

## 2.5-Year-Olds Express Suspense When Others Approach Reality With False Expectations

Henrike Moll, Allie Khalulyan, and Lillie Moffett  
*University of Southern California, California*

The study investigated if 2.5-year-olds are susceptible to suspense and express tension when others' false expectations are about to be disappointed. In two experiments ( $N = 32$  each), children showed more tension when a protagonist approached a box with a false belief about its content than when she was ignorant. In Experiment 2, children also expressed more tension when the protagonist's belief was false than when it was true. The findings reveal that toddlers affectively anticipate the "rude awakening" of an agent who is about to discover unexpected reality. They thus not only understand false beliefs per se but also grasp the affective implications of being mistaken. The results are discussed with recourse to current theories about early understanding of false beliefs.

In his commentary to Premack and Woodruff's (1978) article on theory of mind in chimpanzees, Daniel Dennett brainstormed ways in which a theory of mind could be tested in children. This inspired developmentalists to create what are known today as standard false belief tasks, in which children are asked what a misinformed agent believes or will do (Perner, Leekam, & Wimmer, 1987; Wimmer & Perner, 1983)—with their answers being typically wrong before ages 4–5 (Wellman, Cross, & Watson, 2001). Dennett (1978) did not in fact envision a task in which children are "interrogated." He looked for something "direct, natural, plausible" (p. 569) and ideally nonverbal. What struck him as indicative of false belief comprehension were children's reactions during puppet shows: How they "squeal in anticipatory delight" (p. 569) when a misinformed puppet is about to make an unpleasant discovery. He gives the example of Punch tossing around a box to punish Judy, whom he takes to be inside, with children knowing that Judy sneakily escaped moments before.

Almost 40 years later, tests have been created that dispense with the "interrogation" format and try to tap belief understanding implicitly or indirectly. These include (a) *visual attention tests*, such

as looking time (e.g., Kovács, Téglás, & Endress, 2010; Onishi & Baillargeon, 2005) and preferential looking (Scott, He, Baillargeon, & Cummins, 2012); (b) *interaction tests* in which infants between 1.5 and 2 years help or communicate with a misguided agent (e.g., Buttellmann, Carpenter, & Tomasello, 2009; Happe & Loth, 2002; Knudsen & Liszkowski, 2012a, 2012b; Southgate, Chevallier, & Csibra, 2010); and (c) *anticipation tests* in which 2- to 3-year-olds anticipate where a misguided agent will go (e.g., Clements & Perner, 1994; Southgate, Senju, & Csibra, 2007).

Remarkably, none of the tests capitalized on that which makes puppet shows so intriguing: their affective charge. Not only is Punch's belief about Judy's whereabouts incorrect, but discovering that Judy is not in the box will leave him perplexed and frustrated. All false belief tasks share this moment of surprise or puzzlement (e.g., missing chocolate in the change of location), but with few exceptions, research has ignored this affective dimension. The few studies that focused on this aspect revealed that knowledge about the affective impact of incorrect beliefs lags behind false belief attribution per se. Children under 5 or 6 years fail to predict that an agent with a false belief will be surprised (Hadwin & Perner, 1991; MacLaren & Olson, 1993; Ruffman & Keenan, 1996), and some suggest that children operate with only an ignorance-, not belief-based concept of surprise until ages 7–9 (Ruffman & Keenan, 1996).

But again, these studies reverted to the interrogation method and only acknowledged the agent's

---

A Khalulyan and L Moffett with equal contribution.

Henrike Moll was at the Notre Dame Institute for Advanced Study as a Templeton Fellow when this article was written. This research was partially supported by a grant from the Office of Naval Research Award N000141410555. Special thanks to Anna Schwartz and Kristina Fenske for help with coding. We are indebted to the children's museum *Kidspace* in Pasadena, CA, and to USC's Child Development Center for their cooperation.

Correspondence concerning this article should be addressed to Henrike Moll, Department of Psychology, SGM 501, 3620 S. McClintock Ave., Los Angeles, CA 90089. Electronic mail may be sent to hmoll@usc.edu.

© 2016 The Authors

Child Development © 2016 Society for Research in Child Development, Inc.

