CHILD DEVELOPMENT

Child Development, March/April 2015, Volume 86, Number 2, Pages 486–502

Explicit Theory of Mind Is Even More Unified Than Previously Assumed: Belief Ascription and Understanding Aspectuality Emerge Together in Development

Hannes Rakoczy, Delia Bergfeld, Ina Schwarz, and Ella Fizke
University of Göttingen

Existing evidence suggests that children, when they first pass standard theory-of-mind tasks, still fail to understand the essential aspectuality of beliefs and other propositional attitudes: such attitudes refer to objects only under specific aspects. Oedipus, for example, believes Yocaste (his mother) is beautiful, but this does not imply that he believes his mother is beautiful. In three experiments, 3- to 6-year-olds' (N = 119) understanding of aspectuality was tested with a novel, radically simplified task. In contrast to all previous findings, this task was as difficult as and highly correlated with a standard false belief task. This suggests that a conceptual capacity more unified than previously assumed emerges around ages 4–5, a full-fledged metarepresentational scheme of propositional attitudes.

It is commonly assumed that children acquire a full-blown explicit concept of mental states around age 4 when they begin to master standard verbal tasks that require the ascription of false beliefs (FBs) to other agents. Competence in a variety of such tasks emerges around the same time and is systematically synchronized and correlated—suggesting that it is in fact a fundamental conceptual revolution children are undergoing (Perner, 1991; Perner & Roessler, 2012; Wellman, Cross, & Watson, 2001). The standard interpretation of this revolution is that children at this age acquire an explicit metarepresentational conception of propositional attitudes such as beliefs, desires, and so on (Gopnik & Wellman, 1992; Perner, 1991; Perner & Roessler, 2012; Wellman, 2011; Wellman et al., 2001). However, this assumption has been questioned on both theoretical and empirical grounds.

The Intensionality of Propositional Attitudes

Theoretically, it has been argued that a subject can only be credited with a full-blown conception of propositional attitudes if she understands the essential logical structure of the ways that propositional attitudes represent the world (Fabricius, Boyer, Weimer, & Carroll, 2010; Lalonde & Chandler, 2002). In particular, an essential feature of propositional attitudes is their aspectuality: They represent objects and situations always and necessarily *under some aspect or description* only (Anscombe, 1957; Brentano, 1874/1973; Searle, 1983; for an overview, see McKay & Nelson, 2014). We may believe, to take a famous example, that the Morning Star rises in the morning without thereby believing that the Evening Star (in fact one and the same thing as the Morning Star) rises in the morning. Technically, this phenomenon of reports about propositional attitudes is called intensionality (with an "s"): In most regular, *extensional* contexts, the substitution of coreferential terms does not alter the truth value of a sentence.

(1) Clark Kent weighs 80 kg.
(2) Superman weighs 80 kg.

If (1) is true, then (2) is true, and if (1) is false then (2) is false; (2) follows from (1)—together with the premise Clark Kent = Superman. In contrast, in some contexts, notably propositional attitude reports, substitution of coreferential terms in the content of a reported propositional attitude can alter the truth value of the whole sentence:

This work was supported by the German Initiative of Excellence. The authors thank to Kira Sagolla, Katharina Kentsch, and Alexander Dieball for help with testing and coding.

Correspondence concerning this article should be addressed to Hannes Rakoczy, Department of Developmental Psychology, Courant Research Centre "Evolution of Social Behaviour," University of Göttingen, Waldweg 26, D-37073 Göttingen, Germany. Electronic mail may be sent to hannes.rakoczy@psych.uni-goettingen.de.

 0009-3920/2015/8602-0011
DOI: 10.1111/cdev.12311

(3) Linda believes that Clark Kent weighs 80 kg.
(4) Linda believes that Superman weighs 80 kg.

Here (4) does not follow from (3)—plus the premise Clark Kent = Superman. Even if (3) is true, (4) may well be false (Linda may simply be unaware of the identity of Clark Kent and Superman). Such contexts are called "intensional."

Measuring the Developing Understanding of Intensionality

Empirically, tasks tapping children's understanding of intensionality have been found to be much more difficult than standard FB tests (Apperly & Robinson, 1998, 2003; Kamawar & Olson, 1999, 2009, 2011; Russell, 1987; Sprung, Perner, & Mitchell, 2007). In the most direct approach to this topic, children were asked whether one could say that a protagonist who was looking for someone under one description was looking for that person under some other description as well. For example, children were told (5)–(7) and were asked "Can we say that (8)?" (correct answer: no)

(5) The protagonist is looking for the thief.
(6) The thief is the man with the curly hair.
(7) The protagonist does not know that (6).
(8) The protagonist is looking for the man with the curly hair.

Some subsequent studies, in order to reduce the linguistic demands, have used action prediction questions modeled more closely after standard FB tasks (Apperly & Robinson, 1998; Sprung et al., 2007): Children were told about a protagonist looking for an A (e.g., a rubber). There were two objects in two boxes: one was a plain A; the other was both an A (a rubber) and a B (a dice). The protagonist only knew the latter as a B (one had to feel the object and use it as a rubber in order to know it as such, and the protagonist had only seen the object).

(9) There is an A in Box 1.
(10) There is a B in Box 2.
(11) The B in Box 2 is also an A.
(12) The protagonist knows that (9) and (10), but does not know that (11).

Test question: The protagonist is looking for an A. Where will he go to find an A? (correct answer: Box 1)

From an empirical point of view, despite differences in their surface structure, virtually all of these studies have produced converging evidence that understanding intensionality is more difficult and develops considerably later than competence on standard FB tasks. Four- and 5-year-olds typically master standard FB tests yet fail intensionality tasks, and only from around 6 to 7 years old children have been found to reliably solve the variety of tasks tapping intensional problems.

Interpreting the Developmental Findings From Intensionality Tasks

From a theoretical point of view, there are a number of options how to interpret these findings and, consequently, how to explain the development of understanding intensionality in relation to theory-of-mind capacities measured in more standard FB tasks.

Option 1: Existing intensionality tasks measure children's understanding of the fact that representations only hold under some aspect more generally.

One possibility is that existing intensionality tasks measure understanding of the aspectuality of propositional attitudes and are thus a valid measure of a full-blown metarepresentational conception of propositional attitudes (Apperly & Robinson, 1998, 2003).

Option 2: (Almost all) existing intentionality tasks are difficult due to some extraneous performance factors

An alternative option is that existing intensionality tasks are difficult not because they measure understanding of intensionality but due to other extraneous factors involved in the tasks. Different versions of explanations stress different potential factors of this kind.

Option 2a: Ambiguity, reference resolution, and other linguistic demands

One such explanation has it that existing tasks have failed to uncover early competence due to linguistic and other task demands. In particular, even studies using action prediction measures with less direct linguistic demands might have posed other performance problems such as the need for the resolution of ambiguous reference. When told that there in an A in Box 1 and an object that is both an A and a B in Box 2 and that the protagonist, not knowing the B in Box 2 is also an A, is looking for an A, the child has to engage in reference resolution: Both objects are As, so "an A" refers to both of them, but since the protagonist knows of only one of the As that it is an A it must be this object she will be looking for under the description of "an A." It might thus be that children find such tasks

difficult not because of the intensionality but because of the ambiguity and the reference resolution it requires—which, as we know, even adults often and find difficult (Keysar, Lin, & Barr, 2003).

