True Belief Belies False Belief: Recent Findings of Competence in Infants and Limitations in 5-Year-Olds, and Implications for Theory of Mind Development

Joseph A. Hedger · William V. Fabricius

Abstract False belief tasks have enjoyed a monopoly in the research on children's development of a theory of mind. They have been granted this status because they promise to deliver an unambiguous assessment of children's understanding of the representational nature of mental states. Their poor cousins, true belief tasks, have been relegated to occasional service as control tasks. That this is their only role has been due to the universal assumption that correct answers on true belief tasks are inherently ambiguous regarding the level of the child's understanding of mental states. It has also been due to the universal assumption that nothing in the child's developing theory of mind would lead to systematically incorrect answers on true belief tasks. We review new findings that 4- and 5-year-olds do err, systematically and profoundly, on the true belief versions of all the extant belief tasks. This reveals an intermediate level of understanding in the development of children's theory of mind. Researchers have been unaware of this intermediate level because it produces correct answers in false belief tasks. A simple two-task battery—one true belief task and one false belief task—is sufficient to remove the ambiguity from each task. The new findings show that children do not acquire an understanding of beliefs, and hence a representational theory of mind, until after 6 years of age, or 2 years later than most developmental psychologists have concluded. This raises the question of how to interpret other new findings that infants are able to pass false belief tasks. We review these new infant studies, as well as recent studies on chimpanzees, in light of older children's failure on true belief tasks, and end with some speculation about how all of these new findings might be reconciled.

J. A. Hedger (✉)
Philosophy Department, Syracuse University, 541 Hall of Languages, Syracuse, NY 13244, USA
e-mail: jhedger@syr.edu

W. V. Fabricius
Psychology Department, Arizona State University, PO Box 871104, 950 S. McAllister Room 237, Tempe, AZ 85287-1104, USA
e-mail: william.fabricius@asu.edu

🙲 Springer

1 The False Belief Task

False belief tasks have been used as the core test of children's ability to attribute representational mental states and use them to understand behavior since the first study of theory of mind in children (Wimmer and Perner 1983). In one such task, the *unexpected transfer task*, children are presented with the following scenario, normally accompanied by dolls to represent the characters in the story, and props to represent the locations: Maxi comes into the kitchen with a piece of chocolate. He wants to put it somewhere so he can find it when he comes back from playing outside. He puts it into the red cupboard, and then goes outside to play. Later, Maxi's mother comes into the kitchen to clean. She moves the chocolate from the red cupboard to the green cupboard while she is cleaning. Then, Maxi comes inside from playing. He is hungry and wants his chocolate. Subjects are then asked the test question: "Where will Maxi look for his chocolate?" The correct answer is, of course, the red cupboard, since Maxi ought to believe (falsely) that this is where his chocolate is located.

Philosopher Daniel Dennett (1978) first sketched the logic and rationale for false belief tasks.[1] Dennett was commenting on studies of theory of mind in chimpanzees (Premack and Woodruff 1978), in which chimps would often predict that an agent who gave cues to his mental state (e.g., reaching for bananas hung from the ceiling) would subsequently engage in the correct behavior (i.e., get a box to stand on). Chimpanzees appeared to attribute mental states, but they might have simply associated familiar behaviors and outcomes. Dennett argued that an unambiguous test of mental state attribution required the chimpanzee to predict that the agent would engage in an incorrect behavior, which could be accomplished by giving the agent a false belief that the incorrect behavior would accomplish the goal. Wimmer and Perner (1983) adopted Dennett's logic to design the unexpected transfer task, which was the first false belief task for verbal children. Thus the field began with a confidence that low-level mechanisms, such as might be used by chimpanzees, had been controlled for by using false belief tasks.

The *unexpected contents task* is a second false belief task (Hogrefe et al. 1986). The child is shown a familiar candy container, such as an M&M bag, and asked what he thinks is inside. After the child says "M&M's," it is revealed that something unexpected is inside, such as a pencil. After the pencil is placed back inside the bag, the subject is informed that a friend of the experimenter's (named "Elmo" e.g.) is waiting outside. The test question is: "If he just looks at it, what will Elmo think is inside the bag?" The correct answer is to attribute to Elmo a false belief that M&M's are inside the bag.

Children under the age of 4 tend to fail false belief tasks. In the examples above, they tend to choose the green cupboard, because that is where the chocolate currently is, despite the fact that Maxi was outside and could not have known that the chocolate was moved. They also say that Elmo will think that there is a pencil in

---

[1] Gilbert Harman (1978) made a similar suggestion in regards to studying chimpanzee theory of mind, and other philosophers had previously stressed the importance of understanding false belief for having the concept BELIEF. Donald Davidson (1975), e.g. argues that "Someone cannot have a belief unless he understands the possibility of being mistaken, and this requires grasping the contrast between truth and error—true belief and false belief" (p. 170).

the bag, because that is what the bag actually contains, even though Elmo has no reason to think that. We will refer to this as *Reality Reasoning (RR)*. A meta-analysis of 178 false belief studies by Wellman et al. (2001) revealed that the age at which 50% of children pass false belief tasks is 3 years, 8 months.

Importantly, the meta-analysis also revealed that none of the various ways researchers have used to try to make the tasks more understandable or motivating for younger children reliably eliminated RR and allowed younger children to pass at rates comparable to older children. Six variables were found to affect the age at which children passed false belief tasks, but no variation was able to improve the performance of 3 year, 5 month old subjects to chance or greater. Thus, Wellman et al. (2001) argued that the consistent failure of younger children was due to a lack of conceptual understanding. Most psychologists and neuroscientists (e.g. Gopnik and Wellman 1992; Perner 1988, 1991; Saxe et al. 2004) have long agreed that at around 4 years of age children acquire the concept of representational mental states and transition from RR to *Belief Reasoning (BR)*. BR signals a full-fledged theory of mind in which understanding and predicting the behavior of others is accomplished by attributing to them mental states, especially false beliefs.

2 The True Belief Task

Dennett (1978) felt that true belief tasks would be too easy to pass and would not demonstrate anything of theoretical importance. That is true as long as there are only two ways that children could think about mental states—to not think about them at all (RR), or to correctly understand their representational nature (BR)—both of which predict success on true belief tasks. But when Fabricius and Imbens-Bailey (2000) proposed that children might be using what they called *Perceptual Access Reasoning (PAR)* instead of BR when they passed false belief tasks, it introduced an intermediate level of understanding and entailed the counterintuitive prediction that children should fail true belief tasks.

