook her hand; in the Loc_FB and Asp_TB condition). As explained above, because both trials were ambiguous, a third trial was administered in these cases. All in all, 64 children (32 each in the non-aspectuality and aspectuality conditions) received at least one valid (including ambiguous) trial and for these children at least one trial was also unambiguous.

Main analyses. For the main analysis, the first valid and unambiguous trial of the 64 children to which this applied was used. Invalid and ambiguous trials were replaced by the second trial if this was a valid trial. When possible, one-tailed tests were conducted whenever directed a priori hypotheses were tested (such that children perform more often "open box 1" in contrast to "open box 2/give object" behavior in the TB than in the FB conditions). Figure 3 depicts children's helping behavior in the first valid and unambiguous trial. Children's helping behavior differed significantly between non-aspectual FB and TB conditions (p = .037, Φ_{Cramer} = .38, one-tailed Fisher's exact test). Concerning the aspectual conditions, in contrast, helping behavior did not differ significantly between FB and TB (p = .166, Φ_{Cramer} = .26, one-tailed Fisher's exact test).

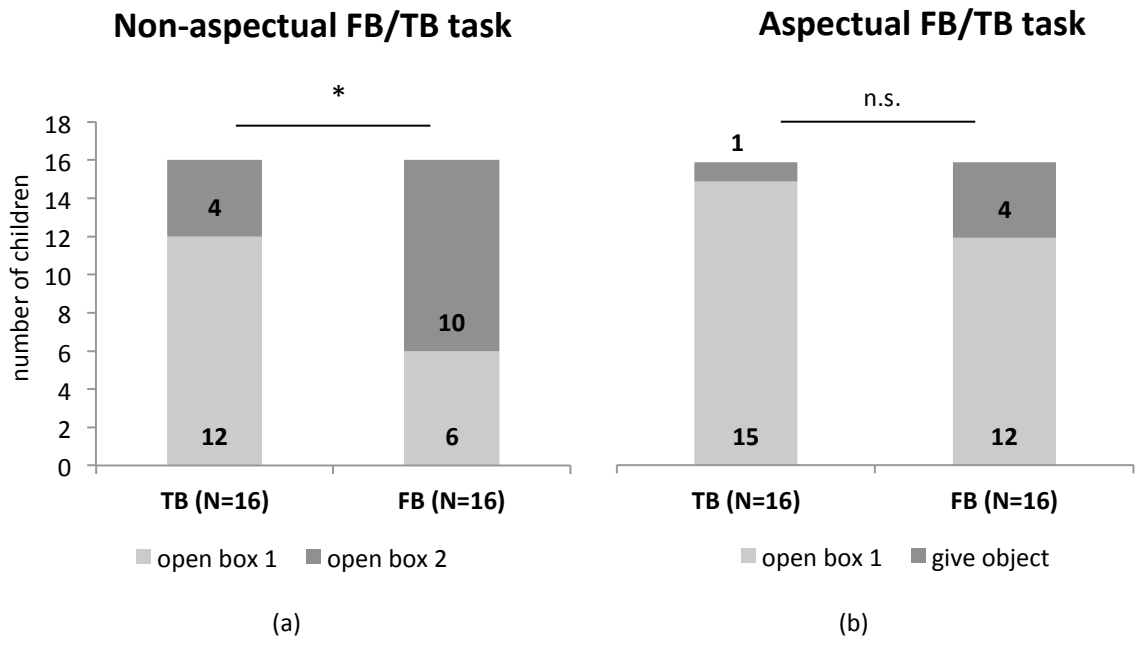


Figure 3. Number of children showing the different kinds of helping behavior in the first valid and unambiguous trial of (a) the non-aspectual tasks, (b) the aspectual task. [$^{*}p < .05$]

Complementary analysis. When comparing the non-aspectual and aspectual tasks for FB and TB, respectively, results revealed that children tended to perform differently in the location than in the aspectual FB tasks (opening box 1 less often in the location than in the aspectual FB tasks, p = .037, one-tailed Fisher's exact test). No differences were found for the two TB tasks (p = .166).

Analyses with ambiguous trials included. The difference between conditions in the non-aspectual tasks was less clear when ambiguous responses were included in a two-tailed exact test as a third category (see Table 1), with a trend remaining (p = .073, V_{Cramer} = .39, two-tailed Fisher's exact test). In the aspectual task, conditions did not differ significantly also when including ambiguous trials in the analysis (p = .110, V_{Cramer} = .39, two-tailed Fisher's exact test).