Option 2b: Higher order perspective taking

Another possibility is that virtually all existing intensionality tasks are difficult because they require higher order perspective taking—perspective taking nested in another form of perspective taking (Sprung et al., 2007). According to this possibility, the first level of perspective taking required in such tasks is to adopt the perspective of the protagonist (in particular, his beliefs). The second level of perspective taking involved in most tasks is due to the fact that usually the protagonist has only partial knowledge of a person or an object (the thief/man with curly hair, the rubber/pen, etc.). This partial knowledge pertains to the sortal (kind) concepts used to refer to the object (e.g., "rubber" vs. "pen"). Sortal concepts themselves, the account goes, present conceptual perspective problems in a way other concepts (such as concepts pertaining to an object's properties) do not do (for details, see Perner, Brandl, & Garnham, 2003; Perner, Mauer, & Hildenbrand, 2011). Roughly speaking, and using Perner and colleagues' mental file card metaphor, the reason is the following: when a sortal concept ("firefighter") is used to pick out an object/person, a new mental file card is opened up that refers to the individual picked out specifying it as the kind of thing given by the sortal ("firefighter"). On this file card, additional information about the individual can be represented ("is brave," "has fought against many fires," etc.). Crucially, however, when the same individual (assume the firefighter is the elementary school teacher) is picked out under other circumstances under a different sortal ("elementary school teacher"), another file card is opened up, which in turn can contain information about the individual under this sortal description ("gives boring classes," "requests too much coursework," etc.). Young children can then easily reason about the object under each sortal description and use the information available as premises. However, what they cannot do that easily is reason about the identity relation of the two sortal descriptions—because this would require the representation of the relation between the two file cards, and this relation can only be captured in metarepresentational terms (one has to represent the file cards as mental representations). So, according to this account, the capacity to understand that two sortal descriptions both refer to a given individual, since it constitutes a metarepresentational problem, should emerge together with other metarepresentational capacities, typically indicated in mastery of the FB task. Empirical evidence for this claim comes from two lines of studies: First, it has recently been found that children's understanding of sortal identity statements emerges at the same time as and correlates with mastery of the FB task (Perner et al., 2011). In a typical scenario of these studies, children would witness a toy character referred to under one description (e.g., the firefighter), would then learn something about someone not visible under a different description (e.g., that Peter has lost his bag), and finally learn that the two descriptions refer to the same person (Peter is the firefighter) and are asked to draw corresponding inference (e.g., whom, out of several toy characters, among them the firefighter/Peter, to give the bag).

Second, children have been found to fail sortal alternative naming games until they master the FB task as well (Perner, Stummer, Sprung, & Doherty, 2002): In such tasks, there are two sortal characterizations of an individual both of which children know (e.g., "dog" and "animal") and the task is simply to respond to the partner's use of one sortal with the other one (importantly, children of this age have been found to have no problem in solving property control versions of the task, in which there are two property predicates, e.g., "is big" and "is blue" and one has to use the one the other person has not used). The specific failure only in the sortal version of the task is interpreted as showing that young children, while being able to reason about an object under one sortal description or under another, cannot explicitly relate these two descriptions to each other.

As a consequence of this analysis, according to Perner and colleagues' embedded perspective account it is not the intensionality per se that makes virtually all existing intensionality tasks hard, but the fact that the intensionality pertains to two sortal concepts. Alternative versions of intensionality tasks without such sortal contrasts using, say, two property descriptions, will then only present regular perspective taking problems of first rather than higher order and will thus be easier—in fact as easy as standard perspective taking tasks.

Which Interpretation Is Correct?

Existing findings leave open which of the interpretations is the correct one. Regarding the contrast between 2b and the other options, there is only one finding so far that seems to be inconsistent with Option 1 and to speak in favor of Option 2b: In a recent study, 4- and 5-year-olds found an action

prediction task (modeled after Apperly & Robinson, 1998) difficult when the partial knowledge pertained to the sortal description of the object (rubber/dice), but found it easy—as easy as a standard FB task—when it pertained to two property descriptions (is red, is blue; Sprung et al., 2007).

Regarding the contrast between Option 1 and Option 2, more generally, what is lacking so far are systematic attempts to reduce the extraneous performance factors of intensionality tasks. The question whether Option 1 or some form of Option 2 is correct has important broader theoretical ramifications for the question of what develops in children's theory of mind: If Option 1 is in fact correct and 4-year-olds can master FB tasks without any understanding of intensionality, without understanding that subjective representations hold only under some aspects or descriptions, this would suggest that the theory-of-mind capacities measured by standard FB tasks are less comprehensive and less sophisticated than generally assumed. Four-year-olds' conceptual grasp of beliefs would be severely limited, falling short of a full-blown metarepresentational understanding of propositional attitudes including their aspectuality. Strictly speaking, if this was true, one would have to qualify or even withdraw the claim that a proper understanding of propositional attitudes emerges at this age at all. If, in contrast, Option 2 turned out to be correct, and if the capacity to solve standard FB tasks and modified intensionality tasks with reduced performance factors turned out to be of comparable complexity and to be intimately related (such that once children pass FB tasks they also understand intensionality), this would yield radically different conclusions: It would corroborate the standard interpretation of the 4-year cognitive revolution as the emergence of a truly metarepresentational conception of propositional attitudes.

Rationale of the Present Studies

In order to test these open questions, we devised intensionality tasks with radically reduced performance factors, modeled in structure after standard FB tasks as closely as possible and inspired thematically by standard examples of aspectuality. Imagine, for example, that Eve sees Clark Kent going into his house and somewhat later Superman flying out of this house to the beach. Where will she think Clark Kent is? Crucially, if she does not know about the identity, she will think he is still in the house. This basic narrative structure was realized in our tasks in the following way:

(13) There is an object that is both an A and a B.
(14) The protagonist does not know (13).
(15) The protagonist encounters the object as an A, puts it in Box 1 and leaves the room.
(16) After her return, the protagonist sees the object—as a B—being transferred to Box 2.

Test question: The protagonist is looking for the A. Where will she look for the A? (correct answer: Box 1)

In two control conditions, similar action prediction questions were asked, but now under the aspect available to the protagonist ("Where will she look for the B?") or under the aspect not perceptually available ("Where will she look for the A?"), but in situations where the protagonist knew about the dual identity.