According to the PAR hypothesis (Fabricius and Imbens-Bailey 2000; Fabricius and Khalil 2003; Fabricius et al. 2010), false belief tasks are confounded in a way that allows children to pass without truly understanding false beliefs. In the unexpected transfer task, for example, children might only reason that when Maxi returns he will not see his chocolate in the new location and will therefore not know it is there, and that as a consequence of his ignorance he will "get it wrong" and look in the empty location. But it just so happens that the one "wrong" location is also the location where children who attribute a false belief to Maxi should say he will look. Thus children can pass the task by attributing only ignorance, predicting "wrong" behavior, and using a process of elimination. This line of reasoning does not involve attributing false belief. Likewise children can pass the unexpected contents task by using PAR. They can reason that Elmo cannot see the pencil inside, so he will not know it is a pencil inside, so when given the forced choice (of a pencil or M&M's) he will "get it wrong" and say that M&M's are inside instead. PAR is thus composed of two rules:

Rule 1: Seeing→Knowing (and Not Seeing→Not Knowing)
Rule 2: Knowing→Getting it Right (and Not Knowing→Getting it Wrong)

Importantly, these two rules contain no reference to false beliefs. In the first test of the PAR hypothesis, Fabricius and Khalil (2003) modified all the various false belief tasks (unexpected transfer, unexpected contents, appearance-realty, and representational change) by including a third, irrelevant alternative. Children using PAR should have no reason to choose the false belief alternative over the irrelevant alternative, but children using BR should. In three studies of over 150 5- and 6-year-olds, Fabricius and Kahlil found that, consistent with the PAR hypothesis, many children passed the standard tasks and failed the modified versions by choosing the irrelevant alternatives.[2]

The two PAR rules also contain no reference to true beliefs. The rules predict correct behavior on the basis of the child's decision that the other knows, but this does not presume that the child uses an adult conception of knowing as true belief; namely, as a representational mental state that the person maintains over time and throughout changes in external reality. Instead, knowing in PAR is caused by what the person has perceptual access to in the current situation, and both perceptual access and hence knowing change as the situation changes.[3] Thus PAR entails the prediction that children will judge that a protagonist will get it wrong regardless of whether the protagonist has a false belief or a true belief, as long as the situations in both cases result in comparable lack of perceptual access. To see this consider the various true belief versions of the false belief tasks (unexpected contents, unexpected transfer, and appearance-reality) used in Fabricius et al. (2010).

The true belief version of the contents task began like the false belief version. The child was shown an M&M bag and asked what he thought was inside, and then was shown that it contained a pencil. It became a true belief task when the child watched the experimenter remove the pencil and fill the bag with M&M's before being asked the test question. Children using BR should reason that seeing the familiar candy container will cause Elmo to have a true belief that it contains M&M's. Children using RR should also answer M&M's, but simply because that is what the bag actually contains. Children using PAR should reason that Elmo cannot see the M&M's inside, so he will not know M&M's are inside, so when given the forced choice (of a pencil or M&M's) he will "get it wrong" and say that a pencil is inside. Using PAR, in other words, should cause children to fail the true belief version of the contents task just as it should cause them to pass the false belief version, because the protagonist is equally deprived of perceptual access to the contents in both cases.

Fabricius et al. (2010) used two slightly different true belief versions of the transfer task. In the first (used in Study 1), Maxi watches his mother place the chocolate in the red cupboard before leaving. During his absence, she takes it out and considers moving it to the green cupboard, but then changes her mind and returns the chocolate to its original location. Therefore this true belief transfer task is

---

[2] Perner and Horn (2003) tested the PAR hypothesis using a variation of the three-alternative false belief procedure designed to be simpler than that used by Fabricius and Khalil, and they concluded that their findings refuted the hypothesis. Their study had several unexplained findings, however, which Fabricius and Khalil (2003) argue cast doubt on that conclusion. In addition, Fabricius and Frederick (under review) failed to replicate Perner and Horn's findings against the hypothesis.

[3] There is independent evidence that the child's understanding of 'know' undergoes much development during these years regarding its relation to both perceptual access and correct versus incorrect behavior (e.g., Perner 1991).

similar to the false belief transfer task in that the chocolate is moved during Maxi's absence, but it is not moved to a new location. In the second task (used in Study 2), Maxi watches his mother place his chocolate in the red cupboard, and then continues to watch as she changes her mind and moves it to the green cupboard. Then Maxi leaves and there is no subsequent movement of the chocolate. Hence, the chocolate is moved to a new location just as in the false belief version, but Maxi is present to witness the transfer. Thus, in both true belief tasks, when Maxi returns he should have a true belief that the chocolate is still in the red cupboard. Therefore children using BR should pass both versions of this task. But for children using PAR, when Maxi leaves his perceptual access to the situation is broken, and they do not attribute a belief to him that persists after he leaves the situation. When Maxi returns, children using PAR see him in a new situation, and when asked where Maxi will look, they should reason that since Maxi can't see the chocolate now, he doesn't know where it is, and thus he will "get it wrong" and look in the empty cupboard.

The true belief version of the appearance-reality task that Fabricius et al. (2010) used (in Study 3) involves no hidden objects, and uses tactile instead of visual perceptual information. Children were shown a fake rock and asked what they thought it was. After answering "a rock," they were allowed to feel that it was actually a sponge. Then the sponge was put away and the experimenter brought out a similar-looking real rock, saying, "Here is a real rock." Children were allowed to handle it briefly. They were then asked what another child who only looked at it would think it is. Children using PAR should reason that since the other child will not touch the object, he won't know that it is a real rock, and will get it wrong when asked whether it is a rock or a sponge. Therefore children using PAR should say that when the other child looks at the rock he will think that it is a sponge, while those using either RR or BR should once again pass the task.

The PAR hypothesis specifies that children only analyze the current situation to determine whether someone does or does not have perceptual contact with the object in question, and consequently whether he will be right or wrong about that object. For example, in the true and false belief transfer tasks, Maxi's return prompts children to see him as being in a *new* situation, and to use PAR about this new situation (i.e., "He doesn't see the object, so he doesn't know where it is, so he'll be wrong") without any reference to the prior situation in which Maxi acquired his true or false belief. Conversely, if Maxi did not leave, but simply placed his chocolate in the red cupboard and stayed in the kitchen, the simple disappearance of the chocolate into the cupboard would *not* constitute a new situation. The hypothesis specifies that when children do not decide that the situation has changed they default to their initial conclusion that the protagonist will get it right. While technically speaking it is true that at the moment an object disappears into a container, or a protagonist stops touching an object, he is out of visual or tactile perceptual contact with it, it would make no sense for children to re-apply PAR and conclude that he now doesn't know where or what it is and that he will be wrong about it. There would be no ecological validity to applying PAR on such a moment-by-moment basis. Furthermore, feedback from others' behavior would quickly extinguish any tendency to do so. The PAR hypothesis does not specify the cues children use to decide what constitutes a new situation, and thus what cues "trigger" the PAR rules, nor how these cues are learned (or whether they are innate).

