The knowledge ("true belief") error in 4- to 6-year-old children: When are agents aware of what they have in view?

Michael Huemer\textsuperscript{a,b,}\textsuperscript{*}, Lara M. Schröder\textsuperscript{a,b}, Sarah J. Leikard\textsuperscript{a,b}, Sara Gruber\textsuperscript{a}, Anna Mangstl\textsuperscript{a}, Josef Perner\textsuperscript{a,b}

\textsuperscript{a} Centre for Cognitive Neuroscience, University of Salzburg, Hellbrunnerstraße 34, 5020 Salzburg, Austria
\textsuperscript{b} Department of Psychology, University of Salzburg, Hellbrunnerstraße 34, 5020 Salzburg, Austria

\section*{A R T I C L E   I N F O}

\textbf{Keywords:}
Theory of mind
False belief
True belief
Knowledge

\section*{A B S T R A C T}

The standard view on explicit theory of mind development holds that children around the age of 4 years start to ascribe beliefs to themselves and others, typically tested with false belief (FB) tasks. The present study (\(N = 95\), 53 female, 41 male, Austrian, 41 to 80 months) systematically investigated the puzzling phenomenon that FB achievers (FB+) fail knowledge (often subsumed under "true belief") tasks: Despite the story protagonist witnessing the displacement of an object these children predict that the protagonist will look for it in its original location. We replicate this result in Experiment 1. Interestingly, some of our children indicated uncertainty about the protagonist’s awareness of the relevant event. Thus, in Experiment 2 a new active watching condition was designed to help children understand that the protagonist attended to the critical event. This practically eradicated the knowledge error. Experiment 3 successfully replicated these results. Implications for existing explanations, perceptual access reasoning (PAR, Fabricius et al., 2010) and pragmatic difficulties (Oktay-Gür & Rakoczy, 2017) are discussed.

1. Introduction

Once upon a time – nearly 40 years ago –, Wimmer and Perner (1983) introduced the false belief (FB) task, which became the standard measure for assessing someone’s ability to ascribe beliefs to themselves and others. For example, Lisa puts her teddy into the red box and leaves for the playground. In her absence her brother moves the teddy to the yellow box, and leaves. Lisa returns, and the child is asked where Lisa will look for her teddy. The standard view holds that children start to master FB tasks around the age of 4 (Wellman, Cross, & Watson, 2001). This goes along with the emerging competence in a variety of tasks requiring the ability to distinguish different perspectives (Doherty & Perner, 2020; Perner & Roessler, 2012), suggesting a conceptual change to metarepresentational thinking (Perner, 1991).

Younger children typically fail this sort of task but show no problems with structurally similar tasks where Lisa’s belief about the location of the teddy is correct (e.g., Fabricius et al., 2010; Oktay-Gür & Rakocy, 2017, Experiments 1 & 2; see also Wellman et al., 2001). Two versions of such tasks can be found in the literature, to which we refer as knowledge and true belief (TB) tasks. In knowledge tasks, Lisa witnesses the transfer of the teddy and therefore knows where she can get the teddy. To keep these tasks structurally as similar as possible to FB tasks, Lisa is absent either before or after the teddy is moved (e.g., Fabricius et al., 2010, Experiment 2; Friedman, Griffin, Brownell, & Winner, 2003, Experiment 2; Oktay-Gür & Rakoczy, 2017; Rakoczy & Oktay-Gür, 2020). In TB tasks, the teddy is moved but put back in the original location in Lisa’s absence. For instance, Lisa who knows of the teddy in the red box does not witness her brother playing with it and then putting it back in the red box (Fabricius et al., 2010, Experiment 1). In this situation, Lisa has a justified true belief about the teddy being in the red box but does not have proper knowledge of it. Scenarios like this resemble “Gettier cases” introduced in epistemology to demonstrate that knowledge cannot be equated with justified true belief (Gettier, 1963). We emphasize this distinction between knowledge and TB tasks because the proposed accounts of the observed phenomena apply to these two tasks in different ways.\footnote{Corresponding author at: Centre for Cognitive Neuroscience, University of Salzburg, Hellbrunnerstraße 34, 5020 Salzburg, Austria.
E-mail address: michael.huemer2@stud.sbg.ac.at (M. Huemer).}

\footnote{Here, we only use knowledge tasks, which have been the focus of interest in almost all relevant studies (for an overview, see Fabricius et al., 2021). Our results are not generalizable to the rarely studied Gettier-type TB task.}

\footnote{https://doi.org/10.1016/j.cognition.2022.105255
Received 13 November 2021; Received in revised form 9 August 2022; Accepted 10 August 2022
Available online 8 September 2022
0010-0277/© 2022 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).}
1.1. A developmental puzzle, and how to explain it

Recent evidence shows that while younger children typically fail FB tasks (FB-children) but show no problems when Lisa entertains an accurate belief, children who pass FB tasks (FB+ children) start to fail TB tasks (Fabricius et al., 2010, Experiment 1) and knowledge tasks (e.g., Fabricius et al., 2010, Experiment 2; Fabricius et al., 2021; Friedman et al., 2003; Oktay-Gür & Rakoczy, 2017; Rakoczy & Oktay-Gür, 2020). For an overview, see the recent meta-analysis on children's performance in TB and knowledge tasks provided by Fabricius et al. (2021).

There are two principal ways to explain this paradoxical phenomenon – competence vs. performance limitations. Proponents of competence limitations argue that children do not have a fully developed concept of belief by age 4 and use simple heuristics instead. For instance, perceptual access reasoning (PAR) posits a U-shaped developmental pattern of performance on TB and knowledge tasks in preschool children. Three-year-olds succeed in these tasks, basing their answers on reality. Four- and 5-year-olds fail because they reason from Lisa's lacking perceptual access that she will act incorrectly. From 6 years on children use belief reasoning and therefore master TB and knowledge tasks (Fabricius et al., 2010; Fabricius et al., 2021; Fabricius & Imbens-Bailey, 2000; Fabricius & Khalil, 2003; Hedger, 2016; Hedger & Fabricius, 2011). In contrast, performance limitation explanations hold that 4-year-old children have the conceptual prerequisites to master belief tasks but fail because of extraneous performance factors in knowledge tasks. The pragmatic account (Oktay-Gür & Rakoczy, 2017; Rakoczy & Oktay-Gür, 2020) attributes children's problems to trivial test questions, which refer to where Lisa thinks her teddy is, or to predict Lisa's action when she shares common ground and could not possibly err.