In order to test Option 2b specifically, closely following Sprung et al. (2007), we systematically contrasted cases in which the two descriptions A and B constituted a sortal contrast ("an A" vs. "a B") with cases in which the two descriptions referred to two properties of an object of the same kind ("red sock" vs. "blue sock"). In order to test the relation of understanding intensionality and belief ascription measured in classical theory-of-mind tests, standard FB tasks (first and second order) were also administered. By testing children in the age range in which they have been found to come to master standard FB tasks of first and second order and intensionality tasks (around ages 4–6), the different theoretical options can be tested against each other by exploring the relative difficulty and contingencies between the tasks: Option 1 predicts that standard first-order FB tasks will be significantly easier than all intensionality tasks (i.e., that there are some children solving the former but failing the latter and no/few children showing the opposite pattern). Option 2b predicts that standard first-order FB tasks will be as difficult as (and correlated with) property intensionality tasks, and both will be easier than second-order standard FB and sortal intensionality tasks. Finally, Option 2a predicts that suitably modified intensionality tasks, both sortal and property ones, will be as difficult as (and correlated with) standard first-order FB tasks, all of which will be easier than second-order standard FB tasks.

Study 1

Participants

Sixty 4- to 5-year-olds (48–69 months, $M = 59$; 29 male) from mixed socioeconomic backgrounds

were recruited from a databank of children whose parents had previously given consent to experimental participation. Children were randomly assigned to one of three conditions (see below), resulting in three samples, each with $n = 20$ and a mean age of 59 months and equal gender distributions (10/10 or 9/11). Children were tested by a pair of female experimenters in a quiet room of their day care.

Design and Procedure

The basic design concerning the intensionality tasks was a 3 (condition: intensional—extensional—true belief control) × 2 (contrast: sortal—property) design, with the former as between-subjects and the latter as within-subjects factor. Each child received four trials in total, two trials of sortal and two trials of property tasks (order counterbalanced across subjects). In addition, each child got a standard change-of-location FB task (Wimmer & Perner, 1983) at the end of the session.

Intensionality Tasks. The basic logic and set up of the three intensionality conditions are depicted in Figure 1 (for details, see Appendices A and B). Like in a standard FB task, an object was put into a box (Box 1) and then transferred to another box (Box 2) and the crucial test question was where the protagonist would look for the object. In contrast to a standard FB task, the protagonist did witness the transfer of the object. But in the crucial ("intensional") condition, the protagonist did not know that the object she saw being transferred (under aspect B) was identical to the object she was looking for (under aspect A). In order to solve this task, the subject thus had to take into account under which aspect the protagonist was looking for the object ("A") and whether she knew that both aspects ("A" and "B") actually pertained to the same object she saw being transferred—reasoning like this: "She is looking for the A but does not know that the B that was just moved to Box 2 is the A, so she thinks the A is still in Box 1 and will look for it there." One control condition ("extensional") was exactly matched to the intensional condition with the exception of the test question, which was now where the protagonist would look for the object under the B-aspect. The other control condition ("true belief") was likewise exactly matched to the intensional condition including the test question, with the exception that the protagonist knew about the dual identity of the object.

The following control and test questions were administered:

1. *Control Question 1*: Does the protagonist know that the A is also a B?
2. *Control Question 2*: Where did we put the A in the beginning?
3. *Control Question 3*: Where is the A now?
4. *Test questions*: Where will she look for her . . . [A] (intensional/true belief condition)/[B] (extensional condition)?

No feedback was given after each trial. The boxes were removed and E1 brought out a pair of boxes to start the new trial. Across trials and across subjects, it was counterbalanced which of two boxes was used as Box 1/ Box 2 and where Box 1 was placed (left–right).

Standard FB Task. After the last trial of the intensionality tasks, E1 helped Susi to retrieve the object and transformed it back to her A-aspect. This object was then used (only under the A-aspect) for a standard FB change-of-location task (Wimmer & Perner, 1983) always administered at the end of the session: Susi put the object into one box and left. E1 then transferred it in her absence to the other box. Upon her return, E1 asked two control questions ("Where did she put the A [e.g., the bunny] in the beginning?" and "Where is the A now?") and test question ("Where will she look for the A?").

Results

Intensionality Tasks

Control questions and consistency across trials. Children answered the control questions correctly in 92% of the trials of the intensionality tasks (and in 95% of the standard FB tasks; the following results refer to the whole sample, but the same results hold when those trials are removed from the analyses in which one or more control questions were answered incorrectly). The consistency in performance of children over Trials 1 and 2 was moderate to high in the intensional condition (Φs = .38 and .63 for the sortal and the property contrasts, respectively), but tended to be low to moderate in the other two conditions (extensional: Φs = .36 and .13, for the sortal and the property contrasts, respectively; true belief: Φs = .02 and .21, for the sortal and the property contrasts, respectively).

Main analyses. The mean number of trials in which children answered that Susi would look for the object in Box 1 as a function of conditions and type of contrast is depicted in Figure 2. A 3 (condition: intensional—extensional—true belief) × 2 (contrast: sortal—property) mixed factors analysis of

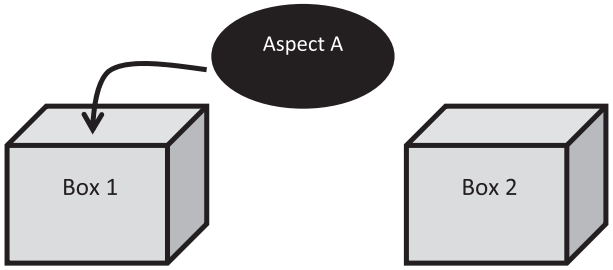
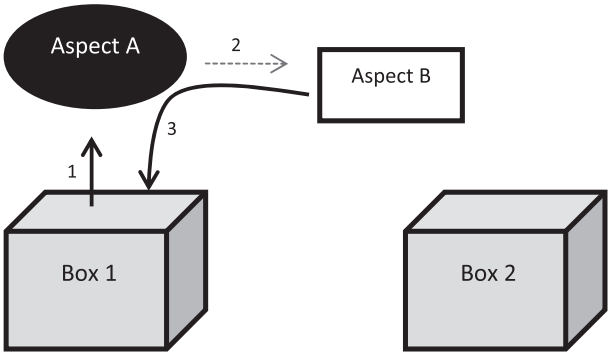
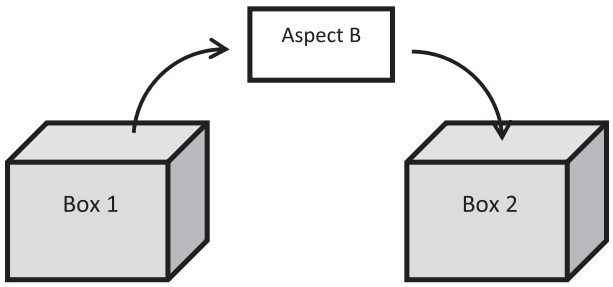
		intensional	extensional	true belief
1	Familiarization with the objects: child learns that the object has two aspects and how to turn one into the other 			
2		Protagonist present	Protagonist present	Protagonist present
3		Protagonist **absent**	Protagonist **absent**	Protagonist **present**
4		Protagonist present	Protagonist present	Protagonist present
test	Where will protagonist look for…	…[aspect A]?	…aspect [B]?	…aspect [A]?

Figure 1. Basic experimental logic and set up of the three conditions in Study 1. Regular arrows signify spatial transfer (in and out of the boxes). Dashed arrows (signify transformations of the object (turning it inside out) to reveal its A- or B-aspect.