Most importantly, the PAR hypothesis does specify that without a concept of mental representation there is no "place," as it were, for children to put the

information about what Maxi's perceptual access in the prior situation caused him to know, just as there is no place for them to put the information about what the visual appearance of a container or object causes someone to know about what is inside or what it really is. Consequently, the hypothesis should be easy to refute since it entails the unique prediction that children who use PAR will fail true belief tasks.[4] Specifically, it predicts a U-shaped developmental pattern of performance in true belief tasks: 3-year-olds should pass by using RR, 4- to 5-year-olds should fail by using PAR, and 6-year-olds should pass by using BR.

Fabricius et al. (2010) tested this prediction with 108 children in three studies. In Study 1 they found the predicted pattern in the true belief transfer task and in the true belief contents task (see Fig. 1). Note that in Fig. 1 it appears as if 5 1/2-year-olds understand the location task, but not the contents task. Fabricius et al. (2010) demonstrated in Study 2 that this difference is only apparent. The version of the true belief location task used in Study 1 inadvertently had twice as many references to the correct location than the incorrect one in the control questions, which apparently primed 5 1/2-year-olds (but not 4 1/2-year-olds) to choose the correct location. The researchers used the second version of the true belief location task in Study 2 which avoided that problem, and compared the results against a true belief contents task which was identical to the one used in Study 1. In Study 2, 5 1/2-year-olds scored the same on the two tasks (M=65%), which falls favorably along the U-shaped function.
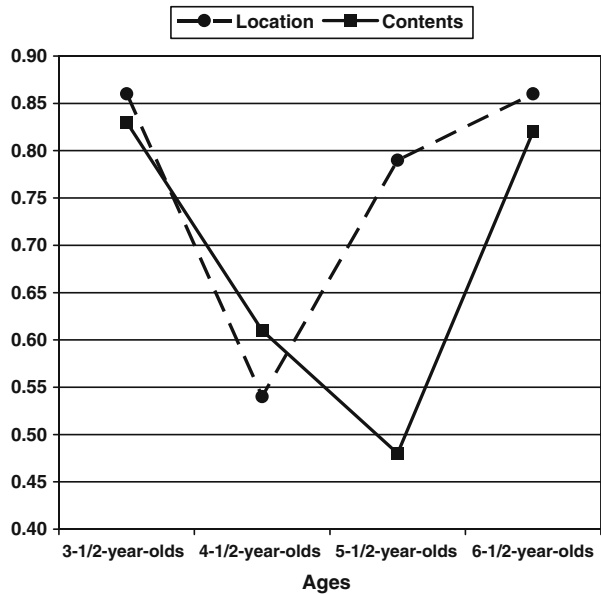
Fabricius et al. (2010) also constructed and tested two methods for determining whether individuals used RR, PAR, or BR. The first method involved giving each child both a false belief and a true belief task. Passing the false belief task and failing the true belief task indicates PAR. Passing the true belief task and failing the false belief task indicates RR. Passing both tasks indicates BR. The second method involved asking children for their justifications for their answers only on the true belief task (e.g., "Why will Elmo think that?"). Passing the task and giving a correct explanation (e.g. "Because it's an M&M bag") indicates BR. Passing the task but giving an incorrect justification (e.g. "Because it has M&M's in it") or no justification indicates RR. Failing the task regardless of justification given is sufficient to indicate PAR, because that is the only approach that gives the incorrect answer to true belief tasks.

The two methods yielded similar rates for each type of reasoning. For example, in Study 3, the first method (classifying children according to their performance on false and true belief tasks) indicated that 13% of 5 1/2-year-olds used RR, 25% used PAR, and 58% used BR (only 4% failed both tasks). The second method (classifying children according to their justifications on the true belief task) obtained similar results: 17% of 5 1/2-year-olds used RR, 29% used PAR, and 54% used BR. The average rates of BR across all three studies from Fabricius et al. (2010) were 10% for 3 1/2-year-olds, 29% for 4 1/2-year-olds, 44% for 5 1/2-year-olds, and 64% for 6 1/2-year-olds.

The PAR hypothesis provides a consistent and principled explanation for the difficulty children who pass the standard, confounded false belief tasks have with 3-alternative

---

[4] Two previous studies reported children's failure on true belief tasks (Friedman et al. 2003; Ruffman 1996), although both sets of researchers assumed that children's difficulty with true belief was unique to the particular tasks used in their studies, and was not indicative of a general pattern of reasoning used across different tasks. See Fabricius et al. (2010) for further discussion.

Fig. 1 Children's performance on true belief tasks. (From Fabricius et al. 2010, American Psychological Association, reprinted with permission.)

versions of false belief tasks (Fabricius and Khalil 2003) and with true belief tasks (Fabricius et al. 2010). It is principled because it provides clear criteria for distinguishing PAR from RR and BR. It is consistent because it provides the same explanation across all the task variations of the extant belief tasks. Tests of the hypothesis to date have also produced consistent results (except as noted for the Perner and Horn (2003) results). In contrast to the meta-analytic data showing that when they turn 5 years of age 75% of children pass the confounded false belief tasks (Wellman et al. 2001), research using un-confounded tasks (Fabricius and Khalil 2003; Fabricius et al. 2010) has consistently found that only about 35% understand belief.

3 Is Theory of Mind Innate? New Considerations in Light of New Findings

There has been a resurgence of a long-standing view (Fodor 1992; Leslie 1987) that despite the failure of 3-year-olds on false belief tasks, humans have an innate mechanism which attributes representational mental states in order to understand and predict behavior. The resurgence began with two studies that found that infants can pass false belief tasks by 13 months (Surian et al. 2007) and 15 months of age (Onishi and Baillargeon 2005). These and most of the subsequent infant studies used the Violation of Expectation (VOE) paradigm. It is known that infants will look significantly longer at novel or unexpected stimuli than at stimuli which are familiar or expected. Thus, if infants expect Maxi to go to one location, then they will look longer if Maxi goes to the other location. Theorists who believe that a theory of mind mechanism is an innate human capacity cite these positive infant findings in support of their view (Leslie 2005; Scott and Baillargeon 2009; Surian et al. 2007). According to this view the failure of 3-year-olds on false belief tasks is not a sign that they lack *competence* in using BR, but is simply a kind of *performance error*

(Chomsky 1965; Fodor 1992) in which something interferes with this competence to cause them to fail false belief tasks. Leslie (1987; Roth and Leslie 1998) hypothesized that true belief is the default attribution of our innate theory of mind mechanism, and that an inability to inhibit this default attribution in verbal tasks is what causes 3-year-olds to fail; in other words, failure on false belief tasks is due to misattribution of true belief. Perhaps infants do not have trouble inhibiting true belief because the infant studies use eye gaze as a measure of understanding instead of verbal reports (Leslie 2005).