All rights reserved. 0009-3920/2016/xxxx-xxxx

DOI: 10.1111/cdev.12581

affect. What makes puppet shows unique testing grounds is the children's own affective investment. The anticipatory delight that Dennett noted hints at the workings of a psychological tool used in the narrative arts known as induction of suspense (Harsh, 1944; Truffaut, 1967). Suspense is the tension experienced by an observer who has privileged "foreknowledge" of the protagonist's fate while being unable to intervene. Regardless of whether it takes shape as empathy, *schadenfreude*, or some emotional blend (Zillmann, 1996), this tension indicates that the observer is aware of the clash between reality and how the protagonist views it.

Moll, Kane, and McGowan (2016) tested if 3-year-olds are susceptible to suspense. When a protagonist puppet was about to make an unexpected discovery, children expressed tension by, for example, biting their lips or furrowing their brows. They rarely did so when the protagonist knew reality or had no specific expectation (was merely ignorant), suggesting that the expressions sprang from the perceived clash between reality and its misrepresentation by the protagonist and thus manifested belief understanding.

The goal of this study is to test if suspense of anticipation can be induced in 2.5-year-olds. If toddlers this age show affective anticipations in light of others' false beliefs, then the divide between implicit and explicit belief comprehension is even greater than was assumed. It would demonstrate that children affectively grasp the consequences of erring about 5 years before they are able to explicitly predict them and 2 years prior to passing the standard false belief tasks. Such a finding would challenge reductive accounts of early belief reasoning and increase the demand for theoretical advances that explicate the cognitive basis of earlier in relation to later emerging abilities.

In two experiments, 2.5-year-olds observed how a protagonist returned to a box containing an object that an antagonist had devalued (e.g., damaged) in her absence (false belief condition). The same children were assigned to one of two control groups: Those in the true belief group additionally received stories in which the protagonist saw the manipulation (true belief condition), and those in the ignorance group received stories in which the protagonist lacked specific expectations about the object (ignorance condition). We hypothesized that children in both groups would show more tension in the false belief than in the respective control condition. In Experiment 1, the antagonist *maliciously* devalued the protagonist's object (adopting the classic character constellation of a protagonist who is

hindered by "opposing forces"). Because children in the true belief group failed to discriminate and showed similar tension whether the protagonist's belief was true or false, the antagonist's malicious motive was removed in Experiment 2 and the object's devaluation was marked as *accidental*. As will be shown, this effectively reduced children's expressions in response to true beliefs.

## Experiment 1

### *Method*

#### *Participants*

Participants were thirty-two (16 female) 2.5-year-olds ( $M = 30;07$ , range = 26;00–32;23 for the true belief group;  $M = 30;12$ , range = 28;23–32;08 for the ignorance group). They were recruited through the university's child-care center and a database of parents who had volunteered to have their child participate in developmental studies. The racial/ethnic composition was highly diverse, as is representative of the Greater Los Angeles Area, in which testing took place: Thirteen children were biracial/mixed, 9 White, 7 Hispanic, 2 African American, and 1 Asian. Their socioeconomic status as measured via family income varied greatly, with incomes ranging from < \$20,000 to > \$120,000 (and a median in the upper middle class range). No child was excluded from the sample. Data were collected between May 2014 and February 2015.

#### *Materials and Design*

A female experimenter (*E*) acted out four stories (*Reduction* story, *Deconstruction* story, *Replacement* story, *Breaking* story) on a white puppet theater (74 cm wide, 71 cm high, and depth of stage: 14 cm). There were three puppets (all between 26 × 16 × 13 cm and 55 × 30 × 17 cm) and various props in each story. Two of them (protagonist, antagonist) were used in all conditions (false belief, true belief, ignorance); the third (a protagonist's friend/relative) appeared only in the ignorance condition. For the *Reduction* story, the puppets were Cookie Monster (protagonist), a man named Tony (antagonist), and Elmo (friend). Props were a cylindrical cookie box (18 cm high, 7 cm in diameter) and nine wooden cookies. For the *Deconstruction* story, the puppets were a boy named Max (protagonist), a penguin named Pengu (antagonist), and a Grandfather. Props were a round container (14 cm high, 13 cm in diameter) and a small construction (3.5 cm