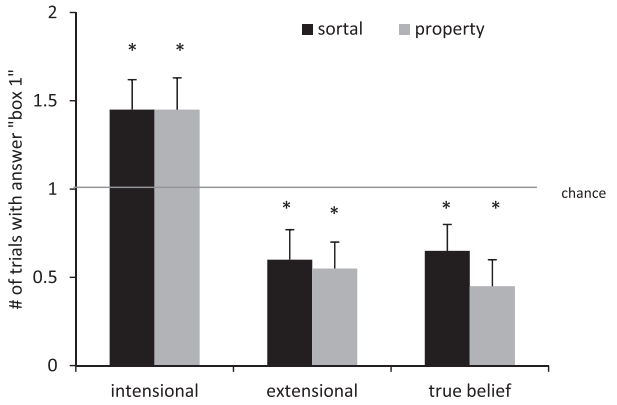


Figure 2. Mean performance (number of trials with answer "Location 1") in the different tasks in Study 1.

variance (ANOVA) revealed a main effect of condition, $F(2, 57) = 11.73$, $p < .001$, $\eta_p^2 = .29$. Post hoc Tukey least significant difference tests showed that this main effect was due to the fact that the intensionality condition differed both from the extensional and the true belief conditions (ps < .001), which did not differ from each other. There was no significant effect of contrast type, $F(1, 57) = 1.18$, $p = .28$, and no interaction effect, $F(2, 57) = .68$, $p = .54$.

Comparisons against chance performance showed that children answered "Box 1" significantly more often than expected by chance in the two intensional conditions: sortal, $t(19) = 2.65$, $p < .05$, $d = .59$, and property, $t(19) = 2.44$, $p < .05$, $d = .55$, and significantly less often than expected by chance in the extensional conditions: sortal, $t(19) = 2.37$, $p < .05$, $d = .53$, and property, $t(19) = 2.93$, $p < .05$, $d = .66$, and the true belief conditions: sortal, $t(19) = 2.33$, $p < .05$, $d = .52$, and property, $t(19) = 3.582.44$, $p < .05$, $d = .80$.

Intensionality and FB Tasks

Of the 60 children, 36 answered the standard FB task correctly (15 in the intensional condition, 11 in the extensional condition, and 10 in the true belief condition), which were not significantly different from each other, $\chi^2(2, N = 60) = 2.91$, $p = .23$.

Differences between intensionality tasks and the standard FB task. In order to test whether the intensionality tasks (of which there were two trials per condition) were more difficult than the standard FB task (of which there was only one trial), the first trial of each intensionality condition was compared to the standard FB test (the rationale for this comparison is that the first trial of a given task presumably presents the purest measure, unspoiled by learning effects; however, the same results hold for separate analyses in which the second trial of each intensionality condition is compared to the standard FB task). The contingency patterns of correct–incorrect answers can be seen in Table 1. There were no significant differences between the intensionality tasks and the standard FB in the intensional conditions (sortal: $p = 1$; property: $p = 1$, McNemar's tests), the extensional conditions (sortal: $p = 1$; property: $p = .38$), and the true belief condition (sortal: $p = .58$; property: $p = .15$).

Correlations between intensionality tasks and the standard FB task. The correlations of the sum scores (over two trials) of correct answers in each intensionality condition with each other and with the standard FB task are depicted in Table 2. In the intensionality condition, as can be seen from the table, not only were the sortal and the property tasks closely related, but also both were strongly related to the standard FB task.

Discussion

Study 1 had four main results. First, 4- and 5-year-old children performed competently in the intensionality tasks: They were sensitive to the aspect under which a protagonist was looking for an object, inferring that she would be looking in the empty location (Box 1) when she did not know that the object she had seen being transferred under Aspect B was actually the A-aspect object she was

Table 1
Contingency Pattern of Sortal and Property Tasks (Trial 1) and Standard False Belief (FB) Task as a Function of Condition in Study 1

	Intensional				Extensional				True belief			
	Sortal		Property		Sortal		Property		Sortal		Property	
	Fail	Pass	Fail	Pass	Fail	Pass	Fail	Pass	Fail	Pass	Fail	Pass
Standard FB												
Fail	3	2	3	2	2	7	1	8	2	8	1	9
Pass	3	12	2	13	6	5	4	7	5	5	3	7

Table 2
Raw Correlations (Pearson Correlation Coefficients and Partial Correlations Correcting for Age) Between the Sortal and Property Task and the Standard False Belief (FB) Task as a Function of Condition in Study 1

	Intensional			Extensional			True belief		
	Sortal	Property	Standard FB	Sortal	Property	Standard FB	Sortal	Property	Standard FB
Sortal		.84** (.83**)	.51* (.49**)		.75** (.75**)	−.33 (−.31)		.36 (.34)	−.23 (−.17)
Property			.61** (.59**)			−.44* (−.42)			−.37 (−.31)

*$p < .05$. **$p < .01$.

looking for. In contrast, in the control conditions, when the protagonist was looking for the object under the B-aspect, or knew about the identity of both aspects, children claimed correctly that she would be looking in the right place (Box 2). Second, in contrast to previous studies, the crucial intensionality task (both sortal and property) was no more difficult than a standard FB task. Third, the two types of tasks were strongly correlated. And fourth, in contrast to one recent finding (Sprung et al., 2007), there was no evidence that the kind of aspectual contrast (sortal vs. property) made a difference to intensionality tasks.

These findings taken together thus clearly speak in favor of Option 2a discussed earlier—that previous intensionality tasks might have produced false negatives due to extraneous performance factors. However, the present study has some methodological limitations that call for caution when interpreting these findings: (a) The order of the intensionality tasks and the standard FB task was not counterbalanced, which complicates the comparison of the two types of tasks. (b) There was only one trial of the standard FB task in contrast to two trials of the intensionality tasks. (c) Only a first-order standard FB task was used. Given that one account (Option 2b above) explicitly claims that some (sortal) intensionality tasks measure second-order theory of mind, second-order FB tasks should be administered as well. (d) The correlations did not control for covariates such as verbal ability.

Study 2

Method

In order to overcome the limitation of Study 1, therefore, children's performance in intensionality tasks (both with sortal and with property contrasts) was investigated in relation to their performance in standard FB tasks, with the following modifications: The order of intensionality and standard FB tasks was counterbalanced. There were two trials in each condition of the intensionality and standard FB tasks. In addition to a standard first-order standard FB task, a second-order FB task was also administered. Verbal ability was tested as a potential covariate in partial correlations between the tasks.

Performance in the crucial intensional condition needs to be interpreted against the background of relevant control conditions to rule out that children always and invariably would give one kind of answer ("Box 1"), even under circumstances under which it would not be appropriate. As the control conditions of the intensionality task (extensional; true belief) in Study 1 had clearly ruled out such an alternative, we focused here on the target intensional condition.

Participants

Twenty-three 3- to 6-year-olds (43–72 months, $M = 56$; 13 male) from mixed socioeconomic backgrounds were recruited from a databank of children whose parents had previously given consent to experimental participation. Children were tested by a pair of female experimenters in a quiet room of their day care.