Table 1. *Helping behavior in the first valid trial including ambiguous responses.*

	Non-aspectual task		Aspectual task	
	True belief	False belief	True belief	False belief
open box 1	12	6	12	12
open box 2/give object	4	9	1	4
ambiguous response	0	1	3	0

Aggregate score including all trials. Since for a large number of children at least one trial had been coded invalid or ambiguous (N = 10 in the non-aspectual task; N = 4 in the aspectual task) a sum score of two trials did not constitute a valid measure of children's overall performance. In order to also include the second valid and unambiguous trials in the analysis, an aggregate score was calculated for which children were separated into two

groups: children who consistently referred to box 2 (non-aspectual task)/ gave the object (aspectual task) in their valid and unambiguous trials and children who did not do so (see Table 2). One-tailed Fisher's exact tests confirmed the results of the main analysis: in the non-aspectual task helping behavior differed significantly between the TB and FB condition conditions (p = .037, Φ_{Cramer} = .38), whereas in the aspectual task no such difference was found (p = .50, Φ_{Cramer} = .11).

Table 2. *Aggregate score of children's helping behavior in both trials.*

	Non-aspectual task		Aspectual task	
	TB	FB	TB	FB
consistently referred to box 2/give object	4	10	1	2
did not consistently refer to box 2/give object	12	6	15	14

Experiment 2

Method

Participants 137 children (61 females) were included in the final sample[4]. Children's mean age was 26 months (*SD* = 1.92; *range* = 24-30). Five additional children were tested but not included in the analysis because they were uncooperative.

Design. The design was the same as in Exp. 1. In Exp. 2 children received up to two subsequent trials of one of the four conditions to which they were randomly assigned: non-aspectual FB, non-aspectual TB, aspectual FB, aspectual TB. (Two short additional tasks were administered during warm-up and at the end of the session, which focused on another research question and are therefore not reported here.)

Materials. In the non-aspectual tasks, two different toys (whose reversible function was not shown) per trial (a bear, bunny, tiger or pig) were used as stimuli and one (green or orange) box. In the aspectual task, one reversible mini soft toy per trial and one box were used. Boxes had several doors and openings: one opening on the upper surface and three doors which could be opened by pulling a grip (see Figure 4). All openings were covered with a tissue on the inside and additional tissue panels were spanned inside the box so that it was not easy to determine if the box was empty or not. In contrast to Exp. 1, E2 herself played the role of the protagonist to avoid that the younger children would be afraid of the puppet

[4] Since we assumed the new method of Experiment 2 would generally make the tasks more demanding, competence more difficult to detect and effect sized thus smaller, we reasoned that a significantly bigger sample than in Experiment 1 would be appropriate.

Figure 4. Example of the boxes used in Exp. 2.

Procedure. For warm-up games and the subsequent test trials parent and child were asked to take a seat at a table. The parent had been instructed not to interfere by saying or showing anything to the child. Each session began by the child's playing two warm-up games (puzzle, picture book) with both experimenters. In contrast to Exp. 1 no warm-up trial with a hiding procedure was administered because this was considered to possibly differentially simplify correct helping only in the non-aspectual tasks as opposed to the aspectual tasks (see Appendix 2 for details).

Test trials

(1) Non-aspectual task. During the last warm-up game, E1 left the table and went behind curtains, where she was not visible for the child. Now, E2 told the child that she forgot something and left the room. E1 appeared from behind the curtains, greeted the child and sat down at the table. Accidentally, she placed a tissue on one side of the table. Then she took two soft toys out of a bag, showed them to the child, placed them on the table and returned behind the curtains. At this point, E2 re-entered the room and discovered the objects and picked them up excitedly. Then she looked disappointed, announcing that she

forgot something again and had to leave. She looked around and placed one box at one side on the table. Then she told the child that she would put the objects into the box for the time of her absence and put them into the upper-side opening. The following sequence varied between conditions (see Fig. 5):

- FB condition: E2 left and E1 reappeared from behind the curtains. She greeted the child and proposed to 'play a trick' on E2, took one of the objects out of the box and placed it under the tissue (in such a way that it was neither visible that this object was under the tissue nor whether there was any object there), giggling and gesturing and whispering "Shh".
- TB condition: E1 performed the very same action of taking the out of the box and placing it under the tissue as in the FB condition, but the crucial difference was that this was witnessed by E2 who did not leave the room. Rather, E2 stayed and E1 appeared from behind the curtains, telling E2 she wanted to show her something first. Then, but saying "Look, what I do now!", while E2 observed this closely and said "Ah I see! Yes!" Then E1 returned behind the curtains and E2 told the child she now really had to leave and left the room for several seconds.