wide, 7.5 cm high) made from Lego® Blocks, Lego, Billund, Denmark. Figure 1 shows the stimuli used in this story. For the *Replacement* story, the puppets were a boy named Franz (protagonist), a girl named Jen (antagonist), and a woman named Fiona (friend). Props were a round tin box (14 cm high, 13 cm in diameter), a colored plush ball (11 cm in diameter), and a dirtied table tennis ball (4 cm in diameter). For the *Breaking* story, two girl puppets named Lucy (protagonist) and Twyla (antagonist) and a monkey puppet named George (friend) were used. Props were a cardboard box (6.5 cm high, 7 cm in diameter) and a yellow ring (3 cm wide, 5.5 cm long) with an inserted light (that could be switched on and off).

Each story came in three versions: false belief, true belief, and ignorance. Every child received the false belief version of two stories and either the true belief or ignorance version of the remaining two stories. Children were randomly assigned to the true belief or the ignorance group. The order of stories and conditions was counterbalanced.

#### Procedure

The child was seated in a chair 90 cm in front of the puppet theater. E announced the start of the show, opened the curtains, and acted out the first story.

All stories followed the same general structure: The protagonist entered the stage from the child's left and introduced herself. In the false belief condition, she retrieved an object, engaged with it briefly, placed it in a box on stage and left. In the true belief condition, the protagonist did the same but instead of leaving, she positioned herself on a dowel in the left corner from where she could observe the stage. In the ignorance condition, the protagonist came on stage and took position on the dowel. Her friend entered with a box, told the protagonist what it contained (e.g., cookies), and that she could engage with it later. The two then exited without the protagonist having seen the box's content. In all conditions, 4 s after the protagonist left/was placed on the dowel (true belief), the antagonist entered from the right and introduced herself. She then intentionally manipulated (e.g., reduced, see Appendix) the object with a malicious "heh-heh" sound. In the true belief condition, the protagonist made a "hmm" sound twice (once right before and once right after the manipulation) to convey that she registered the act. The antagonist placed the object in the box and left. In the true belief condition, the protagonist stated that she had to leave and exited. In all conditions, the protagonist returned 4 s later, marking the beginning of the response phase. She expressed her intention to fetch



Figure 1. Stimuli (puppets and props) used in one of the stories: *Deconstruction* story.

the object (e.g., the cookies), traversed the stage left to right, grasped the container (“I’m taking this outside”), and exited. The response phase lasted approximately 15 s. Between stories, *E* closed the curtains, and then prepared and announced the next story.

The specific procedures for the four stories can be found in the Appendix.

#### Scoring and Reliability

Children were video-taped and their expressions in the response phase were coded by a rater who was unaware of condition. (She did not see what the child observed and the sound was off.) The rater judged if the child exhibited at least one tense expression in the trial (1 = *expressions present*, 0 = *expressions absent*) using Moll et al.’s (2016) coding scheme. Behaviors listed in the coding scheme included brow furrowing, brow raising, sudden mouth opening or shutting, biting/“tucking in” the lip, placing the fingers/hand over or in the mouth, tense tongue motion/biting, smirking, making a fist, and nervous or agitated body movement. A behavior was only coded if it reflected tension from observing the puppet show (e.g., mouth opening while looking away was not coded), and if it was not continually present throughout the story (e.g., having the mouth open during most of the show was not coded). Multiple expressions per trial were noted, but did not affect scoring.

To assess interrater reliability, a second rater, also unaware of condition, coded the expressions of 16 (50%) randomly chosen children. The raters disagreed on 4 of the 64 trials, equaling an agreement of 94% ( $\kappa = .76$ ). Reliability was thus good. Because one child only completed three of the four trials total, we calculated the proportion of trials with tension by adding trials scored as “1” and dividing them by the number of completed (usually 2) trials for each condition. Possible values were thus 0, 0.5, and 1.