Design and Procedure

Verbal Ability. At the beginning of the session, children completed a vocabulary test (the vocabulary subscale of the Kaufman Assessment Battery for Children; Kaufman & Kaufman, 1999).

Intensional Tasks. Each child received the exact same four trials of the intensional tasks (two sortal, two property; order counterbalanced) as in Study 1.

Standard FB Tasks. In addition, children received two trials of standard FB tasks with a first- and second-order FB test question (after Perner & Howes, 1992; Sprung et al., 2007). In our adapted version of the task, Susi the puppet put an object (a regular object without any dual identity such as a marble) into Box 1 and left. E1 then transferred the object in her absence to the other box. Upon her return, E1

asked two control questions ("Where did she put the marble in the beginning?" and "Where is the marble now?") and two test questions:

Test question first order: "If we ask Susi: Where is the marble? What will she say?" [correct: Box 1]
Test question second order: "If we ask Susi: Do you know where the marble is? What will she say?" [correct: yes]

The second-order test question was only asked when children had answered the first-order questions correctly since the second-order question is only meaningful when the first-order question has been solved. Children who do not even pass the first-order question could give the correct answer to the second-order questions for totally wrong reasons: Without a concept of an FB, they would answer "Box 2" to the first-order question, assuming that people cannot fail to believe the truth, and then they would naturally answer "yes" to the second-order question. Children who failed the first-order question of a given trial were thus automatically scored as failing the second-order task as well. Children received two trials (one involving a marble as object hidden and transferred, the other one involving a toy horse), resulting in scores of 0–2 for first- and second-order FB ascription, respectively. Across trials and across subjects, it was counterbalanced which of two boxes was used as Box 1/Box 2 and where Box 1 was placed (left–right).

Results

The data of all 23 children in the intensionality tasks were entered into the analyses. Regarding the standard FB tasks, the data of 22 children were available (the data of 1 child could not be analyzed due to an experimental error).

Differences Between Intensionality and Standard FB Tasks

First of all, the consistency in performance of children over Trials 1 and 2 was very high for all tasks (Φs > .61). The mean number of trials with correct answers in the different intensionality and standard FB tasks are depicted in Figure 3. A one-factorial ANOVA with task as within-subjects factor (intensional sortal—intensional property—FB first order—FB second order) revealed a significant difference between the tasks, $F(3, 21) = 12.31$, $p < .001$, $\eta_p^2 = .37$. Post hoc t tests showed that this difference was due to the fact that the second-order standard FB test was significantly more difficult than the other three tasks, ts(21) > 3.91, ps < .05 (Bonferroni corrected), ds > .84, which did not differ from each other, ts(22) < .81, ps > .43. T tests against chance performance showed that children performed significantly above chance in the two intensional tasks and in the first-order standard FB task, ts (22) > 2.24, ps < .05, ds > .47, and significantly below chance in the second-order FB task, t (21) > 2.49, $p < .05$, $d = .53$.

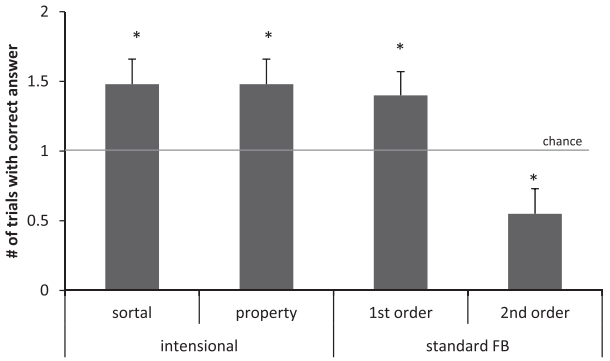


Figure 3. Mean performance (correct answers) in the different tasks in Study 2.

Correlations Between Intensionality and Standard FB Tasks

The correlations between the intensionality and the standard FB tasks are depicted in Table 3. As can be seen from the table, the two intensionality tasks and the first-order FB correlated strongly with each other, and these correlations remained high even after controlling for age and verbal ability.

Discussion

Study 2 thus replicated and extended the basic findings of Study 1 so that the two studies together supply converging evidence that intensionality tasks, when suitably modified, are as difficult as and strongly correlated with standard FB tasks (and much easier than second-order FB tasks). However, there is one fundamental concern regarding the kinds of stimuli used in Studies 1 and 2: What is conceptually crucial for testing children's understanding of intensionality is that there are objects to which two aspects/descriptions apply simultaneously (such that another protagonist might know the object under one but not the other). This clearly applies to Superman/Clark Kent, Yocasta/Oedipus's mother, the thief/man with curly hair, rubber/dice, and so on. It applies

Table 3
Raw Correlations (and Partial Correlations Controlling for Age and Verbal Ability) Between the Intensionality and the Standard False Belief (FB) Tasks in Study 2

	Intensional sortal	Intensional property	Standard FB first order	Standard FB second order
Intensional sortal		.68** (.46*)	.81** (.70**)	.29 (.07)
Intensional property			.69** (.51*)	.13 (−.11)
Standard FB first order				.37† (.22)

†$p < .10$. *$p < .05$. **$p < .01$.

to the objects used here in some sense as well (the reversible toy is both a toy rabbit and a toy carrot), but it is possible that children made sense of these objects in simpler ways: Rather than seeing an object as an object with both A- and B-aspects, they might have seen it as somehow shifting identities over time (at Time 1 it is a carrot, then at Time 2 it turns into a rabbit, at Time 3 it is a carrot again . . .), or they might have seen it as an object with neither A- nor B-aspects (it is a piece of cloth that at some times looks like a carrot, and at other times like a rabbit). While theoretically possible, these two alternatives seem implausible since they would predict that children would fail the control question, "Where is the A now?"—yet children answered this question virtually always correctly. A third alternative, however, would be consistent with children's mastery of the control questions, namely, that children might have seen the objects as really a B-object with another A-object hidden inside it (it is really a toy rabbit, which has a toy carrot hidden inside it). As long as such a simpler construal cannot be ruled out, it remains unclear what the present findings actually show.

Study 3

Method

Study 3, therefore, followed up on the previous studies with stimuli that were more closely matched to those used in previously published intensionality studies such that they did not allow such simpler construals on the part of the children.

Participants

Twenty-six 3- to 5-year-olds (43–64 months, $M = 54$; 16 male) from mixed socioeconomic backgrounds were recruited from a databank of children whose parents had previously given consent to experimental participation. Children were tested by a pair of female experimenters in a quiet room of their day care.

Design and Procedure

The general design and procedure were the same as in Study 2. The only exceptions were the following: First, in the intensionality tasks, objects were used that were theoretically more suitable for testing children's understanding of intensionality. Second, since the sortal and property conditions consistently produced the same results in the previous two studies, this distinction was dropped as an independent variable and only sortal contrasts were used now (sortals are the theoretically more interesting case for present purposes since according to Option 2b they should present more complex second -order perspective-taking problems and therefore be more difficult than standard first-order FB task). Two new kinds of intensionality tasks were developed with objects that more unambiguously had two identities and that therefore were more clearly suited for testing children's understanding of others' representing these objects under one versus the other description (see Figure 4 and Appendix C).