According to the PAR account (Fabricius et al., 2010; Fabricius et al., 2021; Fabricius & Imbens-Bailey, 2000; Fabricius & Khalil, 2003; Hedger, 2016; Hedger & Fabricius, 2011), there are three developmental stages. In the first, until age 4, children respond to the question where Lisa will look for her teddy using their own knowledge about reality. This yields wrong answers in FB tasks and correct answers in TB and knowledge tasks. In the second stage, children of age 4 and 5 base their reasoning in these tasks on agents' perceptual access (this idea originates from Ruffman, 1996). Children using PAR assume that agents without perceptual access lack knowledge, and will therefore act in the wrong way (“doesn’t see ⇒ doesn’t know ⇒ gets it wrong”), while agents with perceptual access possess knowledge, which enables them to get things right (“does see ⇒ does know ⇒ gets it right”). No ascription of beliefs is involved in this way of reasoning. In the third stage, from 6 years on, children engage in belief reasoning leading to correct answers in FB, TB and knowledge tasks.

In the PAR account, there is no presumption of children having a mature concept of knowing as a representational mental state that one maintains over time. Instead, knowing in PAR is caused by what one has and knows about the location of the teddy, and thus will get it right. A direct comparison of a return with a no-return condition showed that 4.5-year-olds found the return condition (82% wrong responses) more difficult than the no-return condition (53% wrong) (Friedman et al., 2003, Experiment 2).

This can account for the pattern that 4- and 5-year-olds respond correctly in FB tasks and incorrectly in TB tasks. In TB scenarios, the teddy is returned to its original location after it has been moved in the absence of Lisa, her perceptual access has to be established anew when she returns. On her return, she can’t see the teddy and therefore children predict she will look for it in the wrong place. Children's difficulties in knowledge scenarios, where Lisa witnesses the displacement of the teddy and leaves the scene after the transfer, can be accounted for in the same way: Children think that when Lisa returns she is in a new situation where she cannot see her teddy, thus she does not know where to find her teddy, and will therefore act incorrectly even though she has seen the transfer. In contrast, in the knowledge stories where Lisa does not leave the room after the transfer, children will not perceive her as being in a new situation and, hence, they use their initial analysis that Lisa had perceptual access, knows about the location of the teddy, and thus acts correctly (e.g., Fabricius et al., 2010; Fabricius et al., 2021; Hedger & Fabricius, 2011). According to the PAR account, the sequence of the three developmental stages – reality reasoning, PAR, belief reasoning – would lead to a U-shaped pattern of correct responses from before 4 years to after 6 years. This pattern should occur in TB tasks and in those knowledge tasks where Lisa leaves and returns after the transfer.

An alternative to the conceptual deficit explanation is the pragmatic account of Rakoczy and Oktay-Gür (Rakoczy, 2022; Rakoczy & Oktay-Gür, 2020). The critical questions asked are not sincere regular questions to find out the answer but are second-order test questions to find out whether the listener knows the answer (Perner, Leekam, & Wimmer, 1984). The point of this kind of question may well escape young children (Siegal, 1999). If the question is understood as a sincere question, then children might get confused when they are asked where Lisa thinks her teddy is, or to predict Lisa's action in a situation where she shares common ground, without the possibility of erring. Since the answer to the test question seems too obvious, especially because it asks about behavior or beliefs which could hardly be otherwise, children may ask themselves whether this is what the experimenter could really have meant, or they may think they must have missed something, and respond with the answer different from the obvious one.

This may make the test questions in knowledge tasks confusing for 4- to 7-year-old children with immature pragmatic abilities. Since emerging pragmatic expertise depends on developing theory of mind competence (e.g., Fernández, 2013; Happe, 1993; Winner & Gardner, 1993), children who have come to master FB tasks should be particularly affected by pragmatic confusion in these tasks. This would lead to a negative correlation between FB and knowledge tasks, which has been empirically confirmed in several studies (Rakoczy, 2022; Rakoczy & Oktay-Gür, 2020). No such pragmatic confusion occurs in 3-year-old children. Since they have not yet acquired a concept of belief and their pragmatic abilities are not yet sufficiently developed, they answer the test question based on reality anyway. Children from around 8 to 10 years of age are then able to overcome the difficulties of the younger children. It seems that these children can resolve the pragmatic confusion of knowledge tasks when reaching a higher level of theory of mind beyond second-order (Schidlok, Proft, & Rakoczy, 2022). The pragmatic account thus would predict a U-shaped development of performance in knowledge tasks, which found supporting evidence in their experiments (Oktay-Gür & Rakoczy, 2017; Rakoczy & Oktay-Gür, 2020).

1.2. Previous findings

To test the PAR account against their pragmatic account, Oktay-Gür and Rakoczy (2017, Experiments 1 and 2; see also Rakoczy & Oktay-Gür, 2020. Experiments 2 and 4) used knowledge tasks in which the child was asked where Lisa will look for her teddy immediately after the teddy was
displaced. The PAR account predicts a knowledge error only in scenarios where Lisa leaves after the object has been transferred, because only there would children see Lisa in a new situation on her return and use PAR (‘doesn’t see ⇒ doesn’t know ⇒ gets it wrong’). Since she cannot see the object children predict she will look in the empty box. In contrast, when Lisa does not leave after the transfer but stays around, children will not perceive this as a new situation and therefore ascribe perceptual access to Lisa and predict that she will go to the box with the teddy.

Oktay-Gür and Rakoczy (2017; 2020) found that 4- to 8-year-olds showed the knowledge error even if Lisa did not leave after the transfer. Fabricius et al. (2021, p. 36) have pointed out that in Experiments 1 and 2 of Oktay-Gür and Rakoczy (2017), a “Where will she look first for her…” phrasing was used in the test question. The word “first” in the test question may encourage use of PAR ‘by implying that the first search would be wrong’ (Fabricius et al., 2010, p. 1414). This argument is supported by findings of Rai and Mitchell (2004), who, in a knowledge task in which Lisa stays after the teddy’s displacement, found that 5-year-olds were 85% correct when asked a “will do next” test question, but were only 25% correct when given a “look first” test question. However, this objection does not apply to Rakoczy and Oktay-Gür (2020, Experiment 2 and Experiment 4, TB first condition), where a knowledge error was found when children were asked where Lisa believed/thought her teddy was; contrary to what the PAR account predicts.

In contrast, these results fit Rakoczy and Oktay-Gür’s pragmatic account quite well as all three factors – academic test questions, trivial questions, and being about an agent’s belief – are present in their knowledge tasks. Significantly, the triviality of the test question seems to be even stronger in those tasks where children are asked where Lisa will look for the teddy directly after the object has been transferred. Children may find this situation particularly puzzling, having just witnessed how Lisa has seen the teddy being put into the new box and that everybody else around has noticed that Lisa has seen where the teddy has been placed. In this situation, there is simply no possibility that anyone could be mistaken. The question where Lisa will look for her teddy now seems very odd. Conversely, children’s problems with the knowledge task might be eased if Lisa leaves the scene after the transfer and the test question is asked when she returns. The time delay between the transfer and the test question may make the answer seem less trivial. The experimenter or Lisa might have forgotten where the teddy had been put.