#### Results and Discussion

There were no effects of trial (learning/fatigue),  $p = .92$ , or condition order,  $p = .40$ . A total of 48 expressions of tension (24 in the true belief group and 24 in the ignorance group) were observed. The most common were biting or tucking in the lip (21%), smirking (19%), and brow furrowing (15%). Table 1 shows how many children in each group received a given combination of scores. In the true

Table 1

Experiment 1: Number of Children Per Group (True Belief vs. Ignorance) Who Received a Given Proportional Score in the Two Conditions

		False belief condition		
		0	0.5	1
Control condition	0	3	3	1
	0.5	4	5	—
	1	—	—	—
True belief	0	3	3	1
	0.5	4	5	—
	1	—	—	—
Ignorance	0	1	7	5
	0.5	3	—	—
	1	—	—	—

Note. There were two trials per condition, so possible proportions are 0, 0.5, and 1.

belief group, 4 of the 16 children expressed suspense more often in the false belief than in the true belief condition, 8 children expressed suspense equally in the two conditions, and 4 showed more suspense in the true than the false belief condition. The mean proportion of trials in which expressions were shown was .28 ( $SD = .26$ ) for the true belief and .31 ( $SD = .31$ ) for the false belief conditions. A Wilcoxon signed-rank test revealed no significant difference between the two conditions,  $W = -4.31$ ,  $p = 1.00$ .

In the ignorance group, 12 of the 16 children showed more suspense in the false belief than the ignorance condition. One child expressed suspense equally in the two conditions, and three children expressed more suspense in the ignorance than in the false belief condition. The mean proportion of trials with expressions was .09 ( $SD = .20$ ) for the ignorance and .53 ( $SD = .39$ ) for the false belief conditions. A Wilcoxon signed-rank test was significant, with children showing more tension in the false belief than the ignorance condition,  $W = -90.50$ ,  $p = .011$ .

To compare the two groups (true belief vs. ignorance) to each other, a Mann–Whitney *U* test with difference scores (proportion of expressions in the false belief condition minus the proportion of expressions in the true belief or ignorance condition) was conducted. It showed that children in the ignorance group differentiated significantly better between false belief and control trials than children in the true belief group,  $U = 70.00$ ,  $p = .019$ .

These results are inconclusive. On the one hand, children showed more tension when the protagonist had a false belief than when she was merely

ignorant—suggesting that the 2.5-year-olds anticipated the ill-informed agent’s affective response to unexpected reality. On the other hand, children showed similar tension when the agent represented the world correctly as when she misrepresented it. The indiscrimination between true and false belief casts doubt on the interpretation of tension as suspense, which relies on a clash between reality and another’s misconception of it. One possibility is that children’s expressions in the true belief condition are of a different origin. Here, children may have sympathized with the victim who proceeds to gather what the mischievous antagonist left her with. Introducing a malicious motive may thus have created an independent cause of tension when the protagonist knew what had happened.

To remove this problem, we had an antagonist devalue the objects accidentally in Experiment 2. Such “mishaps” should not leave the protagonist resentful, so children should not be inclined to show tension in the true belief condition.

## Experiment 2

### Method

#### Participants

Participants were thirty-two (16 female) 2.5-year-olds ( $M = 31;13$ , range = 28;23–34;15 for the true belief group;  $M = 29;27$ ; range = 26;23–33;07 for the ignorance group). Sources for recruitment were the same as in Experiment 1. The racial/ethnic makeup was again highly diverse: 15 children were White, 10 biracial/mixed, 4 Hispanic, 2 African American, and 1 Asian. The socioeconomic background as measured by income varied as widely as in Experiment 1, with a median in the upper middle class range. No child was excluded. Data were collected between February and April 2015.

#### Design and Procedure

The only procedural difference from Experiment 1 was that the antagonist devalued the object *accidentally*. In the *Reduction* story, she dropped the stack of cookies on the ground (“Oh no, they fell!”), so they had to be discarded. In the *Deconstruction* story, the antagonist accidentally dropped the Lego® construction (“Oh no, it fell apart!”). In the *Replacement* story, the ball slipped from the antagonist’s hands (“Oh no, it rolled away!”). In the *Breaking* story, the antagonist accidentally broke the light of the ring (“Oh no, the light broke!”).

#### Scoring and Reliability

The scoring and reliability procedures were the same as in Experiment 1. A second rater, unaware of the condition, coded the expressions of 16 (50%) children. The raters disagreed on 3 of the 64 trials, equaling an agreement of 95% ( $\kappa = .87$ ). All children completed all four trials. To be consistent and facilitate comparability with Experiment 1, proportional scores were again calculated (sum of trials scored as “1” divided by 2).