Dual-Function Intensionality Tasks. The objects in these tasks—closely modeled after dual-function objects used in Apperly and Robinson (1998) and Sprung et al. (2007)—had one obvious function (e.g., pen) and another function that could only be detected upon closer inspection (e.g., the pen was also a rattle—which could only be found out by rattling). The subject knew about both identities, but the protagonist only knew about the obvious one. The general logic and procedure of the tasks was similar to those of Studies 1 and 2 with the following exception: E moved the object under the nonobvious aspect (the aspect the protagonist was ignorant about) from Box 1 to Box 2. Since this had

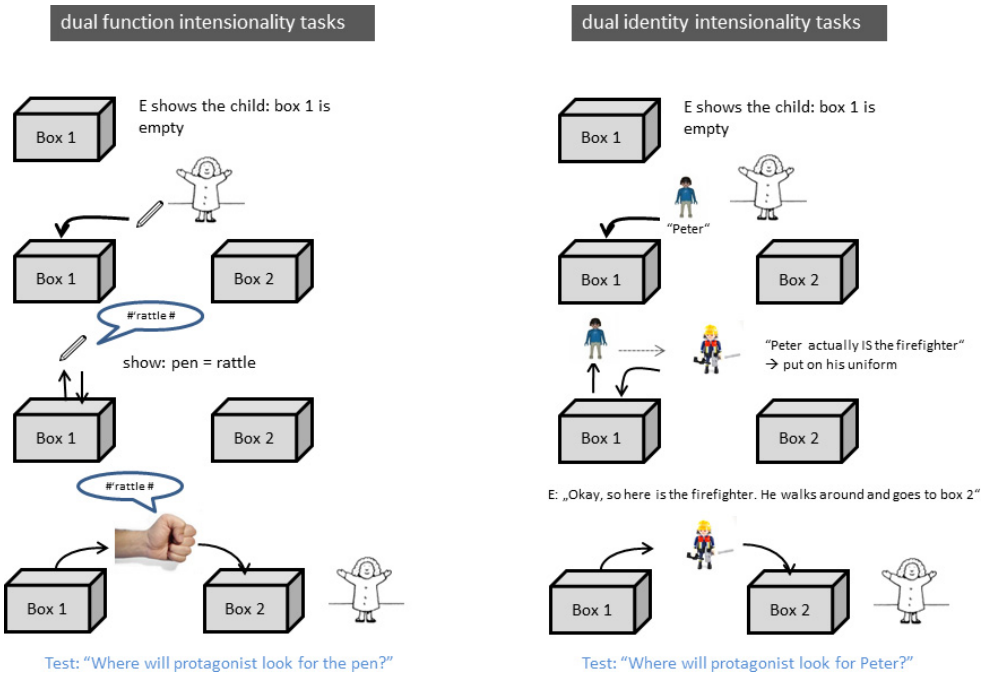


Figure 4. Basic experimental logic and set up of the intensionality tasks in Study 3.

to be done in such a way that the protagonist would not become perceptually aware of the fact that the two functions pertain to the same object, E moved the object hidden in her hand while demonstrating and describing the nonobvious function ("Look, this is the rattle" [rattles]. "I'll move it to Box 2").

Dual-Identity Intensionality Tasks. The objects in these tasks were inspired by standard aspectuality examples such as Clark Kent/Superman. Toy figurines were used that (much like Clark Kent vs. Superman) had two visibly distinguishable identities: For example, one toy figurine was introduced as Peter (A) without any uniform and as the firefighter (B) with his firefighter uniform—where, crucially, the firefighter-uniform-wearing toy figurine by itself would not be recognizable as Peter (just like Superman who by himself is not recognizable as Clark Kent). The general logic and procedure of the tasks, inspired by and matched also to the procedure of Perner et al. (2011), was very similar to those of Studies 1 and 2: The toy figurine as Peter (A) would enter Box 1, then in the absence of the protagonist (ignorant about the dual identity A = B) put on his uniform (see "Peter is also the firefighter"), and then upon the protagonist's return, would move as the firefighter (B) from Box 1 to Box 2.

In both dual-function and dual-identity tasks, the same kinds of control questions (whether the protagonist knew that A = B, where A had been in the beginning and where it was now) and test questions ("Where will the protagonist look for A?") as in Studies 1 and 2 were asked. Each child received two dual-function intensionality tasks, two dual-identity intensionality tasks, two first-order standard FB, and two second-order standard FB tasks (order counterbalanced across children).

Results

Differences Between Intensionality and Standard FB Tasks

First of all, the consistency in performance of children over Trials 1 and 2 was high for all tasks (Φs > .43). Children answered the control questions correctly in 100% of the cases in the intensionality and FB tasks (due to an experimental error, one control question was forgotten in one trial of two children who otherwise also answered all control questions correctly). The mean number of trials with correct answers in the different intensionality and standard FB tasks are depicted in Figure 5. A one-factorial ANOVA with task as within-subjects factor (dual-function intensionality—dual-identity intensionality—F first order—FB second order) revealed a significant difference between the tasks, $F(3, 25) = 28.60$, $p < .001$, $\eta_p^2 = .53$. Post hoc t tests

showed that this difference was due to the fact that the second-order standard FB test was significantly more difficult than the other three tasks, ts(25) > 5.35, ps < .05 (Bonferroni corrected), ds > 1.05, which did not differ from each other, ts(25) < 1.44, ps > .16. T tests against chance performance showed that children performed significantly above chance in the two intensional tasks and in the first-order standard FB task, ts(25) > 2.74, ps < .05, ds > .54, and significantly below chance in the second-order FB task, $t(25) = 3.35$, $p < .05$, $d = .66$.

Correlations Between Intensionality and Standard FB Tasks

The correlations between the intensionality and the standard FB tasks are depicted in Table 4. As can be seen from the table, the two intensionality tasks and the first-order FB correlated strongly with each other, and these correlations remained high even after controlling for age and verbal ability.

Discussion

The results of Study 3 closely replicated the findings of Studies 1 and 2—but now with stimuli that were carefully modeled after standard textbook examples and after those stimuli used in previous intensionality studies and were therefore less ambiguous and open to alternative construals. But are there not still some alternative construals children could have applied to those objects, solving the task in simpler ways? Theoretically, for one type of task (dual identity), one could wonder whether—like in the worries concerning the stimuli of Studies 1 and 2—children might have construed the relation between A and B as one of spatial containment: There is Peter and there is the firefighter, where "firefighter" is misunderstood as only referring to the uniform, and the former is conceived of as spatially in the latter. But it seems very implausible, to say the least, that children would come up with such exotic misconstruals: First of all, there is no a priori reason children suffer from such confusions regarding the ontological status of uniforms. Second, everyday experience and recent experimental findings (Perner et al., 2011) strongly suggest children this age have no problems in understanding such mundane identities as, "Peter is the firefighter." Third, if children reconstrued the relation of A and B as mere containment, they should have found this task easier than (and performance should have been independent from) tasks for which such construals are not possible. Now, the dual-function tasks used here are clearly such that these kinds of containments construals seem out of the questions, and still performance in the two types of tasks was comparable and strongly correlated.