Thus, the triviality explanation predicts that children will more likely be making errors when the test question is asked immediately after the transfer than when it is asked after a delay, due to Lisa leaving and returning. The PAR account, by contrast, predicts fewer errors for immediate than for delayed questioning. See Fig. 1 for an overview of how the studies reported here manipulated the timing of the test question in relation to the transfer.

Whether there is a difference between delayed and immediate asking of the test question, and if so, in which direction a possible difference goes, can be hardly answered by the existing literature. In their meta-analysis, Fabricius et al. (2021) compared, among others, conditions in which Lisa stays or leaves without returning after the teddy’s displacement with conditions in which Lisa leaves and returns after transfer. For this comparison, those conditions are particularly interesting in which the factors highlighting and first look formulation of the test question do not occur, as both factors have an additional influence on the response behavior in knowledge tasks. Highlighting refers to the child being reminded immediately before the test question that Lisa saw the teddy bear being placed in the second box (where it actually ends up). This highlights Lisa’s prior perceptual access, and interferes with children’s reasoning about Lisa’s current lack of perceptual access (Fabricius et al., 2021; Friedman et al., 2003). The issue of the look first formulation we have already discussed above. Fabricius et al. (2021, p. 46) report that children’s performance (estimated at 54 months) in the stay/no-return conditions was higher (94% correct) than in the five return conditions (78% correct).

This result is in line with the PAR account’s prediction that children are more likely to respond incorrectly in return conditions. According to the PAR account, in return conditions children consider the situation to be new and, hence, reapply PAR, leading them to conclude from Lisa’s lack of perceptual access that she will err, which in turn leads to incorrect responses to the test question. It is worth noting, however, that children in the five return conditions in the meta-analysis are clearly more likely to answer correctly (78% correct) than not. Without knowing whether these children were succeeding in the FB task or not, it is not possible to decide whether these findings are inconsistent with the PAR account, or whether among these children there are many whose answers reflect reality-based reasoning.

More recent studies, not included in the meta-analysis, found knowledge errors both in delayed conditions, where Lisa leaves and returns after the teddy’s displacement (Fabricius et al., 2021, Experiment Section VII), and in immediate conditions, where Lisa stays after the transfer (Rakoczy & Oktay-Gür, 2020, Experiment 2 and Experiment 4, TB first condition; Schidkelo, Schönemann, Rakoczy, & Proft, 2021, Experiment 2). Overall, the available empirical evidence seems to be incoherent and complicated. Part of the issue is the fact of only having comparisons of conditions across different studies with different procedures. We now put both the PAR account and the triviality hypothesis (based on Rakoczy and Oktay-Gür’s pragmatic account) to a direct test by experimentally contrasting an immediate with a delayed condition.

2. Experiment 1

We investigated whether the timing of the test question could have an influence on FB+ children’s performance in the knowledge task by varying the time between the transfer (witnessed by Lisa) and the test question. According to the triviality hypothesis, if the test question is posed immediately after the transfer, then the answer appears to be more obvious and trivial resulting in more knowledge errors than if Lisa leaves after the transfer, returns, and then the test question is posed. Hence, the prediction is that FB+ children have less difficulties with the delayed compared to the immediate condition.

According to the PAR account the opposite pattern should emerge. Only in the delayed condition, where Lisa leaves after the transfer and the question is posed after her return, Lisa finds herself in a new situation, where she does not see, therefore does not know where the object is, and consequently will act wrongly. This should not happen in the immediate condition where Lisa stays in the same situation. Thus, the prediction is that FB+ children should give mostly correct answers in the immediate condition but wrong answers in the delayed condition. For FB- children, both the PAR account and the triviality hypothesis predict they would give correct responses regardless of the timing of the test question. Neither hypothesis makes any claims about the timing of the questions for these children. Since the PAR and the triviality hypothesis make different predictions for FB+ children but not for FB- children, the focus of our interest was on the FB+ children.

2.1. Method

2.1.1. Participants

Forty children (24 female, 16 male) aged 48 months to 79 months (M = 64.4 months, SD = 7.15) were included in the final sample. Children were recruited from a kindergarten in Linz (Austria). Parents gave written consent to their child’s participation.

---

Fabricius et al. (2021, pp. 450) report that they tested whether stay and no-return conditions differed, but found no effects. They, thus, collapsed the two conditions for the subsequent analysis.
2.1.2. Design

Each child faced one standard change-of-location FB and two knowledge tasks of the same structure as the FB task except that the protagonist was present when the object was transferred. The two knowledge tasks differed with respect to the timing of object transfer and the posing of the test question; the test question followed either immediately or with a delay.

In all three experiments of the current study, the three belief tasks were used together with five tasks of another study measuring higher-order belief understanding in children. The order of the eight tasks was counterbalanced for all three experiments as a block with the restriction that the two knowledge tasks could never follow each other. There were no detectable order effects in any of the three experiments (Kruskal-Wallis Test: all $p \geq .226$).

2.1.3. Procedure

2.1.3.1. False belief tasks. The FB tasks were modeled after Wimmer and Perner’s (1983) original Maxi task. We used animated stories (also used in Perner, Mauer, & Hildenbrand, 2011; Priewasser, Fowles, Schweller, & Perner, 2020; Priewasser, Roessler, & Perner, 2013) which were presented on a laptop. Three stories with different protagonists and material were used. The procedure (see Fig. 2) is here described, using the “Lisa”-story as an example.

Lisa put her teddy in the red box and left the scene. In her absence, her brother took the teddy out of the red box and moved it to the yellow box. After Lisa’s brother left the scene, the child was asked the following control questions:

- **Memory question:** “Where did Lisa put the teddy in the beginning?”
- **Reality question:** “Where is the teddy now?”

Then Lisa returned to the scene and wanted to play with her teddy. The child was then asked the test question:

- **Test question:** “Where is Lisa going to look for her teddy?”

2.1.3.2. Knowledge tasks. The basic scheme of both knowledge tasks was adopted from the FB task described above, with the only difference being that Lisa was present during the transfer. The two conditions – immediate and delayed – differed with regard to the time elapsed between the transfer and the test questions. In the immediate condition, Lisa left and returned to the scene before the transfer, and the test question followed immediately after the transfer. In the delayed condition, Lisa left and returned after the transfer, leading to a delay between transfer and the test question (see Fig. 2, knowledge passive, middle row).

2.2. Results and discussion

Data of all three experiments are available on doi: 10.17632/3jk5pmnc3.1.

2.2.1. FB task

Children performed perfectly on the memory and reality question in the FB task. Four children answered the test question incorrectly and were classified as FB- (mean age = 62 months, SD = 5.0). Thirty-six children passed the test question and were classified as FB+ (mean age = 65 months, SD = 7.4).