By her return (in both conditions), E2 approached the table, said "Okay" and reached into the box though the upper side opening. She took the remaining object out and put it beside the box and began to search in the box again (moving her hand inside the box, trying to look inside, saying "Hmm, Eh? I don't understand… but where is…."). As in Exp. 1, this event is referred to as the 'signal'. If the child did not react, E2 sat down and expressed disappointment: "Hm. Oh no!" If still no reaction followed, she asked "Can you help me?" The task was repeated once with a different box and two different target toys if E2 was not sure whether the child's reaction to the signal could clearly be assigned to one of the two

central categories: helping to *search in the box* (e.g. by pulling one of the door grips) or *giving the object* under the tissue to E2.

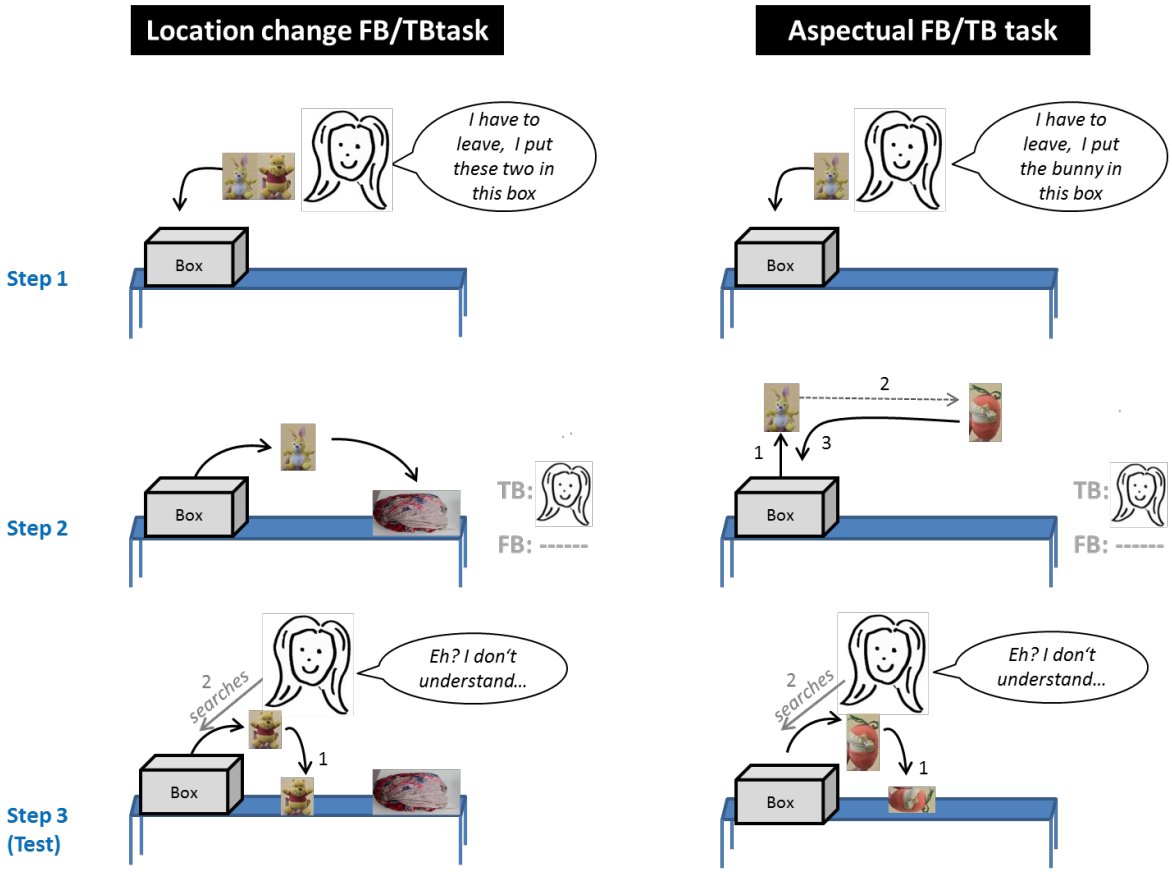


Figure 5. Schematic event sequences of the location and aspectual FB/TB tasks in Exp.. 2.