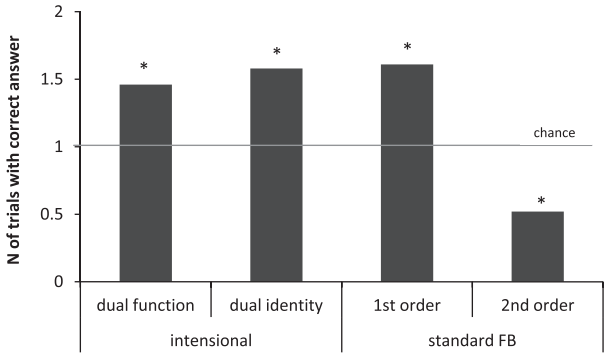


Figure 5. Mean performance (correct answers) in the different tasks in Study 3.

General Discussion

Summary of Findings

The present studies had five main findings: (a) Children aged 3–5 were competent at solving simplified intensionality problems, taking into account under which aspect a protagonist was representing an object she was looking for. Two control condi-

Table 4
Raw Correlations (and Partial Correlations Controlling for Age and Verbal Ability) Between the Intensionality and the Standard False Belief (FB) Tasks in Study 3

	Dual-function intensionality	Dual-identity intensionality	Standard FB first order	Standard FB second order
Dual-function intensionality		.80** (.80**)	.78** (.79**)	.41† (.34)
Dual-identity intensionality			.82** (.80**)	.34† (.34)
Standard FB first order				.35† (.33)

†$p < .10$. **$p < .01$.

tions in Study 1 directly showed that in fact children did not invariably give any particular answer but were sensitive to information regarding the aspect under which the protagonist represented the object and to information concerning whether the protagonist knew about the identity of the different aspects. (b) The intensionality tasks were no more difficult than standard first-order FB tasks. (c) They were significantly easier, however, than second-order standard FB tasks. (d) There were no differences whatsoever between intensionality tasks using different contrast (sortal vs. property, and dual function vs. dual identity). In general, the different tasks in the different studies varied considerably in stimuli and structure yet produced highly converging results—suggesting that a robust cognitive phenomenon was tapped underlying performance in these superficially different tasks. (e) The different intensionality tasks (sortal, property, dual function, dual identity) and the standard first-order FB tasks were strongly correlated, even if age and verbal ability were controlled for.

Relation of the Present Findings to Previous Work

Taken together, the present findings are thus in contrast to most existing studies on the development of understanding intensionality that have found that intensionality tasks were much more difficult and thus solved later than standard FB tasks. Why is this so? Most plausibly, the intensionality tasks used in the present study are much less demanding in terms of linguistic and other performance factors than all previous studies. In contrast to the study by Russell (1987) with its, "Can we say. . .?" measure, for example, the tasks in the present study posed only minimal linguistic task demands. And in contrast to the action prediction measures used in other studies (e.g., Apperly & Robinson, 1998; Sprung et al., 2007), this study had a much simpler logical and conversational structure. In previous action prediction tasks, there were two objects, one of which was a pure A and the other one was both an A and a B, and the protagonist only knew the latter as a B. The protagonist was then looking for "an A" and the child had to predict where the protagonist would search. In contrast, in this study there was only one object that was both an A and a B, and the protagonist only knew about it as an A and had witnessed it as such being placed in one box. She then saw a B being transferred from that box to another one and was looking for "her A," and the child had to predict where she would look. Both tasks seem to measure the child's understanding of intensionality, but the task used in this study seems to be less demanding in at least two respects: First, as there is only one object involved, there are fewer memory demands. Second, there is no ambiguity (of "an A") and the resulting need for reference resolution (to which A does "an A" refer in this context?). Rather, the single object falls under two descriptions ("A" and "B") and the child has to decide whether a given description can refer to the object from the protagonist's perspective or not future research will need to directly test which of these factors—individually or combined—made the present tasks easier than previous intensionality tasks.

All in all, the results of the present studies suggest that a basic form of understanding intensionality develops considerably earlier than previously assumed. They thus favor Option 2 as discussed in the Introduction: the general claim that previous negative findings on intensionality tasks were due to extraneous performance factors. The present findings are highly compatible with the more specific Option 2a—that the crucial performance factors pertained to ambiguity, reference resolution, and the like figuring strongly in previous tasks and drastically reduced in ours. What about Option 2b then—the embedded perspective account by Perner and colleagues? Prima facie, the present findings seem compatible with the embedded perspective account in the following general respect: They show that not any form of understanding intensionality is as difficult as suggested by most previous research and that suitable modified intensionality tasks appear to measure the same competence as basic theory-of-mind (standard FB) tasks. However, the present findings seem much less compatible with the embedded perspective account in other respects: In particular, in contrast to the claims of the account, our findings show that sortal intensionality problems (a) are no more difficult than property intensionality problems and (b) are as difficult as and strongly related to standard FB problems.

Children's Conception of Propositional Attitudes Is a More Unified Phenomenon Than Previously Assumed

More generally, the present findings have fundamental theoretical implications for the description and interpretation of theory-of-mind development. Most importantly, they imply that children's developing understanding of propositional attitudes typically measured with standard FB and related tasks is a much more unified and sophisticated capacity than suggested by previous findings. The previous

empirical situation was somewhat paradoxical: On one hand, children from age 4 solved FB and related tasks—that quite plausibly require the application of propositional attitude concepts like "belief." On the other hand, however, they failed tasks measuring one of the fundamental logical properties of propositional attitudes without which, strictly speaking, one cannot really understand what a propositional attitude actually is—namely, their aspectuality. There was thus at the same time evidence for and against the claim that around 4 years children acquire the basics of our propositional attitude folk psychology. The present findings release that tension by revealing that under suitable circumstances (using simplified tests) tasks measuring belief ascription and tasks measure an understanding of the intensionality of beliefs do not only differ in difficulty, but in addition correlate strongly—suggesting that a common underlying capacity is in fact measured. And plausibly, this underlying capacity is the conceptual apparatus of grasping propositional attitudes and their aspectuality. The 4-year conceptual revolution is thus a real and comprehensive revolution.