2.2.2. Knowledge tasks

All memory and reality questions were answered correctly. Fig. 3 (upper left panel) shows percentages of children choosing box 2 (the object’s current location) in the test question in both the immediate and the delayed condition of the knowledge tasks. FB- children predicted that Lisa will look for her teddy in box 2 in both the immediate and delayed condition at 100%, whereas FB+ children chose box 2 in the immediate condition at 36% (not different from chance, Binomial test: $p = .132$) and in the delayed condition at 25% (different from chance, Binomial test: $p = .004$). For FB+ children there was no reliable difference between the two conditions (Related-samples Wilcoxon signed-rank test: $Z = 2, p = .125, r = .33$). This provides no evidence for the triviality hypothesis; the small, nonsignificant difference happened to go in the opposite direction than expected under this hypothesis. The data...
also do not support the PAR account. The observed response pattern does not match the prediction that FB+ children have problems only with the delayed condition but give mainly correct answers in the immediate condition. Interestingly, some children asked the experimenter during the test sessions whether Lisa (or Maxi) had seen the transfer. This indicates that at least some children were not sure if Lisa was aware of the transfer despite being present during the whole transfer with nothing hindered her sight. Certainly, doubt that Lisa was aware of the transfer is not out of the question; she stands there passively while her brother moves her teddy without her permission. So, if some children did not interpret the protagonist as witnessing the critical event – the transfer of the object – the knowledge task would be for them the same as the FB tasks, giving wrong answers for good reasons. This would be the case for both conditions. Interestingly, with children doubting whether Lisa had noticed the transfer, the data become consistent with predictions from the PAR account. The PAR account then predicts that children will think that Lisa will go to the empty box does not show a lack of understanding belief, but reflects a difference in how they interpret the stimuli. With a new procedure in Experiment 2 we tried to eradicate this misinterpretation. In the new active watching version of the knowledge task, Lisa accompanies her brother during the transfer and watches what he is doing closely. In this condition, Lisa is not directly involved in the transfer of the teddy, for instance, by carrying it; she only watches the teddy being displaced, made evident by her walking alongside her brother. As a manipulation check, we asked children whether Lisa could see the transfer. This question also reinforced the manipulation by drawing children’s attention to the importance of Lisa’s access to this information. Lisa’s active watching and the additional question are intended to serve the purpose of bringing children’s interpretation of events in line with how we had intended them to be interpreted.

It is important to note that the see-question in the active watching condition is asked before Lisa leaves and returns. According to the PAR
account, Lisa’s return “prompts children to see [her] as being in a new situation, and to use PAR about this new situation [...] without any reference to the prior situation” (Hedger & Fabricius, 2011, p. 433). Children do not attribute any memory to others once the situation leaves and returns in our active watching condition is not the same as “highlighting” (Fabricius et al., 2021) because it is asked before the final situation in which the test question is asked. Highlighting consists of reminding children of Lisa’s seeing the transfer immediately before the test question. In fact, a question similar to our see-question has also been asked by Fabricius et al. (2010; see also Fabricius et al., 2021); they asked their question after Lisa had left but before she returned.

On our misinterpretation hypothesis we expect FB+ children to show markedly fewer knowledge errors than in Experiment 1. In case knowledge errors persist we also manipulated the delay between transfer and test question again, as in Experiment 1. This will provide another potential test of the triviality and the PAR hypothesis with the active watching versions.

3.1. Method

3.1.1. Participants

Nineteen children (9 female, 10 male) aged 47 months to 79 months (M = 66.0 months, SD = 10.63) were included in the final sample. Children were recruited from two kindergartens in and around Linz (Austria). Parents were informed beforehand and gave written consent to their child’s participation.

3.1.2. Design

Each child was presented with one FB and two knowledge tasks. The two knowledge tasks had the same structure as the FB task, except that the protagonist was present during the object’s displacement. Again, the two knowledge tasks differed with respect to the timing of object transfer and the positing of the test question; the test question followed either immediately or with a delay.

3.1.3. Procedure

The FB tasks were the same as in Experiment 1. The following modifications were made to the procedure of both knowledge conditions of Experiment 1 (see Fig. 2, knowledge active, bottom row): Lisa actively followed her brother Tom displacing the teddy instead of passively standing in the middle of the scene; and, immediately after the transfer, children were asked whether Lisa had seen the transfer:

See-question: “Did Lisa see that Tom put the teddy in the yellow box?”

3.2. Results and discussion

3.2.1. FB task

Children’s performance on the memory and reality questions was perfect. Three children failed the test question and were classified as FB- children (mean age = 66 months, SD = 4.5). Sixteen children answered the test question correctly and were classified as FB+ children (mean age = 66 months, SD = 11.5).

3.2.2. Knowledge tasks

Children performed perfectly on the memory and reality questions. All children responded with “yes” to the see-question. Percentages of children choosing box 2 in the test question are depicted in Fig. 3 (upper right panel). As in Experiment 1, FB- children responded with “box 2” in both the immediate and delayed condition at 100%. However, in contrast to Experiment 1, now almost all FB+ children also predicted that Lisa will look for her teddy in box 2, 81% in the immediate and 88% in the delayed condition (both above chance, Binomial tests: all p < .021). There was no reliable difference between conditions (Related-samples Wilcoxon signed-rank test: Z = 1, p = 1, r = 0.25). For FB+ children, the joint mean over the two conditions of 84% in Experiment 2 is reliably higher than the joint mean of 31% in Experiment 1 (Mann-Whitney U test: U = 118.5, p < .001, r = 0.52).

Our finding that FB+ children predicted in both conditions that Lisa will look for her teddy in box 2 demonstrates that the knowledge error all but vanishes when one ensures that children understand Lisa has seen the relevant event, the displacement of the teddy. Again, as in the
passive onlooker conditions used in Experiment 1, the timing of the test question played no role. The data support neither the triviality hypothesis nor the PAR account. If anything, the strong clue that Lisa is fully attending to the transfer should make the test question even more trivial and thus increase wrong answers instead of reducing them.

4. Experiment 3

In Experiment 2 we found a strong reduction in knowledge errors under active watching as opposed to the passive presence of Lisa in Experiment 1. This suggests that the knowledge error hinges on how children interpret Lisa’s passive presence. To them, it does not indicate that Lisa could see the transfer of the teddy. This important finding, however, rests on a comparison of different experiments. Experiment 3 checks whether it holds up under experimentally controlled conditions administered to the same sample.

4.1. Method

4.1.1. Participants

Thirty-six children (20 female, 16 male) aged 41 months to 80 months ($M = 65.6$ months, $SD = 10.64$) were included in the final sample. Children were recruited from two kindergartens in Salzburg (Austria). One additional five-year-old child was tested but had to be excluded due to language difficulties. Parents were previously informed and gave written consent to their child’s participation.