However, the new true belief findings complicate this picture. Why would children fail true belief tasks if they have an innate theory of mind mechanism, especially one whose default attribution is that people have true beliefs? According to Roth and Leslie (1998), false belief tasks are more difficult than true belief tasks because of the need to inhibit the default true belief attribution, whereas this is not necessary in true belief tasks. Yet beginning at around age 4 children tend to pass false belief and fail true belief tasks. There is no explanation on offer for why 4- and 5-year-olds who are able to inhibit true beliefs in false belief tasks should inhibit true beliefs in true belief tasks as well. Any such explanation would have to show that the true belief tasks, despite their apparent simplicity and face validity, had some cue, feature, or input that caused 4- and 5-year-olds to mistakenly conclude that the protagonists in those tasks had false beliefs.

To guard against just this possibility that some feature of any one true belief task might inadvertently disrupt children's attributions of true beliefs, Fabricius et al. (2010) used several true belief tasks (two transfer tasks, a contents task, and an appearance-reality task), all with different features. Nevertheless, Study 2 of Fabricius et al. was devoted to testing several task-specific possibilities to insure that their coincidental confluence did not explain the overall pattern of results. These included whether the forced choice question in the true belief contents task, "Will he think it is a pencil or M&M's?" might have seemed odd or arbitrary and misled children to think that "pencil" was an option worth considering, and whether features of the first transfer task might have confused children. Importantly, Study 2 also included a detailed consideration of the possibility that the first transfer task and the contents task required two inhibitions each whereas their false belief counterparts required only one, which would have made it more difficult to correctly attribute beliefs in the former than in the latter (Leslie et al. 2005). There was no evidence that any of these potential factors accounted for children's difficulties in the true belief tasks.

If 3-, 4-, and 5-year-olds do not understand false beliefs, then how are we to account for the recent findings that infants can pass false belief tasks? This prompts the question of what could be the developmental precursor of PAR in young children, to which we turn next.

4 Developmental Precursor of PAR

The recent findings that infants pass false belief tasks, along with other well-established findings that young children who fail explicit false belief tasks unconsciously look first to the false belief location (e.g., Clements and Perner

1994), suggests to us that a PAR-like mechanism might be present in infancy. The two PAR rules (*see/not see→know/not know* and *know/not know→get it right/get it wrong*) are bridged by the concept KNOW, which children begin to acquire and link with perceptual access by about 3 1/2 years of age (e.g., Pillow 1989; Pratt and Bryant 1990). Our hypothesis is that before children acquire the concept KNOW they implicitly use the condensed rule *see/not see→get it right/get it wrong*, which we will call *Rule A*. Our theory of Rule A follows directly from combining the two PAR rules used by preschoolers into a mechanism that could be used by nonverbal organisms. The theory receives some support from findings that have already emerged in studies of infants and chimpanzees (discussed below), but it is not a post hoc attempt to explain those findings. The theory makes testable predictions of new findings that should emerge from new methods with these populations and is therefore, like PAR, potentially easily falsifiable.

Rule A is hypothesized to be a mechanism with the function of producing short-term, situation-embedded predictions of another's behavior in reference to a goal that has been attributed to the other (by some other mechanism). The input to Rule A is a representation of a conspecific with a goal. The mechanism initiates an analysis of the perceptual contact the other has with that goal in the present situation. The output is a prediction of whether the other's behavior will be successful in obtaining the goal or not. In terms of function, input, output, and constraints (below), Rule A is identical to PAR. PAR however, is accessible to conscious processing and incorporates what the child has learned about 'know.'

Two constraints on Rule A make it adaptive. The first is that the input is constrained to be a conspecific with a goal.[5] Rule A is thereby saved from the computational explosion of monitoring everything in the surroundings that the other does and does not see and predicting behavior in relation to every object. This adaptively focuses Rule A on the other's consequential behavior towards important objects.

The second constraint regards the eliciting context. Rule A is situation-specific, and is re-triggered when the subject perceives that the situation has changed, as determined by some other mechanism attuned to threshold cues for situation change. This saves Rule A from re-analyzing the other's perceptual contact with the goal at every head turn or other momentary lapse in the other's perceptual contact. This is adaptive because organisms remember things and do not require constant perceptual contact with a goal in order to obtain it. Furthermore, Rule A *must* be situation-specific, just as PAR must be, because the theory specifies that there is no attribution to the other of a mental representation of the situation. When Rule A is re-triggered by a change in the situation, no representation of the prior situation is attributed to the other.

It is not part of the theories of PAR and Rule A to specify what the cues for situation change are for chimps, infants or young children. Those are empirical

---

[5] For chimpanzees, the input is perhaps further constrained to goals that are shared by the subject, as evidenced by indications that better understanding of others' mental states is elicited in experimental tasks that involve *competition* with conspecifics (e.g., Kaminski et al. 2008). For human infants, the other's goal does not have to be one that the infant also wants, as evidenced by their success in passive, looking time tasks. For instance, studies involving pointing by preverbal infants appears to demonstrate understanding of goals which they don't share and *cooperation* with adults (Liszkowski et al. 2008).

questions. In all cases, however, it would be adaptive for the mechanism that is attuned to cues for situation change to be capable of learning which cues are more reliable. When the situation changes so as to result in the other actually having a false belief about the goal, Rule A will correctly predict that the other will not achieve the goal. But when the situation does not change enough to render the other's belief false (or to cause the other to forget) and Rule A is re-triggered nonetheless, then the subject will incorrectly predict that the other will not achieve the goal. In the case of a subordinate chimp predicting the behavior of a dominant chimp toward a mutually-desired piece of food, that error can be painful. For chimpanzees, learning about the situational cues for re-triggering Rule A would allow individuals to refine their prediction of conspecifics' behavior. For humans it might contribute to the eventual acquisition of a concept of mental representation.

This analysis of the mechanism that re-triggers Rule A leads to an important methodological imperative. A simple two-task battery—one true belief task and one false belief task—is sufficient to distinguish PAR and BR, but *only if* both tasks contain comparable cues for situation change and thus equal likelihood of eliciting Rule A. The infant false belief research has frequently included true belief tasks, but to date the true belief tasks have contained much weaker cues for situation change than the false belief tasks. There are fewer false belief studies of chimpanzees, but they are less open to this concern, and we consider them first.

Consistent with Rule A, chimpanzees in Kaminski et al. (2008; Study 2) showed correct performance on false belief tasks, and incorrect performance on comparable true belief tasks. Three cups were presented to two chimpanzees (subject and competitor), and both watched and observed each other watching the experimenter place a high-quality reward in one cup. Then the competitor's view was blocked (the subject could see this) while the experimenter either lifted and then replaced the reward back into the initial cup (true belief task) or moved it to a different cup (false belief task). Thus the true belief task and the false belief task had comparable cues for situation change because the object was moved out of a container and into a container during the competitor's absence. The competitor was the first to choose, and the subject never witnessed the competitor's choice. After the competitor made her choice, it was the subject's turn to choose. If the subject reasoned that the competitor had gotten it wrong, he should choose the cup where the high-quality reward was last placed; if he reasoned that the competitor had gotten it right, he could choose the "consolation prize," which was a second, low-quality reward that he had seen placed in a fourth cup that was available only to him. Thus in the false belief task the correct choice was the high-quality cup, but in the true belief task that choice was incorrect.