(2) Aspectual task. The procedure of the aspectual task was exactly the same as for the non-aspectual task, except for the following differences: When E1 appeared from behind the curtains for the first time, she introduced only one reversible mini soft toy to the child and – in contrast to the non-aspectual task - transformed it to its 2nd aspect saying: "Look! The X is also an Y!" (e.g., bunny and carrot). She did that in a sneaky way, whispering "Shh" and told the child: "'E2 does not know that, right?", again, to ensure that the child

would not assume E1 to know all aspects of the object. The child was then asked to transform the object into its original state: “Can you make it so that it is a bunny again?” If the child was not able or not willing, E1 helped and both transformed the object together. Then E1 placed the object in its original aspect on the table and returned behind the curtains. The following sequence was the same as in the non-aspectual task, despite the fact that E2 – upon her return - discovered only one object and put this one into the box before she left again in the FB condition. But instead of changing the location of one object, E1 appeared from behind the curtains, took the object out of the box, transformed it to its 2nd aspect and put it back into the box. This was done either in the absence (FB) or presence (TB) of E2 and in exactly the same sneaky/not sneaky way as in the non-aspectual task. When E2 reentered the room she behaved exactly the same as in the non-aspectual task.

Table 2. Schmematic procedure of the tasks in Exp. 2

	Non-aspectual, change-of-location task	Aspectual task
1	**Setting** E2 (protagonist) sat vis á vis the child on the lap of her parent at a table. During two warm-up games (puzzle, picture book) E1 also sat on the short side of the table. During the last warm-up game E2 left the table and went behind curtains where she was not visible for E2 and the child.	
2	**Introduction of two single-aspect target objects** E2 announces: "Oh no, I forgot something! I have to leave for short." and leaves. E1 appears from behind the curtains, greets the child and sits down at E2's chair. Accidently, she places a tissue on one side of the table. Then she takes two soft toys out of a bag saying "Look what I got here, what's that? Yes, a bunny! And look what else I have! A bear! I put them both here. Okay" E1 leaves behind the curtains. E2 re-enters the room and sees the toys: "Wow! A bunny, it's great! And a bear! It's great too!" Then she announces: "Oh no! I again forgot sth., I have to leave for short. I put these two in this box" (she takes one box from under the table and places it at the other side than the tissue).	**Introduction of the dual aspect objects** E2 announces: "Oh no, I forgot something! I have to leave for short." and leaves. E1 appears from behind the curtains, greets the child and sits down at E2's chair. She takes one soft toys out of a bag saying "Look what I got here, what's that? Yes, a bunny! A bunny! And now I show you sth.: the bunny is also a carrot. Where is the bunny now? Can you make it so that it's a bunny again?" The child transforms the object, if necessary with the help of E1. E1 emphasizes: "And Susi doesn't know that, right? Shh! I put the bunny here. Okay" E1 leaves behind the curtains. E2 re-enters the room and sees the toy: "Wow! A bunny, it's great!" Then she announces: "Oh no! I again forgot sth., I have to leave for short. I put the bunny in this box" (she takes one box from under the table and places it at one side of the table).
3	**Location-change** **False belief:** E2 leaves. E1 appears from behind the curtains and says "Shhh! Hello [name of child], now we play a trick on [name of E2] Look, what I do now. Okay. Hihi. Shh!" She takes one toy out of the box and places it under the tissue in a sneaky way emphasizing at the end: "And Susi can't see that, right? Shh!" E1 leaves behind the curtains. **True Belief:** E1 appears from behind the curtains and says "Look [names of E2 and child] what I do now! Like this and like this!" She takes one toy out of the box and places it under the tissue, thereby alternating gaze between the child and E2. E2 observes this event closely saying "Ah I see! Yes!" E1 leaves behind the curtains again. E2 says "But now I have to leave, I'll be right back" and leaves.	**Aspect-change** **False belief:** E2 leaves. E1 appears from behind the curtains and says "Shh! Hello [name of child], now we play a trick on [name of E2] Look, what I do now. Okay. Hihi. Shh!" She takes the toy out of the box, transforms it into its 2nd aspect in a sneaky way and puts it back emphasizing at the end: "And Susi can't see that, right? Shh!" E1 leaves behind the curtains. **True Belief:** E1 appears from behind the curtains and says "Look [names of E2 and child] what I do now! Like this and like this!" She takes the toy out of the box, transforms it into its 2nd aspect in a sneaky way and puts it back, thereby alternating gaze between the child and E2. E2 observes this event closely saying "Ah I see! Yes!" E1 leaves behind the curtains again. E2 says "But now I have to leave, I'll be right back" and leaves.
4	**The signal** E2 re-enters the room, stands at the table, says "Okay!" and grasps into the box. She takes one object out of the box and puts it beside the box (the object now laying at the center of the table). Then she moves her hand further in the box, as if searching for sth. saying "Hm? Eh? I don't understand. But where is?" If the child had not reacted until now, E2 sits down at her chair says "Hmm", looks at the child and asks "Can you help me?" She repeats this once if the child does not react. If the child helps to search in the box or gives the object under the tissue (non-aspectual task)/the object at the table (aspectual task), E2 thanks the child. If the child shows no reaction or a reaction different from these three categories, E2 says "Come on we play another game and removes the box and object to begin a second trial with a different box and different target objects.	