4.1.2. Design and procedure

As in the previous two experiments, each child was tested on one FB and two knowledge tasks. The two knowledge tasks had the same structure as the FB task, except that the protagonist was present during the object’s displacement. To directly assess the effect of active watching over passive presence, we compared the scenario from Experiment 1, where Lisa was just passively standing in the center of the scene during the transfer, with the scenario from Experiment 2, where Lisa followed her brother while he was moving the teddy and children were asked the see-question, whether the protagonist could see the transfer (see Fig. 2). Both conditions were presented in the delayed version. The see-question was asked in the active watching condition only.

4.2. Results and discussion

4.2.1. FB task

Children’s performance on the memory and reality questions was perfect. Five children failed the test question and were classified as FB- children ($mean \ age = 61$ months, $SD = 14.1$). Thirty-one children answered the test question correctly and were classified as FB+ children ($mean \ age = 66$ months, $SD = 10.3$).

4.2.2. Knowledge tasks

All children gave correct responses to the memory and reality questions in both conditions. Almost all children (except one of 36) responded with “yes” in the see-question. Fig. 3 (bottom left panel) shows the percentages of children correctly choosing box 2 in response to the test question for each condition. FB- children were consistently correct regardless of condition, whereas for FB+ children conditions mattered strongly. In the passive presence condition, only 35% of them responded correctly with “box 2” (not different from chance, Binomial test: $p = .150$), while 84% did so in the active watching condition (above chance, Binomial test: $p < .001$). The difference between conditions was significant (Related-samples Wilcoxon signed-rank test: $Z = 3.87, p < .001, r = .70$).

This response pattern confirms the suggestion from Experiment 1 and 2 that FB+ children interpreted Lisa’s passive presence differently than her active watching. When Lisa’s active watching makes clear that she has the relevant knowledge about the teddy’s displacement, these children correctly predict that Lisa will look for her teddy where it actually is. When this is not made clear, they assume Lisa is not aware of the transfer and consequently correctly predict in line with this assumption that Lisa will look for the teddy in its original place. Hence, their responses in Experiment 1 do not demonstrate that their belief understanding is flawed or that they are misled by a trivial test question. Rather, their knowledge “errors” result from how they interpret the information gain for passive onlookers, which differs from our adult intuitions.

5. General discussion

5.1. Summary of the main findings

We found in Experiment 1 that FB- and FB+ children responded very differently in both versions of the knowledge tasks – immediate vs. delayed posing of the test question. While FB- children predicted that Lisa will look for her teddy in the place where it actually is, FB+ children indicated that Lisa will look in the location where the teddy was in the beginning, thus demonstrating the same decline in children’s performance in knowledge tasks once they start to master FB as in a number of previous studies (Fabricius et al., 2010; Fabricius et al., 2021; Friedman et al., 2003; Oktay-Gür & Rakocy, 2017; Rakocy & Oktay-Gür, 2020; Schidelko et al., 2021). However, we noticed that some children asked the experimenter whether Lisa has actually seen the transfer. This suggests that Lisa’s mere presence was not reliably interpreted by these children as being aware of the relevant event, the displacement of the teddy.

In Experiment 2, we therefore introduced a new scenario. Instead of being just passively present during the displacement of the teddy, Lisa accompanies her brother and watches what he is doing. Additionally, we asked children whether Lisa could see the transfer. We contrasted again, as in Experiment 1, an immediate vs. a delayed condition. Both FB- and FB+ children now easily mastered this new active-engagement version, regardless of condition. Then, in Experiment 3, we contrasted directly the passive presence condition with the new condition of active watching, finding that FB+ children’s responses depended on the condition: When Lisa was only passively present, children indicated that she would look where the teddy was initially, whereas in the active watching condition children said that she would look for her teddy in the place where it actually was.

Across all three experiments, the 13 FB- children always responded that Lisa will look for her teddy in the location where it actually is. They did this regardless of whether the question was asked immediately or delayed after the transfer, or whether Lisa was only passively present during the displacement or was accompanying her brother when he moved the teddy and we asked if she could see that. This result clearly underlines that the FB- children give their answers based solely on reality. In contrast, the FB+ children’s responses are heavily dependent on whether Lisa was merely passively present or actively engaged, however, contrary to the predictions of the two hypotheses tested, they were indifferent to the time at which the test question was posed.

5.2. Implications from contrasting immediate with delayed test question

The aim of Experiment 1 was to test opposite predictions from the PAR and the triviality hypothesis for immediate and delayed test questions. As some children were asking the experimenter whether Lisa had seen the displacement of her teddy, it seems that at least some children doubted that Lisa was aware of the displacement despite the fact that she was present. That does not seem implausible. In this scenario, Lisa’s brother takes her teddy without asking and puts it in the other box, to which Lisa does not react in any way. Typically, one would expect children of this age to protest massively if they were in Lisa’s place. If children did not perceive the situation consistently in a way that Lisa noticed the transfer, then this task is a FB task for them, and they
therefore state that Lisa will look in the box where the teddy was at the beginning. Since the procedure of Experiment 1 leaves in doubt whether the children thought that Lisa saw the transfer or not, neither the predictions of the triviality hypothesis nor the PAR approach could be subjected to a critical test.

The change of procedure in Experiment 2, with Lisa accompanying her brother during the transfer and asking the see-question, ensured that children interpreted the scenario as it was intended, namely that Lisa perceived the transfer. This allowed us to assess anew the predictions from the triviality hypothesis and the PAR account that go in opposite directions. With the active watching versions in Experiment 2, we found that the majority of FB+ children anticipated Lisa looking for her teddy in its actual location. The knowledge error virtually disappeared by ensuring that children understood that Lisa attended to the relevant event, the moving of her teddy.

This finding seems hard to explain with the pragmatic account of Rakoczy and Oktay-Gür (Oktay-Gür & Rakoczy, 2017; Rakoczy & Oktay-Gür, 2020), from which our triviality hypothesis was derived. Tests of further hypotheses derived from the pragmatic view showed mixed results (Rakoczy & Oktay-Gür, 2020): With their manipulations the knowledge error was reduced in Experiment 3a and vanished in Experiment 1 and in Experiment 5 (in TB first, but not in FB first condition), but was not reduced in Experiments 3b, 3c, and 4.

One possibility to reconcile the pragmatic account with Experiment 2 is that our manipulation of Lisa actively watching the transfer may have alerted children that the protagonist’s perspective is of relevance, and that her perspective is part of the discourse in this situation. Under this assumption the pragmatic account (Oktay-Gür & Rakoczy, 2017; Rakoczy & Oktay-Gür, 2020) is consistent with the findings. According to this account, children have difficulties with the test question because it poses a trivial academic test question about a rational agent’s perspective; we do not normally engage in such discourse about subjective mental perspectives unless there is at least the possibility of error or deviance. The lack of such an obvious possibility in the knowledge task implies that there might be some hidden perspective difference and thus children may think they missed something. Our manipulation could have eliminated this problem by establishing that the perspective of the protagonist is something that is part of this discourse.