Subjects correctly chose the high-quality cup 60% of the time in the false belief task, and incorrectly chose the high-quality cup 60% of the time in the true belief task. They apparently reasoned, in accord with Rule A, that the competitor had gotten it wrong in both cases because she could not see the reward in the new situation that resulted from the occluded movement. Each task also had a control condition in which the competitor's view was not blocked during the lifting (true belief) or moving (false belief). Thus in these control conditions the subject could see that the competitor had perceptual contact with the high-quality reward up to the point of its being placed in the final cup, and knew that the competitor would be

allowed to search immediately after that. In these control conditions, subjects apparently reasoned (correctly) that the competitor had gotten it right because they chose the high-quality cup less than 45% of the time. Hence the simple disappearance of the object into the container did not re-trigger Rule A, as it shouldn't if Rule A is to be adaptive.

Chimpanzees also performed in accord with Rule A in a different false belief task in Hare et al. (2001; Study 1). These researchers did not include a true belief task, but they did include an ignorance task. The competitor was a dominant chimp and the subject was a subordinate chimp. In the false belief ("Misinformed") condition, the experimenter hid food in one of two containers, and both witnessed and observed each other witnessing the hiding. The experimenter then moved the food to the other container, but the dominant's view was occluded (the subordinate could see this) and only the subordinate watched it being moved. In the control condition (similar to Kaminski et al. 2008) they both continued to watch as the experimenter moved the food to the other container. Subordinates were then given a slight head start, so that they had to predict where the dominants would go. Subordinates went for the food more often in the false belief condition than in its associated control condition. Subordinates performed essentially the same in an ignorance ("Uninformed") condition and its associated control. They predicted that the dominants would get it wrong in both the false belief and ignorance conditions, and that the dominants would get it right in the control conditions. They apparently reasoned, in accord with Rule A, that the competitor would get it wrong in both cases because she had not seen the occluded final placement.

As predicted, Rule A appears to be situation-specific in chimpanzees. The evidence is the difference in responses to the control versus belief conditions across the two studies. The control conditions in both studies are similar because the competitor saw the final hiding and there was *no* subsequent movement of the reward occluded from the competitor's view before she was allowed to search. In these control conditions the subjects tended to reason that the competitor would get the reward. These control conditions differ from the belief conditions, in which after the competitor saw the hiding there *was* movement of the reward that was occluded from the competitor's view. In response to this movement subjects tended to reason that the competitor would not get the reward. Thus when the reward simply disappears from view momentarily (control conditions) it is not a cue that the situation has changed, in which case chimpanzees default to their initial conclusion that the competitor will get it right. This is adaptive because competitors in such situations are in fact likely to get it right given chimpanzees' evolved memory skills for the location of food. When there is some movement of the reward that the competitor does not witness, regardless of whether the reward switches locations (false belief) or not (true belief), this is a cue that the situation has changed. Rule A is re-applied to this new situation to predict that the competitor will get it wrong.

The true belief tasks in the infant false belief studies to date have contained much weaker cues for situation change than the false belief tasks. Consequently the true belief tasks serve as controls for low level responding but do not allow strong inferences that infants attributed beliefs to agents. The first two studies (Onishi and Baillargeon 2005; Surian et al. 2007) included true belief tasks that resembled the chimpanzee control conditions (Hare et al. 2001; Kaminski et al. 2008); i.e., the

object simply disappeared from the agent's view momentarily and she immediately searched for it. To support the argument that infants attribute true and false beliefs, it needs to be shown that they pass a true belief task in which there is some interruption in the agent's connection to the situation that is comparable to what occurs in the false belief task but that preserves the agent's true belief. Kaminski et al. (2008) met this requirement in location tasks with chimpanzees by incorporating an occluded, irrelevant movement of the reward. Fabricius et al. (2010) did the same thing with older children. Both Fabricius et al. (2010) and Friedman et al. (2003) met the requirement in location tasks with older children by having the protagonist simply leave the situation and return (see Fabricius et al. 2010 for further discussion). Without meeting this methodological requirement, studies cannot determine whether infants are using Rule A (re-triggered in the false belief task but not in the true belief task) versus attributing beliefs.[6]

Two recent studies suggest that infants attribute false beliefs based on misleading appearances (Scott and Baillargeon 2009; Song and Baillargeon 2008); these also do not use true belief tasks capable of distinguishing BR from Rule A use. In Song and Baillargeon (2008) two objects were hidden in two containers. One object was a doll with blue hair, and 14.5-month-olds were familiarized to expect that the agent wanted that object. In the false belief task, it was hidden during the agent's absence in a plain container while the other object was hidden in a container that had blue hair sticking out from under the lid. Infants were surprised when the agent got it right and searched in the plain container. In the true belief task the objects were hidden in the same way but with the agent present the whole time. Thus the target object simply disappeared from the agent's view momentarily and she immediately searched for it, again just like in the chimpanzee control conditions. The proper, and simple, true belief task is to hide the doll in the container with the blue hair, because Rule A reasoning would predict the agent would get it wrong and search in the plain container, while BR would predict she would search in the container with the blue hair.

In Scott and Baillargeon (2009) there were two identical penguins, one disassembled and one preassembled, and 18-month-olds were familiarized to expect that the agent wanted the one that was disassembled so she could put a key inside it. In the false belief task, infants observed (in the agent's absence) while that penguin was first assembled and then placed in a transparent box, and the preassembled one was placed in an opaque box. Seeing an assembled penguin in the transparent box would lead the agent to falsely believe that the disassembled one was in the opaque box. Toddlers appeared to attribute such a false belief to the agent because they looked longer when the agent got it right and chose the transparent box than when she got it wrong and chose the opaque box. The true belief task again involved the same transformations with the agent present. In this study however, the true belief task involved "hiding" the correct object in a transparent box so that it never disappeared from sight, even momentarily. It is especially difficult to call this one a

---

[6] Southgate et al. (2007) pointed out that infants in those two studies could have used a specific rule applicable to transfer tasks that they presumably had learned from other experiences (i.e., that agents who do not witness an object being moved will look in the wrong place for it). As the authors tacitly acknowledge, the true belief tasks in those studies do not control for that possibility either.

"true belief" task; it only requires infants to remember which penguin the agent wanted. The proper true belief task would have simply involved placing the preassembled penguin in the transparent box in the agent's absence, because Rule A would predict the agent would get it wrong and search in the transparent box, while BR would predict she would search in the opaque box.