Coding. *A*s in Exp. 1, the first reaction to the signal was coded and it was determined whether it fell into one of the two central categories:

- help to search in the box: if the child, in response to the signal, touched the box, pulled a grip, opened one of the doors, tried to look inside etc.
- give the object (under the tissue in the non-aspectual task/on the table in the aspectual task): if the child, in response to the signal, took the object under the tissue or on the table or pointed to the object.

Compared to Exp. 1, the setting of Exp. 2 allowed for a larger range of behaviors which could not clearly be coded as one of the above categories and were cautiously assigned to one of the following two other categories:

- ambiguous: this coding applied

  (i) if the child showed a behavior that clearly was a reaction to the signal (i.e. the child's behavior began after E2's searching behavior and not already before) but which did not fit either of the two above categories (e.g. the child did perform both above behaviors, for example pulling on the tissue/taking the not-referred-to object (aspectual tasks) and touching the box simultaneously, and it was not clear which reaction was a response to the searching behavior of E2) and

  (ii) if the child's behavior was not clearly a reaction to the signal (i.e. the child referenced the hidden/transformed object or began to explore the box already at a time point before E2 started to search further in the box and this behavior merged with the child's behavior after the signal).

  (iii) if the child produced a response to the signal which did not fit either of the two above categories (e.g. giving the first object taken out in the non-aspectual task)

which could not certainly be assigned to one of the categories were coded as ambiguous.

- invalid: this coding was given if the child did not show any reaction, if the child left the scene during the trial or did not pay attention to the event sequence.

A third observer who was blind to the experimental condition and hypotheses coded a 20% sample of 27 randomly selected tapes (36 trials) for reliability which was κ = .96 for the four-category-coding (open box 1, open box 2/give object, invalid, ambiguous).

Results

For 28 children (10 in the Loc_FB, 6 in the Loc_TB, 3 in the Asp_FB and 9 in the Asp_TB condition) both trials were coded as invalid because the child showed no reaction to the signal (*N* = 20), was inattentive (*N* = 3), or afraid (*N* = 1), or because of experimental error (*N* = 4). Data from 109 children were thus included in the final analysis. Mean age was *M* = 26;2 (*SD* = 2.14, 11 girls, 15 boys) in the non-aspectual FB condition, *M* = 26;2 (*SD* = 1.73, 11 girls, 14 boys) in the non-aspectual TB condition, *M* = 26;1 (*SD* = 1.95, 14 girls, 13 boys) in the aspectual FB and *M* = 25;3 (*SD* = 1.87, 14 girls, 17 boys) in the aspectual TB condition. There was no significant age difference between conditions ($F(3) = 1.24$, $p = .300$).

Trials coded as ambiguous or invalid. Of the 109 included children, the first trial was coded as invalid in 13 cases (3 in the Loc_FB, 3 in the Loc_TB, 5 in the Asp_FB and 2 in the Asp_TB condition) due to experimental error (*N* = 2), because a parent or sibling interfered (*N* = 4), the child did not react to the signal (*N* = 4) or was inattentive *(N* = 3*)*. Of these children 2 (1 in the Loc_FB and 1 in the Asp_TB condition) received an ambiguous code in the second trial. For 6 children both trials were coded as ambiguous (1 in the Loc_FB, 3 in

the Asp_FB and 2 in the Asp_TB condition). 22 children (5 in the Loc_FB, 1 in the Loc_TB, 9 in the Asp_FB and 7 in the Asp_TB condition) received an ambiguous code in the first trial. Of those, 5 children (1 in Loc_FB, 1 in Asp_FB, 3 in Asp_TB) received an invalid coding in the second trial. All in all there were 51 children in the non-aspectual conditions and 58 children in the aspectual conditions who received at least one valid (including ambiguous) trial, of which 48 children per condition had at least one valid unambiguous trial in each case.