The PAR hypothesis does not only fail to predict the results of Experiment 2, it also fails to predict the results of Experiment 3. In both conditions, when Lisa leaves and returns after the transfer (delayed posing of the test question), FB+ children should consider Lisa to be in a new situation after her return and apply PAR anew, and thereby not refer back to the previous situation. According to the PAR account, children do not attribute to Lisa any memory of whether she was knowledgeable in the previous situation based on her perceptual access at that time. Upon Lisa’s return, children infer from her lack of perceptual access in the current situation that she does not know and therefore will act wrongly (e.g., Fabricius et al., 2010; Fabricius et al., 2021; Hedger & Fabricius, 2011). Children using PAR should predict that Lisa will look for the teddy in the now empty box 1 (where the teddy was in the beginning) in both conditions of Experiment 3. This prediction is not borne out in the data: In the condition where Lisa watched actively, most children indicated that Lisa will go to box 2, where the teddy is now (as was already observed in experiment 2 in the active/delayed condition).

5.3. Sensitivity to cues of attentive seeing

Experiments 2 and 3 showed that the knowledge error all but vanishes when it is clearly established that Lisa has seen the crucial event of the teddy being moved. In Experiment 1 some children explicitly expressed their uncertainty about Lisa’s awareness of the displacement of the teddy. To the extent that children interpret the scenario in such a way that Lisa was unaware of the transfer, a knowledge task effectively turns into a FB task for them. Therefore, if children did understand the story in this way, then they did not give “wrong” answers when they predicted that Lisa will look for her teddy in the now empty place but within their interpretation “correct” answers. After we modified the task children seemed to have grasped the stories as intended and practically all children (with the exception of only one of 55 children in Experiments 2 and 3) replied with “yes” to the see-question. Children’s answers to the test question differed considerably depending on whether Lisa was only passively present or actively engaged in watching the transfer of her teddy.

Deciding, whether someone can see an object or not requires level 1 visual perspective taking, which develops after the age of 18 months (Moll & Tomasello, 2006) and is well established at 3.5 years (Flavell, Everett, Croft, & Flavell, 1981; Lempers, Flavell, & Flavell, 1977; Masangkay et al., 1974). In both conditions – passive presence and active watching – the displacement of the teddy took place within Lisa’s visual field, that is, she had an unobscured line of sight. Children’s sensitivity to whether or not Lisa observed the transfer seems to go beyond a mere judgement that the transfer occurred in Lisa’s visual field. Although Lisa could have seen the critical event, children might have judged that she did not do so when only passively present, as she exhibits no signs of paying attention to her brother moving her teddy to the new place. Attention is a selective process, not everything that is in our visual field is actually seen. Perception is more than perceptibility (Flavell, 2004).

So far, evidence suggests that even 9- and 11-year-olds do not yet have an explicit understanding of this. When asked whether somebody can look at something, but not see it, the prevalent answer of 9-year-olds was “no”, whereas the majority of the 11-year-olds responded correctly with “yes” but justified this nonmomentarily with reference to problems with vision or lighting (Fabricius, Schick, Prost, & Schwanenflugel, 1997, cited in Flavell, 2004). However, our data now indicate that our FB- children (from about 4 years of age) are at least implicitly sensitive to this.

To our knowledge, this is a novel finding to be pursued in future research. But it is probably no accident that many studies with preschoolers have as part of the procedure that an observer is not just looking at an event (e.g., hiding a reward) but also showing clear signs of paying attention, e.g., by leaning forward, nodding with the head or make a remark like “ah” (e.g., Call & Tomasello, 1999; Fedra & Schmidt, 2019; Krachun, Carpenter, Call, & Tomasello, 2009). This suggests that some researchers have been intuitively aware that children at this age do not automatically infer from a person’s looking at an event that they must have seen what happened unless they shows clear signs of attentiveness.

The see-question has been used in some earlier knowledge stories. In two studies by Fabricius et al. (2010, 2021), the majority of FB+ children failed the knowledge task while almost all FB- children mastered it. Rakoczy and Oktay-Gür (2020, Experiments 3a – 3c) also used a see-question in their tasks when investigating whether the knowledge error disappears when the test questions are sincere basic questions and not test questions. Unfortunately, the use of the see-question was confounded with other modifications of how the critical questions were posed. Since the just cited studies deviated to different degrees from the procedure used in the current study and obtained somewhat different results, further systematic large-scale studies should investigate with 3- to 10-year-old children which of the differences matter.

5.4. Children’s developing theory of mind

The standard view on explicit theory of mind development holds that children around the age of 4 years start to ascribe beliefs to themselves and others. At this age they begin to master FB tasks in which they have to ascribe a mistaken belief to someone else (Wellman et al., 2001). The emerging competence in FB tasks goes hand in hand with the developing ability to master various tasks that also require the understanding of different perspectives, like false sign (Parkein, 1994) or identity tasks (Perner et al., 2011; Weingartner & Haring, 2020; for an overview see
Doherty & Perner, 2020; Perner & Roesler, 2012). It suggests the onset of robust, explicit metarepresentational thinking, allowing children to grasp the notion of propositional attitudes (Perner, 1991; Rakoczy, 2022; Wellman, 2011). Although children at age 4 are able to ascribe mental states to themselves and others, and subsequently pass standard FB tasks, it seems they still do not understand some basic features of mental states. Empirical evidence suggests that children this age still lack competence in taking into account that beliefs about an object depend on the aspect under which the object is represented (e.g., Apperly & Robinson, 1998; Kamawar & Olson, 1999; Russell, 1987) and the recursive nature of higher-order mental states (e.g., “John thinks that Mary thinks that…”; Perner & Winner, 1985). Children overcome their problems with aspecularity of belief when they start passing second-order belief tasks about 2 years later, at around age 6 (Huemer, Perner, & Leahy, 2018; Perner, Huemer, & Leahy, 2015).

The current study shows that FB+ children perceive the knowledge tasks in a different way than FB- children. Across the three experiments, only FB+ children were sensitive to whether Lisa was just passively present or actively watching the teddy being moved. Conversely, the response behavior of FB- children is completely unaffected by whether Lisa has seen the transfer or not, they always give a reality-based response. FB+ children’s sensitivity of whether Lisa actually saw the transfer or not goes beyond deciding whether the relevant event occurred in her visual field. Thereby they are showing awareness that seeing is an active mental process that is affected by attention. It is no surprise that this awareness comes at an age where children begin to understand mental states based on their emerging metarepresentational abilities. Povinelli, Bierschwale, and Cech (1999) argue along these lines when stating that appreciating the mental aspects of seeing includes the understanding that through seeing, new information about the world is gained and this contributes to one’s beliefs of how the world is. Children may not have such an understanding of seeing before the age of four.