#### Results and Discussion

There were no effects of trial,  $p = .87$ , or condition order,  $p = .58$ . A total of 44 tense expressions (23 in the true belief group and 21 in the ignorance group) were observed. The most common were smirking (39%), lip biting/“tucking in” the lip (25%), and protruding or biting the tongue (14%). Table 2 shows, broken down by group, how many children received a given combination of scores. In the true belief group, 11 expressed more suspense in the false than the true belief condition, 3 expressed suspense equally in both conditions, and 2 showed more suspense in the true belief than the false belief condition. The mean proportion of trials with expressions were .06 ( $SD = .17$ ) for the true belief and .44 ( $SD = .36$ ) for the false belief conditions. A Wilcoxon signed-rank test showed that this difference is significant,  $W = -72.98$ ,  $p = .014$ . Figure 2 depicts a child in a false (left) and true (right) belief trial.

In the ignorance group, the ranks were the same as in the true belief group; that is, 11 children showed more suspense in the false belief than the ignorance condition, 3 did not distinguish, and 2

Table 2  
Experiment 2: Number of Children Per Group (True Belief vs. Ignorance) Who Received a Given Proportional Score in the Two Conditions

		False belief condition		
		0	0.5	1
Control condition				
True belief	0	3	8	3
	0.5	2	—	—
	1	—	—	—
Ignorance	0	2	10	1
	0.5	2	1	—
	1	—	—	—

Note. There were two trials per condition, so possible proportions are 0, 0.5, and 1.



Figure 2. A child during the response phase of a false (left) and a true belief (right) trial.

expressed more suspense in the ignorance than the false belief condition. Mean proportions of trials with expressions were .09 ( $SD = .20$ ) for the ignorance and .41 ( $SD = .27$ ) for the false belief conditions. A Wilcoxon signed-rank test showed that this difference is significant,  $W = -71.55$ ,  $p = .020$ .

To compare performance across groups (true belief vs. ignorance), a Mann–Whitney  $U$  test with difference scores (proportion of false belief trials with tension minus proportion of control trials with tension) as dependent measure was conducted. It showed no significant difference: The groups discriminated equally well between false belief and true belief or ignorance, respectively,  $U = 117.00$ ,  $p = .72$ .

We tested the prediction that tension in the true belief condition would diminish if the protagonist saw her object being accidentally instead of maliciously devalued, leaving her with no reason for a resentful attitude. Consistent with this prediction and with work demonstrating young children’s sensitivity to antisocial intentions (Hamlin, Wynn, Bloom, & Mahajan, 2011), the children rarely expressed tension and did so less than in the false belief condition.

### General Discussion

This study shows that 2.5-year-olds are susceptible to suspense and affectively anticipate an ill-informed agent’s discovery of unexpected reality. In two experiments, they expressed more tension when an agent approached an object with false

assumptions than when she was merely ignorant. In Experiment 2, they also showed more tension when the agent’s belief was false as opposed to true, suggesting that their responses reflect knowledge of the oncoming unexpected discovery by the agent. (See Video S1, for expressions observed in false belief trials of the two experiments.)

In a similar study, 39-month-olds also showed anticipatory affect when others acted on false assumptions (Moll et al., 2016). One major question is how this skill set that is shared between 2.5- and 3-year-olds emerges and changes over time. Longitudinal studies ought to explore this question, but based on our observations, we believe that the paradigm might come to its limits when applied to young 2-year-olds and infants because their expressions lack contour and definition. This could reflect an absence of affective anticipation or a broader failure to follow or understand the stories. With age, the toddlers’ expressions not only become more contoured and defined. They are also complemented with verbal warnings and predictive judgments—a behavior that some of the 3-year-olds, but none of the 2.5-year-olds tested here showed. This shows that children gain a more explicit grasp of how reality will affect the agent.