He et al. (2011) tested 2 1/2 year olds. In the false belief task, after familiarization, the agent watched as a toy frog was placed into one of two identical containers by the experimenter. Then a curtain occluded the agent's view as the experimenter counted to ten and moved the toy frog to the other container. Thereafter the curtain slid back to allow the agent to view the scene. The agent exclaimed "Here it is!" as she pointed to one of the two containers. As both BR and Rule A predict, 2 1/2-year-olds looked longer whenever the agent pointed to the container where the frog was currently located, as opposed to when she pointed to the container from which it had been moved. However, in the true belief task, once again the agent watched as the experimenter moved the frog from one container to the other, so that it disappeared from the agent's view only momentarily and she immediately searched for it. Ironically, this time the adequate formulation of the true belief task was used as a familiarization trial (not as a test trial) for the false belief task. That familiarization trial was identical to the false belief task except that the experimenter did not move the frog as she counted, but simply left it in the initial container.

Two recent studies have tested toddlers in situations that elicited helping and understanding the referential communication of agents with false beliefs (Buttelmann et al. 2009; Southgate et al. 2010). As in the infant studies, in both of these studies the agents in the true belief tasks watched the objects' movements, did not leave the situation, and responded immediately. In Buttelmann et al. (2009), for example, the agent needed to unlock a box to get a toy. In the false belief task he was absent when the toy was switched from the box where he had placed it to the other one. He returned to the original box and struggled to open it. Children tended to help him by unlocking the other box. In the true belief task he watched as the object was moved and then immediately struggled to open that box. In this case children helped him open that one. Rule A would lead children to conclude that in the false belief task he was acting incorrectly and therefore in need of help to open the other box with the toy in it. But in the true belief task Rule A would not be re-triggered by the momentary disappearance of the toy into the box, and children would not expect him to be acting incorrectly.[7]

Four recent studies using ignorance tasks, in which the agent has no belief, have found evidence against Rule A in infants and young children. Rule A predicts that in ignorance tasks children should expect agents to get it wrong, yet these studies found evidence to the contrary. In Scott and Baillargeon's (2009) ignorance condition, 18-month-olds watched while the 2-piece penguin was put together and both were placed in identical containers (either both transparent or both opaque) during the

---

[7] An interesting study by Song et al. (2008) indicates that Rule A can be re-triggered for 18-month-olds by witnessing an agent receive verbal communication about a situation. In their task, an agent hid a ball in a box and was absent when it was moved it to a cup. When she returned, if the experimenter told her "The ball is in the cup!" children expected her to get it right and search in the cup; otherwise, they expected her to get it wrong. The statement apparently signaled a new situation and elicited a re-application of Rule A, just as if the agent had gotten direct (visual, tactile) perceptual access to the situation.

agent's absence, so that the agent had no belief about which one was the desired 2-piece penguin. Children did not look longer when the agent chose the correct container, suggesting that they did not expect that the ignorant agent would get it wrong. The authors were concerned that toddlers might have forgotten which one was the correct container when the agent returned because the containers were identical, unlike in the false belief and true belief conditions in which one was transparent and the other opaque. To test that possibility, the authors ran another condition in which the containers were identical but the agent had a true belief; i.e., the two penguins were placed in identical containers during the agent's presence. In this condition toddlers looked longer when the agent chose the incorrect container, indicating they remembered which one was the correct container. But demonstrating that infants remembered when the agent was present for the identical-container hiding cannot establish that they also remembered when the agent was absent. When the agent was present she watched the correct penguin, and toddlers' appreciation that visual perception establishes a connection between the looker and the external world might have helped make the correct penguin's location more salient (cf. Poulin-DuBois et al. 2007). Scott et al. (2010; Study 3) used these same general task features, also with 18-month-olds; i.e., an ignorance condition with identical objects and, to address concerns about memory, a follow-up condition with identical objects but with the agent present during the hiding. A simple solution to these concerns about memory would be to use non-identical containers.

In He et al. (2011) the test trials of the ignorance task were unusual among all the other familiarization and test trials of the three tasks (false belief, knowledge and ignorance). In the familiarization trials of all three tasks, the object was placed in the first location and remained there while the agent waited. In the test trials of both the false belief and knowledge tasks, the object was moved in one motion from the first to the second location. But only in the test trials of the ignorance task, the object was removed from the first location and then held in abeyance between the two locations giving the agent time to look at it before the screen was closed and it continued on its journey to the second location. This suspension procedure was unique to the test trials of the ignorance task and so was likely to have been novel to children. The novelty could have interrupted their memory for the location of the object.

In Friedman and Petrashek (2009), 4- and 5-year-olds were told that a girl, who never entered the scene, did not know which of two boxes her dog was under. Contrary to Rule A, children did not predict that she would get it wrong. However, the findings were also counter to the authors' predictions that children would choose randomly between the two boxes. Instead, children mysteriously predicted that she would get it right. We suspect that rather than reverting to RR, these 4- and 5-year-olds were simply not considering the other box as an equally likely place for the dog to hide. They must have assumed that the girl was likely to know where her dog usually hid. They were told, "her dog likes to hide under these boxes," which implies that it was a familiar occurrence.

There are other studies with toddlers (e.g., Koenig and Echols 2003; Poulin-DuBois et al. 2007) that find that children do expect ignorant agents to get it wrong, so clearly more attention to methods is needed. But we will conclude with the one study that poses the most challenge to the theory of Rule A (Southgate et al. 2007) and the one that provides some preliminary, tentative support (Brandone and Wellman 2009).

Southgate et al. tested 2-year-olds in the only anticipatory looking (AL) design in this group of studies. This study appears to provide the best evidence that children were not using Rule A, but instead might have shown sensitivity to false beliefs. The agent witnessed the object disappearing into one of two containers, and then turned away. The object ultimately left the scene so there was no reality location and therefore no unique wrong location; both locations were wrong and thus consistent with Rule A. Rather than directing their anticipatory looking randomly toward the two locations when it was signaled that the agent was about to search, children looked longer at the location corresponding to the agent's false belief. One explanation is that in the absence of a competing reality response, children were able to demonstrate belief reasoning. This could be further tested with an AL true belief condition (which the authors did not include), in which case the reality response should also not compete with belief reasoning. If the true belief condition were comparable to the false belief condition along the dimensions we have been arguing for here, then Rule A would have to predict that children would expect the agent to get it wrong in the true belief condition.