The observed sensitivity of FB+ children to whether Lisa has actually seen the transfer or not goes beyond the understanding of perception in the PAR account. There, seeing is understood in the way of level 1 visual perspective taking, where children understand that someone can see an object when the line of sight is not obscured, but cannot see it when the line of sight is obscured (Fabricius et al., 2021). The PAR account does not only have problems in explaining children’s emerging conception of seeing as found here; it also has no explanation for why passing the false belief task is related to passing quite different perspective tasks like false sign and identity tasks (Doherty & Perner, 2020; Perner & Roesler, 2012).

Even though the present study makes an important point by showing that the knowledge error vanishes when it is made clear that Lisa could see the transfer, this does not help to provide an explanation for children’s errors in Gettier-type TB tasks (as used in Fabricius et al., 2010, Experiment 1). There, Lisa is absent when the teddy is taken out and then put back again. To our knowledge, in the literature, only the PAR account provides a consistent explanation for this type of error. It is not clear to us how, for example, the pragmatic account of Rakoczy and Oktay-Gür could explain these results. The TB task differs from an FB task only in that in Lisa’s absence the teddy, after being taken out, is put back in the same place instead of in a different place. So it cannot be more trivial than an FB task.

6. Conclusion

We found that the knowledge error all but vanishes when it has been made clear to children that Lisa has been aware of the displacement of her teddy. If this is not made clear children assume Lisa has not recognized the transfer and, under this premise, correctly predict that Lisa would look for her teddy in its original place. The response of the FB- children was not influenced by the timing of the test question and, therefore, gave no support for the use of PAR. A further interesting finding of this study is that FB- children are somewhat sensitive to Lisa’s signs of attentiveness to the critical event and not just whether it happened in her visual field or not.

CRediT authorship contribution statement

Michael Huemer: Conceptualization, Methodology, Formal analysis, Writing – original draft, Writing – review & editing, Visualization, Funding acquisition. Lara M. Schröder: Writing – original draft, Writing – review & editing, Visualization. Sarah J. Leikard: Formal analysis, Writing – review & editing. Sara Gruber: Investigation. Anna Mangstl: Investigation. Josef Perner: Conceptualization, Methodology, Formal analysis, Writing – original draft, Writing – review & editing, Supervision, Funding acquisition.

Data availability

The data set is publicly available, and can be accessed at https://doi.org/10.17632/3jk5pmnfc3.1.

Acknowledgement

This work was supported by the Austrian Science Fund (FWF) [Research Project I 3518-G24], and the Austrian Academy of Science (OAW) [DOC scholarship 24691]. None of these funding sources had any input in the design of the study, the collection, analysis, or interpretation of the data, in the writing of the report, or in the decision to submit the article for publication.

References

Apperly, I. A., & Robinson, E. J. (1998). Children’s mental representation of referential relations. Cognition, 67, 267–309. https://doi.org/10.1016/S0010-0277(98)00030-4
Call, J., & Tomasello, M. (1999). A nonverbal false belief task: The performance of children and great apes. Child development, 70(2), 381–395. https://doi.org/10.1111/1467-8624.00028
Doherty, M. J., & Perner, J. (2020). Mental files: Developmental integration of dual naming and theory of mind. Developmental Review, 56, Article 100909. https://doi.org/10.1016/j.dr.2020.100909
Fabricius, W., Schick, K., Prost, J., & Schwanenflügel, P. (1997). We don’t see eye to eye: Development of a constructivist theory of mind [poster presentation]. In Biennial meeting of the Society for Research in Child Development, Washington, D.C., USA, April 3–6.
Fabricius, W. V., Boyer, T. W., Weiner, A. A., & Carroll, K. (2010). True or false: Do 5-year-old understand belief? Developmental Psychology, 46(6), 1402–1416. https://doi.org/10.1037/a0017648
Fabricius, W. V., Gonzales, C. R., Pesch, A., Weiner, A. A., Pagliese, J., Carroll, K., … Spinard, T. L. (2021). Perceptual access reasoning (PAR) in developing a representational theory of mind. Monographs of the Society for Research in Child Development, 86(3), 7–154. https://doi.org/10.1111/mono.12492
Fabricius, W. V., & Imbens-Bailey, A. L. (2000). False beliefs about false beliefs. In P. Mitchell, & K. Rigg (Eds.), Children’s reasoning about the mind (pp. 267–280). Hove: Psychology Press.
Fabricius, W. V., & Khalil, S. L. (2003). False beliefs or false positives? Limits on children’s understanding of mental representation. Journal of Cognition and Development, 4(3), 239–262. https://doi.org/10.1207/s15327674jod0403_01
Fedra, E., & Schmidt, M. F. H. (2019). Older (but not younger) preschoolers reject incorrect knowledge claims. British Journal of Developmental Psychology, 37(1), 130–145. https://doi.org/10.1111/bjdp.12264
Fernández, C. (2013). Mindful storytellers: Emerging pragmatics and theory of mind development. First Language, 33(1), 20–46. https://doi.org/10.1177/0142723711422631
Flavell, J. H. (2004). Development of knowledge about vision. In D. T. Levin (Ed.), Thinking and seeing: Visual metacognition in adults and children (pp. 13–36). A Bradford Book, The MIT Press.
Flavell, J. H., Everett, B. A., Croff, K., & Flavell, E. R. (1981). Young children’s knowledge about visual perception: Further evidence for level 1 - level 2 distinction. Developmental Psychology, 17, 99–103. https://doi.org/10.1037/0012-1649.17.1.99
Friedman, O., Griffin, R., Brownell, H., & Winner, E. (2003). Problems with the seeing as found here; it also has no explanation for why passing the false belief task is related to passing quite different perspective tasks like false sign and identity tasks (Doherty & Perner, 2020; Perner & Roesler, 2012).

Even though the present study makes an important point by showing that the knowledge error vanishes when it is made clear that Lisa could see the transfer, this does not help to provide an explanation for children’s errors in Gettier-type TB tasks (as used in Fabricius et al., 2010, Experiment 1). There, Lisa is absent when the teddy is taken out and then put back again. To our knowledge, in the literature, only the PAR account provides a consistent explanation for this type of error. It is not clear to us how, for example, the pragmatic account of Rakoczy and Oktay-Gür could explain these results. The TB task differs from an FB task only in that in Lisa’s absence the teddy, after being taken out, is put back in the same place instead of in a different place. So it cannot be more trivial than an FB task.