A further developmental question concerns the range of narratives in response to which children display their belief understanding. The 3-year-olds in Moll et al.’s (2016) study had to grasp that even though the antagonist meant well (e.g., a doctor reduced the number of cookies in a box for health reasons), the effect of her action would still leave the protagonist disappointed (less cookies than

expected). Pilot data collected prior to this study suggested that 2.5-year-olds do not show differential responses to the false and true belief versions of such stories and might need behavioral markers indicating that the object's modification is unfortunate for the protagonist (e.g., by marking it as a mishap as in Experiment 2). Finally, future studies ought to look into developmental differences. It is conceivable that toddlers who show anticipatory concern have advanced linguistic or narrative skills that allow them to understand the stories, and it might be that children showing anticipatory empathy also tend to engage in acts of helping in intervention tasks.

Our data challenge deflationary accounts of early belief understanding. One of them is Fabricius, Boyer, Weimer, and Carroll's (2010) perceptual access reasoning (PAR) hypothesis, according to which children up to age 6 cannot ascribe beliefs, but use the simple heuristic that agents who did not witness critical changes will "get things wrong." Findings cited in support of this view stem from versions of the change-of-location task in which a third location (C) was added to the object's current (B) and original (A) place. Instead of attributing the belief to the agent that the object is in A, children under 6 simply expected her to search unsuccessfully in A and C (Fabricius & Khalil, 2003; but see Perner & Horn, 2003). Our findings speak against the PAR account because the protagonist did not see the critical change in either the false belief or ignorance condition, but children only reliably showed tension (and assumed the agent "had it wrong") in the former.

Another reductive account that seems incompatible with the present findings states that young children directly link people's observations with subsequent behavior via statistical learning, without going through the mental as a mediator (Ruffman, 2014; Ruffman, Taumoepeau, & Perkins, 2012). This interpretation can deal with anticipation tests in which an agent's behavior covaries with her prior perception: She goes to location A when she did not see the object transfer, and to B when she did. Having detected these regularities allows infants to anticipate an agent's motion trajectory (Southgate et al., 2007). In our task, however, the agent behaved identically across conditions: She went to the same box with the same content, yet children's responses differed depending on her epistemic attitude.

A promising lean interpretation is offered by the two-systems theory (Apperly & Butterfill, 2009). System 1 is innate and supports efficient but limited belief reasoning. It captures beliefs not as

propositionally structured representations but as "registrations," which enables even infants to track beliefs about object locations ("The blue car is *in the box*") and respond to their violation. But the system is subject to signature limits: It cannot capture beliefs for which, for example, mode of presentation or appearance–reality distinctions matter (e.g., "There is a ghost in the closet" when the "ghost" is really a white sheet wrapped around a coat hanger). Borderline cases are beliefs with quantifiers or specific properties in the content, such as "There are *many* cookies in the jar" or "There is a *small blue* car in the box," which are rendered incorrect if the quantity or a property is altered (Apperly & Butterfill, 2009, p. 963). It is beliefs of this kind that children have to track—and the violation of which they have to anticipate—in our suspense task: An object undergoes a change in quantity, size, and so on. Further delineations of the boundary of System 1 seem necessary to decide if these cases are manageable for System 1 or afford the later emerging System 2.

We think the children in our task engage in a kind of epistemic perspective taking that is informed by reality. Crucially, a "full-on" or total adoption of the protagonist's current view of the situation would not bring about tension, because the agent is happily approaching the box at the given time. For tension to arise, the children must retain a concurrent awareness of actual reality (the devalued object) and imagine how it will affect the agent. In this respect, the responses differ from anticipatory looking behavior (Southgate et al., 2007), which can be supported by total perspective taking that is unfettered by reality.

Finally, we argue that future work in this field will benefit from acknowledging the puzzlement that is a common consequence of being mistaken. This aspect is particularly salient in false belief scenarios designed for children, in which someone finds out that an object is suddenly missing, altered, or unexpectedly present (Yott & Poulin-Dubois, 2012). The agent's epistemic state might be more accurately conceived as an expectation, because unlike beliefs (which can be "free floating"), expectations imply the anticipation of an experience or event (Hacker, 2013). False expectations might be the setting in which children first learn what it means to be mistaken because of the tangible consequences that erring has in this context.

## References

- Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states?