Brandone and Wellman (2009) used the only true belief task so far with infants in which the agent's connection to the situation was interrupted in a way comparable to the false belief task. They achieved this by having the agent simply leave the situation and return. They reported preliminary data that infants performed the same in the true belief task and the false belief task, expecting the agent to get it wrong in both. In Brandone and Wellman's ignorance condition, however, infants did not expect the ignorant agent to get it wrong. Agent 1 had previously been shown repeatedly hiding and finding the object. Then in the absence of any agent the object disappeared into one container, after which a new, ignorant agent (Agent 2) showed up unexpectedly and searched for the object. Perhaps infants were distracted by the new agent and forgot where the object was. This could be remedied by having the ignorant agent appear at the beginning and then leave while the object was hidden. This would also meet the requirement of a comparable interruption in the agent's connection to the situation in all three conditions (ignorance, true belief, and false belief).

5 Conclusions and Conjectures

Contrary to what philosophers initially thought, the true belief task is a crucial piece of the puzzle as we gain a greater understanding of the psychological mechanisms of theory of mind, and its developmental trajectory. The new findings that children fail true belief tasks for 2 years complicate the position (e.g. Gopnik and Wellman 1992; Perner 1988, 1991; Saxe et al. 2004; Wellman et al. 2001) that theory of mind undergoes a one-step conceptual change from no understanding of representational mental states (RR) to understanding the representational nature of mental states (BR). Development now appears more protracted than that. The new findings also complicate the position (e.g. Fodor 1992; Leslie 2005; Scott and Baillargeon 2009; Surian et al. 2007) that infants are able to attribute false beliefs and thus are sensitive to representational mental states. If infants are able to attribute representational

mental states, then some non-obvious explanation is required for the protracted development of that ability which we now see between 3 and 6 years of age. Proponents of the two positions currently argue about who is right, but it may be too soon to tell. Both have to confront the new findings that 4- and 5-year-olds fail what was supposedly the easiest belief task.

The PAR hypothesis predicted children's difficulty with true belief tasks, and the implication that children use PAR for an extended period of time before they acquire BR may lead us to a more complex and nuanced understanding of development. Neither the infant competence view nor the 3- to 4-year-old stage shift view provide much fuel for developmental theorizing. The PAR hypothesis accords a role for something that may be innate that we may also share with our chimpanzee relatives (Rule A), and it also reveals two intriguing developments; i.e., how infants' Rule A develops into 4- and 5-year-olds' verbal PAR, and how verbal PAR develops into 6- and 7-year-olds' BR. It appears that more work needs to be done at both ends of this developmental continuum.

The PAR hypothesis also invites us to conjecture that older children and even adults might continue to use Rule A. This could be tested by using the "eye gaze" methodology developed for children under 4 years of age (Clements and Perner 1994; Garnham and Perner 2001; Garnham and Ruffman 2001; Ruffman et al. 2001), in which they have passed standard, confounded false belief tasks by showing unconscious anticipatory looking to the correct location. However, in order to know whether correct anticipatory looking in the false belief task indicates attribution of false beliefs or use of Rule A, the methodology needs to include a true belief task in which there is some interruption in the agent's connection to the situation that is comparable to what occurs in the false belief task. The previous eye gaze studies have not included such true belief tasks. It would certainly be surprising if adults showed anticipatory looking to the incorrect location in true belief tasks, but it would be readily interpretable from the PAR hypothesis. It would indicate that Rule A continues to exist in adults, perhaps in some type of module-like structure.[8]

Rule A predicts the future behavior of conspecifics based on an analysis of their perceptual access, but it might serve different functions for infants and chimpanzees. It might be used for deception by chimpanzees such as hiding food from a dominant competitor with the expectation that if he doesn't see it he won't come and get it (Hare et al. 2001; Kaminski et al. 2008). In human infants, it might be used for communication and cooperation. For instance, Liszkowski et al. (2008) describe "infants' early pointing at 12 months [as] already premised on an understanding of others' knowledge and ignorance, along with a prosocial motive to help others by providing needed information." But our analysis of Rule A suggests that infant communication need not depend upon the attribution of knowledge and ignorance per se but only upon prediction of behavior from perceptual access, and suggests that the "uniquely human skills and motivations for cooperation and shared intentionality (e.g., joint intentions and attention with others)" (Liszkowski et al. 2008) may have their computational base in a Rule A mechanism shared with chimpanzees.

At 3 years of age, children's correct anticipatory looking in confounded false belief tasks might indicate that Rule A is present implicitly when children use

---

[8] On modularity and psychology, see Fodor (1983) and Barrett and Kurzban (2006).

explicit RR. At 4 years of age, Rule A would then develop into explicit PAR with the simultaneous (partial) acquisition of the concept KNOW and the differentiation of Rules 1 and 2 from Rule A. Explicit use of PAR may allow 6-year-olds to become aware of its limitations (i.e., inability to distinguish between a true belief, a false belief, and a wrong response), which may in turn motivate the construction of a concept of mental representation and the acquisition of BR.

If adults were found to continue to use Rule A implicitly it could explain a puzzling feature of theory of mind—that at times we appear able to make judgments about the mental states of others quickly, automatically, and effortlessly, while at other times the process is difficult and deliberative. There is evidence that BR is effortful and difficult for adults (Apperly et al. 2006; Keysar et al. 2003). Lin et al. (2010) found that higher working memory capacity can have a positive impact on adult performance in theory of mind tasks, while cognitive load impairs this ability. Perhaps this can be explained by BR and Rule A being different psychological mechanisms, the former explicit and effortful and the latter implicit, automatic and modular. In fact, persons with Asperger syndrome may be able to learn BR or PAR while lacking Rule A. For instance, Senju et al. (2009) found that adults with Asperger syndrome did not demonstrate anticipatory looking toward the correct location during the false belief task, although they verbally passed the task. Neurotypical adults *did* demonstrate anticipatory looking toward the correct location during the false belief task (which they also verbally passed), but a true belief task was not used in this study. Future experiments studying the eye gaze of 5–6 year olds, adults, and autistics during the false and true belief tasks will provide evidence concerning Rule A's nature and development and its relation (if any) to autism spectrum disorders.

References

Apperly, I.A., K.J. Riggs, A. Simpson, C. Chiavarino, and D. Samson. 2006. Is belief reasoning automatic? *Psychological Science* 17: 841–844.

Barrett, H.C., and R. Kurzban. 2006. Modularity in cognition: Framing the debate. *Psychological Review* 113: 628–647.

Brandone, A. C., and H.M. Wellman. 2009. *Concepts of ignorance and false belief in 15-month-old infants*. Poster presented at the 6th Biennial Meeting for the Cognitive Developmental Society, San Antonio, Texas.

Buttelmann, D., M. Carpenter, and M. Tomasello. 2009. Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition* 112: 337–342.

Chomsky, N. 1965. *Aspects of the theory of syntax*. Cambridge: The MIT Press.

Clements, W.A., and J. Perner. 1994. Implicit understanding of belief. *Cognitive Development* 9: 377–395.

Davidson, D. 1975. *Inquiries into truth and interpretation*. Oxford: Oxford University Press.