6. Conclusion

We found that the knowledge error all but vanishes when it has been made clear to children that Lisa has been aware of the displacement of her teddy. If this is not made clear children assume Lisa has not recognized the transfer and, under this premise, correctly predict that Lisa would look for her teddy in its original place. The response of the FB- children was not influenced by the timing of the test question and, therefore, gave no support for the use of PAR. A further interesting finding of this study is that FB- children are somewhat sensitive to Lisa’s
Hedger, J. A. (2016). Perceptual access reasoning: Developmental stage or system 1 heuristic? Phenomenology and the Cognitive Sciences, 15(2), 207–226. https://doi.org/10.1017/s11097-015-9412-4

Hedger, J. A., & Fabrigus, W. V. (2011). True belief belies false belief: Recent findings of competence in infants and limitations in 5-year-olds, and implications for theory of mind development. Review of Philosophy and Psychology, 2(3), 429–447. https://doi.org/10.1007/s13164-011-0069-9

Huemer, M., Perner, J., & Leabey, B. (2018). Mental files theory of mind: When do children consider agents acquainted with different object identities? Cognition, 171, 122–129. https://doi.org/10.1016/j.cognition.2017.10.011

Kamawar, D., & Olton, D. R. (1999). Children’s representational theory of language: The problem of opaque contexts. Cognitive Development, 14, 531–548. https://doi.org/10.1016/s0885-2069(99)00018-0

Krachun, C., Carpenter, M., Call, J., & Tomasello, M. (2009). A competitive nonverbal problem of opaque contexts. Developmental Science, 12(4), 521–535. https://doi.org/10.1111/j.1467-7687.2008.00793.x

Lemper, J. D., Flavell, E. R., & Flavell, J. H. (1977). The development in very young children of tacit knowledge concerning visual perception. Genetic Psychology Monographs, 95, 3–53.

Masangkay, Z. S., McCluskey, K. A., McIntyre, C. W., Sims-Knight, J., Vaughan, B. E., & Flavell, J. H. (1974). The early development of inferences about the visual percept of others. Child Development, 45, 257–266. https://doi.org/10.2307/1127956

Moll, H., & Tomasello, M. (2006). Level I perspective-taking at 24 months of age. British Journal of Developmental Psychology, 24(3), 603–613. https://doi.org/10.1348/026151099165140

Okbay-Gür, N., & Rakoczy, H. (2017). Children’s difficulty with true belief tasks: Competence deficit or performance problem? Cognition, 166, 28–41. https://doi.org/10.1016/j.cognition.2017.05.002

Parkin, L. J. (1994). Children’s understanding of misrepresentation. Doctoral dissertation. Brighton: University of Sussex.

Perner, J. (1991). Understanding the representational mind. Cambridge, MA: The MIT Press.

Perner, J., Huemer, M., & Leabey, B. (2015). Mental files and belief: A cognitive theory of how children represent belief and its intentionalty. Cognition, 145, 77–88. https://doi.org/10.1016/j.cognition.2015.08.006

Perner, J., Leekam, S. R., & Wimmer, H. (1984). The insincerity of conservation questions: Children’s growing sensitivity to experimenter’s epistemic intentions. Unpublished manuscript. In Laboratory of Experimental Psychology, University of Sussex. Presented as poster at the APA (American Psychological Association) convention, August 1984, and at the EPS-developmental section meeting in Lancaster, September 1984.

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

Perner, J., & Roesler, J. (2012). From infants’ to children’s appreciation of belief. Trends in Cognitive Sciences, 16(10), 519–525. https://doi.org/10.1016/j.tics.2012.08.004

Perner, J., & Wimmer, H. (1985). “John thinks that Mary thinks that . . .”: Attribution of second-order beliefs by 5- to 10-year-old children. Journal of Experimental Child Psychology, 39, 427–471. https://doi.org/10.1016/0022-0965(85)90051-7

Povinelli, D. J., Bierschwale, D. T., & Cech, C. G. (1999). Comprehension of seeing as a referential act in young children, but not juvenile chimpanzees. British Journal of Developmental Psychology, 17(1), 37–60. https://doi.org/10.1348/026151099165140

Priebe, B., Fowles, J., Schweller, K., & Perner, J. (2020). Mistaken max befriends Duplo girl: No difference between a standard and an acted-out false belief task. Journal of Experimental Child Psychology, 191, Article 104756. https://doi.org/10.1016/j.jecp.2019.104756

Priebe, B., Roesler, J., & Perner, J. (2013). Competition as rational action: Why young children cannot appreciate competitive games. Journal of Experimental Child Psychology, 116(2), 545–559. https://doi.org/10.1016/j.jecp.2014.05.003

Rakoczy, H. (2022). Foundations of theory of mind and its development in early childhood. Nature Reviews Psychology, 1(4), 223–235. https://doi.org/10.1038/s44159-022-00037-x

Rakoczy, H., & Okbay-Gür, N. (2020). Why do young children look so smart and older children look so dumb on true belief control tasks? An investigation of pragmatic performance factors. Journal of Cognition and Development, 21(2), 213–239. https://doi.org/10.1080/15248372.2019.1709467

Ruffman, T. (1996). Do children understand the mind by means of simulation or a theory? Evidence from their understanding of inference. Mind & Language, 11(4), 388–414. https://doi.org/10.1111/j.1468-0017.1996.tb00052.x

Russell, J. (1987). “Can we say . . .”? Children’s understanding of intensionality. Cognition, 25, 289–308. https://doi.org/10.1016/0010-0277(87)90007-0

Schidelko, L. P., Proft, M., & Rakoczy, H. (2022). How do children overcome their pragmatic performance problems in the true belief task? The role of advanced pragmatics and higher-order theory of mind. PLoS One, 17(4), Article e0266959. https://doi.org/10.1371/journal.pone.0266959

Schidelko, L. P., Schümann, B., Rakoczy, H., & Proft, M. (2021). Online testing yields the same results as lab testing: A validation study with the false belief task. Frontiers in Psychology, 4573. https://doi.org/10.3389/fpsyg.2021.703238

Siegel, M. (1999). Language and thought: The fundamental significance of conversational awareness for cognitive development. Developmental Science, 2(1), 1–14. https://doi.org/10.1111/1467-7687.00048

Weingartner, P., & Haring, S. (2020). Children’s understanding of identity – The kind of referent that children have in mind when understanding identity arguments. Humanities and Social Sciences, 8(2), 51–62. https://doi.org/10.11648/j.

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

Wellman, H. M. (2011). Developing a theory of mind. In U. Goswami (Ed.), The Wiley Blackwell handbook of childhood cognitive development (2nd ed., pp. 258–284). Wiley-Blackwell.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children’s understanding of deception. Cognition, 13(1), 103–128. https://doi.org/10.1016/0010-0277(83)90004-5

Winner, E., & Gardner, H. (1993). Metaphor and irony: Two levels of understanding. In A. Ortony (Ed.), Metaphor and thought (pp. 425–443). Cambridge University Press.