- Psychological Review*, 116, 953–970. doi: 10.1037/a0016923.
- Buttelmann, D., Carpenter, M., & Tomasello, M. (2009). Eighteen-month-old infants false belief understanding is an active helping paradigm. *Cognition*, 112, 337–342. doi: 10.1016/j.cognition.2009.05.006.
- Clements, W., & Perner, J. (1994). Implicit understanding of belief. *Cognitive Development*, 9, 377–395. doi: 10.1016/0885-2014(94)90012-4.
- Dennett, D. C. (1978). Beliefs about beliefs. *Behavioral & Brain Sciences*, 1, 568–570. doi: 10.1017/S0140525X-00076664.
- Fabricius, W. V., Boyer, T. W., Weimer, A. A., & Carroll, K. (2010). True or false: Do 5-year-olds understand belief? *Developmental Psychology*, 46, 1402–1406. doi: 10.1037/a0017648.
- 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, 239–262. doi: 10.1207/S15327647CD0403\_01.
- Hacker, P. M. S. (2013). *The intellectual powers: A study of human nature*. Oxford, UK: Wiley-Blackwell.
- Hadwin, J., & Perner, J. (1991). Pleased and surprised: Children's cognitive theory of emotion. *British Journal of Developmental Psychology*, 9, 215–234. doi: 10.1111/j.2044-835X.1991.tb00872.x.
- Hamlin, K., Wynn, K., Bloom, P., & Mahajan, N. (2011). How infants and toddlers react to antisocial others. *Proceedings of the National Academy of the United States of America*, 108(50), 19931–19936. doi:10.1073/pnas.1110306108.
- Happe, F., & Loth, E. (2002). "Theory of mind" and tracking speakers' intentions. *Mind and Language*, 17(1–2), 24–36. doi: 10.1111/1468-0017.00187.
- Harsh, P. W. (1944). *A handbook of classical drama*. Redwood, CA: Stanford University Press.
- Knudsen, B., & Liszkowski, U. (2012a). Eighteen- and 24-month-old infants correct others in anticipation of action mistakes. *Developmental Science*, 15, 113–122. doi: 10.1111/j.1467-7687.2011.01098.x.
- Knudsen, B., & Liszkowski, U. (2012b). Eighteen-month-olds predict specific action mistakes through attribution of false belief, not ignorance, and intervene accordingly. *Infancy*, 17, 672–691. doi: 10.1111/j.1532-7078.2011.00105.x.
- Kovács, A. M., Téglás, E., & Endress, A. D. (2010). The social sense: Susceptibility to others' beliefs in human infants and adults. *Science*, 330, 1830–1834. doi: 10.1126/science.1190792.
- MacLaren, R., & Olson, D. R. (1993). Trick or treat? Children's understanding of surprise. *Cognitive Development*, 8, 27–46. doi: 10.1016/0885-2014(93) 90003.
- Moll, H., Kane, S., & McGowan, L. (2016). Three-year-olds express suspense when an agent approaches a scene with a false belief. *Developmental Science*, 19(2), 208–220. doi:10.1111/desc.12310
- Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, 308, 255–258. doi: 10.1126/science.1107621.
- Perner, J., & Horn, R. (2003). Ignorance or false negatives: Do children of 4 or 5 years simulate belief with "not knowing = getting it wrong"? *Journal of Cognition and Development*, 4, 263–273. doi: 10.1207/S15327647JCD-0403\_02.
- Perner, J., Leekam, S. R., & Wimmer, H. (1987). Three-year-olds' difficulty with false belief: The case for a conceptual deficit. *British Journal of Developmental Psychology*, 5, 125–137. doi: 10.1111/j.2044-835X.1987.tb01048.x.
- Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 1, 515–526. doi: 10.1017/S0140525X00076512.
- Ruffman, T. (2014). To belief or not belief: Children's theory of mind. *Developmental Review*, 34, 265–293.
- Ruffman, T., & Keenan, T. (1996). The belief-based emotion of surprise: The case for a lag in understanding relative to false belief. *Developmental Psychology*, 32(1), 40–49. doi: 10.1037/h0031267.
- Ruffman, T., Taumoepeau, X., & Perkins, J. (2012). Statistical learning as a basis for social understanding in children. *British Journal of Developmental Psychology*, 30, 87–104. doi: 10.1111/j.2044-835X.2011.02045.x.
- Scott, R. M., He, Z., Baillargeon, R., & Cummins, D. (2012). False-belief understanding in 2.5-year-olds: Evidence from two novel verbal spontaneous response tasks. *Developmental Science*, 15, 181–193. doi: 10.1111/j.1467-7687.2011.01103.x.
- Southgate, V., Chevallier, C., & Csibra, G. (2010). Seventeen-month-olds appeal to false beliefs to interpret others' referential communication. *Developmental Science*, 13, 907–912. doi: 10.1111/j.1467-7687.2009.00946.x.
- Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science*, 18, 587–592. doi:10.1111/j.1467-9280.2007.01944.x.
- Truffaut, F. (1967). *Hitchcock: The definitive study of Alfred Hitchcock*. New York, NY: Simon & Schuster Paperbacks.
- Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development*, 72, 655–684. doi: 10.1111/1467-8624.00304.
- Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13, 103–128. doi:10.1016/0010-0277(83)90004-5.
- Yott, J., & Poulin-Dubois, D. (2012). Breaking the rules: Do infants have a true understanding of false belief? *British Journal of Developmental Psychology*, 30, 156–171. doi: 10.1111/j.2044-835X.2011.02060.x.
- Zillmann, D. (1996). The psychology of suspense in dramatic exposition. In P. Vorderer, W. J. Wulff, & M. Friedrichsen (Eds.), *Suspense: Conceptualizations*,