Dennett, D.C. 1978. Beliefs about beliefs. *Brain and Behavioral Sciences* 1: 568–570.

Fabricius, W.V., and A.L. Imbens-Bailey. 2000. False beliefs about false beliefs. In *Children's reasoning about the mind*, ed. P. Mitchell and K. Riggs, 267–280. Hove: Psychology.

Fabricius, W.V., and S.L. Khalil. 2003. False beliefs or false positives? Limits on children's understanding of mental representation. *Journal of Cognition and Development* 4: 239–262.

Fabricius, W.V., T. Boyer, A.A. Weimer, and K. Carroll. 2010. True or False: Do five-year-olds understand belief? *Developmental Psychology* 46: 1402–1416.

Fodor, J.A. 1983. *The modularity of mind*. Cambridge: The MIT Press.

Fodor, J.A. 1992. A theory of the child's theory of mind. *Cognition* 44: 283–296.

Friedman, O., and A.R. Petrashek. 2009. Children do not follow the rule "ignorance means getting it wrong". *Journal of Experimental Child Psychology* 102: 114–121.

Friedman, O., R. Griffen, H. Brownell, and E. Winner. 2003. Problems with the seeing = knowing rule. *Developmental Science* 6: 505–513.

Garnham, W.A., and J. Perner. 2001. Actions really do speak louder than words—but only implicitly: Young children's understanding of false belief in action. *British Journal of Developmental Psychology* 19: 413–432.

Garnham, W.A., and T. Ruffman. 2001. Doesn't see, doesn't know: Is anticipatory looking really related to understanding of belief? *Developmental Science* 4: 94–100.

Gopnik, A., and H. Wellman. 1992. Why the child's theory of mind really is a theory. *Mind & Language* 7: 145–171.

Hare, B., J. Call, and M. Tomasello. 2001. Do chimpanzees know what conspecifics know? *Animal Behaviour* 61: 139–151.

Harman, G. 1978. Studying the chimpanzee's theory of mind. *Behavioral and Brain Sciences* 1: 576–77.

He, Z., M. Bolz, and R. Baillargeon. 2011. False-belief understanding in 2.5-year-olds: Evidence from violation-of expectation change-of-location and unexpected-contents tasks. *Developmental Science* 14: 292–305.

Hogrefe, G.J., H. Wimmer, and J. Perner. 1986. Ignorance versus false belief: A developmental lag in attributing epistemic states. *Child Development* 57: 567–582.

Kaminski, J., J. Call, and M. Tomasello. 2008. Chimpanzees know what others know, but not what they believe. *Cognition* 109: 224–234.

Keysar, B., S. Lin, and D.J. Barr. 2003. Limits on theory of mind use in adults. *Cognition* 89: 25–41.

Koenig, M., and C.H. Echols. 2003. Infants' understanding of false labeling events: The referential roles of words and the speakers who use them. *Cognition* 87: 179–208.

Leslie, A.M. 1987. Pretense and representation: The origins of 'theory of mind'. *Psychological Review* 94: 412–426.

Leslie, A.M. 2005. Developmental parallels in understanding minds and bodies. *Trends in Cognitive Sciences* 9: 459–462.

Leslie, A.M., T.P. German, and P. Polizzi. 2005. Belief-desire reasoning as a process of selection. *Cognitive Psychology* 50: 45–85.

Lin, S., B. Keysar, and N. Epley. 2010. Reflexively mindblind: Using theory of mind to interpret behavior requires effortful attention. *Journal of Experimental Social Psychology* 46: 551–556.

Liszkowski, U., M. Carpenter, and M. Tomasello. 2008. Twelve-month olds communicate helpfully and appropriately for knowledgeable and ignorant partners. *Cognition* 108: 732–739.

Onishi, K.H., and R. Baillargeon. 2005. Do 15-month-old infants understand false beliefs? *Science* 308: 255–258.

Perner, J. 1988. Developing semantics for theories of mind: From propositional attitudes to mental representation. In *Developing theories of mind*, ed. J.W. Astington, P.L. Harris, and D.R. Olson, 141–172. New York: Cambridge University Press.

Perner, J. 1991. *Understanding the representational mind*. Cambridge: The MIT Press.

Perner, J., and Horn, R. 2003. Knowledge or false negatives: Do children of 4 to 5 years simulate belief with "Not knowing = getting it wrong?" *Journal of Cognition and Development* 4: 263–273.

Pillow, B.H. 1989. Early understanding of perception as a source of knowledge. *Journal of Experimental Child Psychology* 47: 116–129.

Poulin-Dubois, D., B. Sodian, U. Metz, J. Tilden, and B. Schoeppner. 2007. Out of sight is not out of mind: Developmental changes in infants' understanding of visual perception during the second year. *Journal of Cognition and Development* 8: 401–425.

Pratt, C., and P. Bryant. 1990. Young children understand that looking leads to knowing (so long as they are looking into a single barrel). *Child Development* 61: 973–982.

Premack, D., and G. Woodruff. 1978. Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences* 1: 515–526.

Roth, D., and Leslie, A.M. 1998. Solving belief problems: Toward a task analysis. *Cognition* 66: 1–31.

Ruffman, T. 1996. Do children understand the mind by means of simulation or a theory? Evidence from their understanding of inference. *Mind and Language* 11: 388–414.

Ruffman, T., W. Garnham, A. Import, and D. Connelly. 2001. Does eye gaze indicate implicit knowledge of false belief? Charting transitions in knowledge. *Journal of Experimental Child Psychology* 80: 201–224.

Saxe, R., S. Carey, and N. Kanwisher. 2004. Understanding other minds: Linking developmental psychology and functional neuroimaging. *Annual Review of Psychology* 55: 87–124.

Scott, R.M., and Baillargeon, R. 2009. Which penguin is this? Attributing false beliefs about object identity at 18 months. *Child Development* 80: 1172–1196.

Scott, R.M., R. Baillargeon, H. Song, and A.M. Leslie. 2010. Attributing false beliefs about non-obvious properties at 18 months. *Cognitive Psychology* 61: 366–395.

Senju, A., V. Southgate, S. White, and U. Frith. 2009. Mindblind eyes: An absence of spontaneous theory of mind in asperger syndrome. *Science* 325: 883–885.

Song, H., and R. Baillargeon. 2008. Infants' reasoning about others' false perceptions. *Developmental Psychology* 44: 1789–1795.

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

Southgate, V., A. Senju, and G. Csibra. 2007. Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science* 18: 587–592.

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

Surian, L., S. Caldi, and D. Sperber. 2007. Attribution of beliefs by 13-month-old infants. *Psychological Science* 18: 580–586.

Wellman, H.M., D. Cross, and J. Watson. 2001. Meta-analysis of theory-of-mind development: The truth about false belief. *Child development* 72: 655–684.

Wimmer, H., and J. Perner. 1983. Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition* 13: 103–128.