Main analyses. As in Exp. 1, for the main analysis the first valid and unambiguous trial of the 96 children to which this applied was used (see Figure 6). Invalid and ambiguous trials were replaced by the second trial if this was a valid trial. Children's helping behavior differed significantly between FB and TB condition in the non-aspectual task (N = 48, p = .018, Φ_{Cramer} = .354, one-tailed Fisher's exact test). In the aspectual task helping behavior did not differ significantly between FB and TB condition (N = 48, p = .207, Φ_{Cramer} = .163, one-tailed Fisher's exact test).

Complementary analysis. When comparing the non-aspectual and aspectual tasks for FB and TB, respectively, results revealed that children tended to perform differently in the location than in the aspectual TB tasks (opening box 1 more often in the location than in the aspectual TB tasks, p = .049, one-tailed Fisher's exact test). No differences were found for the two FB tasks (p = .379).

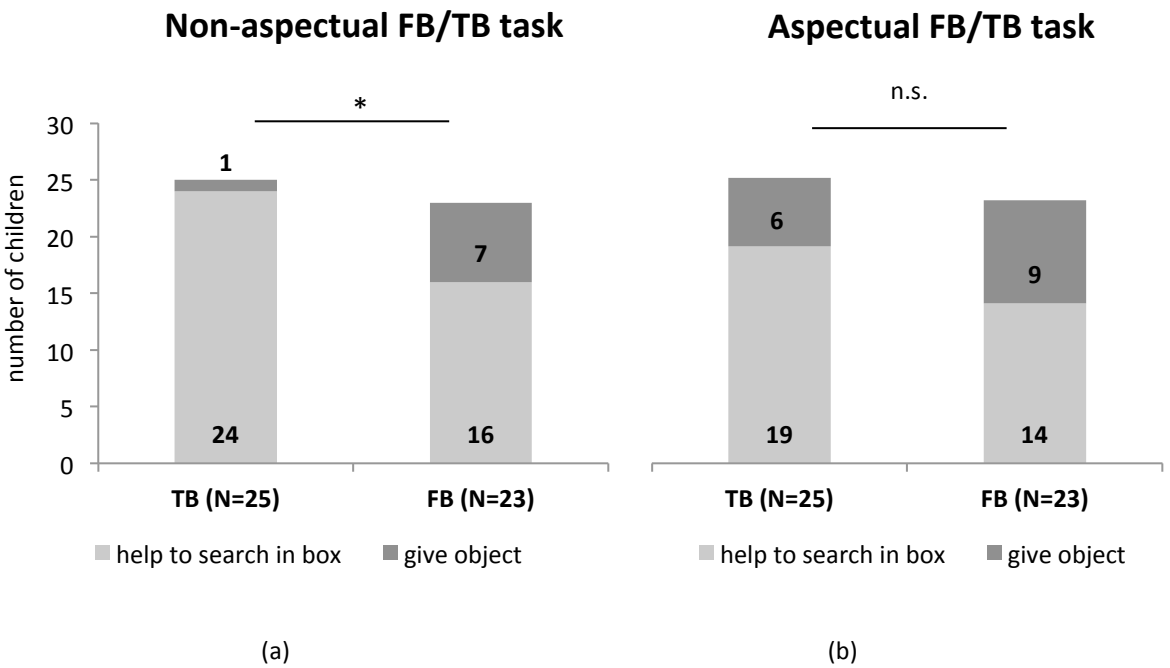


Figure 6. Number of children showing the different kinds of helping behavior in the first valid and unambiguous trial (a) the non-aspectual tasks, (b) the aspectual task. [$^{*}p < .05$]

Analyses including ambiguous trials. The results of the main analysis were confirmed in analyses also including ambiguous responses in the analyses as a third category (see Table 3): the difference between conditions in the non-aspectual task remained significant (N = 51, p = .004, V_{Cramer} = .461, two-tailed Fisher's exact test), whereas in the aspectual task FB and TB conditions did not differ significantly (N = 58, p = .228, V_{Cramer} = .207, two-tailed Fisher's exact test).

Table 3. *Helping behavior in the first valid trial including ambiguous responses.*

	Non-aspectual task		Aspectual task	
	TB	FB	TB	FB
help to search in box	23	13	18	11
give object	1	7	5	7
ambiguous response	1	6	7	10