*theoretical analyses, and empirical explorations* (pp. 199–231). Mahwah, NJ: Erlbaum.

## Appendix

### Reduction Story

Cookie Monster entered with a box of cookies and emptied them on stage, exclaiming: “See how many I have?” He placed the cookies back in the box, and either exited (false belief) or took position on the dowel (true belief). In the ignorance condition, Cookie Monster received a box of cookies from his friend, Elmo, who stated that Cookie Monster could have them later. Both exited. In all conditions, Tony (antagonist) entered and opened the box. He commented on the large number of cookies and mischievously asserted that he will leave only two cookies. He exited with a stack of cookies. In all conditions, Cookie Monster returned from off stage and exclaimed, “It’s time for cookies!” He traversed the stage, grasped the box, and left.

### Deconstruction Story

In the false and true belief conditions, a box was located on stage. Max entered with a Lego® construction, exclaiming: “It took me a long time to make this!” He placed the construction in the box and exited (false belief) or took position on the dowel (true belief). In the ignorance condition, Max received a box with Lego® from his Grandfather, who stated that Max could play with them later. Both exited. In all conditions, Pengu (antagonist) entered and took the construction out of the box. After commenting on how nicely it was put together, he maliciously declared that he will “take this thing apart.” He disconnected the blocks, placed them in the box, and left. In all conditions, Max returned and exclaimed: “I want to play with the Legos now!” He crossed the stage, grasped the box and left.

### Replacement Story

In the false and true belief conditions, a box was located on stage. Franz entered with a ball and commented: “See how big it is?” Franz placed the ball in the box and either exited (false belief) or took position on the dowel (true belief). In the ignorance condition, Franz received a box with a ball inside from his friend Fiona. She set the box down so Franz could play with it later. Both exited. In all conditions, Jen (antagonist) entered, introduced herself, opened the box, and was impressed by the large ball. She clandestinely replaced it with a grimy small ball from her pocket and left. Franz returned and exclaimed: “I want to play with the ball now!” He traversed the stage, grasped the box and left.

### Breaking Story

In the false and true belief conditions, a box was located on stage. Lucy entered and showed her ring, which flashed a light: “Look how it blinks!” After placing it in the box, she exited (false belief) or took position on the dowel (true belief). In the ignorance condition, Lucy received a box with a ring inside from her friend George, who set it down so Lucy could try on the ring later. Both exited. In all conditions, Twyla (antagonist) entered, opened the box, and commented on the ring’s flashing. She maliciously “broke” the light (*E* surreptitiously switched off the light inside the ring), placed the ring back in the box and left. Lucy returned and exclaimed, “I’m going to try on the ring now!” She moved across the stage, grasped the box (“I’m taking this outside”), and exited.

### Supporting Information

Additional supporting information may be found in the online version of this article at the publisher’s website:

**Video S1.** Examples of Children’s Expressions Observed in the False Belief Conditions of the Two Experiments