Potential Implications for the Contrast Between Implicit and Explicit Theory of Mind

This unity in children's explicit capacity to ascribe beliefs and understand their intensionality raises very interesting theoretical questions concerning the relations between the capacities in implicit theory-of-mind tasks recently documented in infants (Onishi & Baillargeon, 2005); for review, see Baillargeon, Scott, & He, 2010) and those tapped in standard theory-of-mind tasks. The biggest theoretical challenge posed by the new infant findings is how to reconcile the fact that 1-year-olds in their looking behavior reveal sensitivity to others' FB with the fact that only 4-year-olds are capable of explicitly ascribing beliefs to others in standard tasks. Two extreme theoretical reactions to this challenge are, first, nativist modularity accounts claiming that implicit tasks measure the crucial cognitive competence, unmasked by performance factors that make explicit tasks unnecessarily difficult (Leslie, 2005) and, second, skeptical accounts claiming that implicit tasks do not require any ascription of propositional attitudes, but can be mastered in simpler ways and that true theory-of-mind competence is only revealed in explicit tasks (Perner & Ruffman, 2005; Sirois & Jackson, 2007). A third class of accounts presents a more differentiated picture: They claim both that implicit tasks may measure some precocious theory-of-mind competence and that theory-of-mind competence subsequently undergoes substantial qualitative development before the capacities measured in explicit tasks emerge (Apperly & Butterfill, 2009; Butterfill & Apperly, 2013; Low & Watts, 2013; Rakoczy, 2012). The most promising and well worked out account of this type is a two-systems theory according to which the capacities measured in early implicit tasks are perhaps precursors to but not the same thing as the later developing capacities measured in explicit tasks (Apperly & Butterfill, 2009). The former are subserved by a simpler, evolutionarily, and ontogenetically more ancient system (System 1) that operates fast and independently of central cognitive resources (such as language or executive function), but that has clear signature limits. Most importantly for present purposes, it can track belief-like mental states but not full-fledged beliefs and other propositional attitudes. With such a conception of belief-like states, a subject can still master some FB tasks, but only a limited variety—basically those having to do with situations in which a protagonist has registered an object at a certain place and is subsequently mistaken about the new location of the object (for details, see Butterfill & Apperly, 2013). So while theoretically a concept of belief-like states would allow an infant to pass some implicit FB tasks, this concept has its signature limits when it comes to situations where a protagonist has some other type of FB, in particular, beliefs about the identities of objects. System 2, in contrast, develops later, dependent on language and executive function, and it underlies the competence measured in standard explicit tasks. It is the conceptual apparatus of our full-fledge propositional attitude folk psychology that allows us to flexibly and explicitly ascribe all kinds of propositional attitudes about all kinds of contents and has thus no such signature limits as System 1.

This theory makes a straightforward prediction about the performance of infants and older children on implicit and explicit versions of different FB tasks: The competence in explicit tasks tapping an understanding of propositional attitudes and their features, subserved by System 2, should constitute a unitary phenomenon: mastery of standard FB tasks with different contents of the FB to be ascribed (about objects' locations, identities, etc.) should be strongly related (emerge together and correlate). Infants' competence in implicit tasks, in contrast, subserved by System 1, should be more fragmented: Infants should be able to pass some FB tasks (in which a protagonist is mistaken about an object's location) but not others (in which a protagonist is mistaken about an object's identity).

The present findings, taken together with some other recent findings, constitute evidence compatible with this prediction: It was recently found that infants master implicit FB tasks with FB about locations but not those with FB about objects' identities, whereas older children and adults in explicit equally tests master both types of tasks (Low & Watts, 2013). The present findings add to this by showing that the unity in the conceptual capacities measured in standard FB tasks goes even further and includes a basic grasp of intensionality.

One of the exciting questions for future research on the early development of theory of mind will be how children's understanding of the intensionality of propositional attitudes develops, perhaps in some precocious form even before the age tested here.

References

Anscombe, G. E. M. (1957). *Intention*. Oxford, UK: Basil Blackwell.

Apperly, I., & 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

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

Apperly, I., & Robinson, E. (2003). When can children handle referential opacity? Evidence for systematic variation in 5- and 6-year-old children's reasoning about beliefs and belief reports. *Journal of Experimental Child Psychology*, *85*, 297–311. doi:10.1016/S0022-0965(03)00099-7

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

Brentano, F. (1973). *Psychologie vom empirischen Standpunkt* [Psychology from an empirical standpoint]. Hamburg, Germany: Meiner. (Original work published 1874)

Butterfill, S. A., & Apperly, I. A. (2013). How to construct a minimal theory of mind. *Mind & Language*, *28*, 606–637. doi:10.1111/mila.12036

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

Fabricius, W. V., Boyer, T. W., Weimer, A. A., & Carroll, K. (2010). True or false: Do 5-year-olds understand belief? *Developmental Psychology*, *46*, 1402–1416. doi:10.1037/a0017648

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

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

Kamawar, D., & Olson, D. R. (2009). Children's understanding of referentially opaque contexts: The role of metarepresentational and metalinguistic ability. *Journal of Cognition and Development*, *10*, 285–305. doi:10.1080/15248370903389499

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

Kaufman, A., & Kaufman, N. (1999). *Kaufman Assessment Battery for Children* (P. Melcher & U. Preuß, Trans., 4th ed.). Frankfurt am Main, Germany: Swets Test Services.

Keysar, B., Lin, S., & Barr, D. J. (2003). Limits on theory of mind use in adults. *Cognition*, *89*, 25–41. doi:10.1016/S0010-0277(03)00064-7

Lalonde, C. E., & Chandler, M. J. (2002). Children's understanding of interpretation. *New Ideas in Psychology*, *20*, 163–198. doi:10.1016/s0732-118x(02)00007-7

Leslie, A. M. (2005). Developmental parallels in understanding minds and bodies. *Trends in Cognitive Sciences*, *9*, 459–462. doi:10.1016/j.tics.2005.08.002

Low, J., & Watts, J. (2013). Attributing false beliefs about object identity reveals a signature blind spot in humans' efficient mind-reading system. *Psychological Science*, *24*, 305–311. doi:10.1177/0956797612451469

McKay, T., & Nelson, M. (2014). Propositional attitude reports (Spring 2014 Edition). In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Retrieved from http://www.plato.stanford.edu/archives/spr2014/entries/prop-attitude-reports/

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. (1991). *Understanding the representational mind*. Cambridge, MA: MIT Press.

Perner, J., Brandl, J., & Garnham, A. (2003). What is a perspective problem? Developmental Issues in understanding belief and dual identity. *Facta Philosphica*, *5*, 355–378.

Perner, J., & Howes, D. (1992). "He thinks he knows": And more developmental evidence against the simulation (role taking) theory. *Mind & Language*, *7*, 72–86. doi:10.1111/j.1468-0017.1992.tb00197.x

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

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

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

Perner, J., Stummer, S., Sprung, M., & Doherty, M. (2002). Theory of mind finds its Piagetian perspective: Why

alternative naming comes with understanding belief. *Cognitive Development*, *17*, 1451–1472.
Rakoczy, H. (2012). Do infants have a theory of mind? *British Journal of Developmental Psychology*, *30*, 59–74. doi:10.1111/j.2044-835X.2011.02061.x
Russell, J. (1987). Can we say . . ? Children's understanding of intensionality. *Cognition*, *25*, 289–308. doi:10.1016/S0010-0277(87)80007-0
Searle, J. R. (1983). *Intentionality: An essay in the philosophy of mind*. Cambridge, UK: Cambridge University Press.
Sirois, S., & Jackson, I. (2007). Social cognition in infancy: A critical review of research on higher-order abilities. *European Journal of Developmental Psychology*, *4*, 46–64. doi:10.1080/17405620601047053
Sprung, M., Perner, J., & Mitchell, P. (2007). Opacity and discourse referents: Object identity and object properties. *Mind & Language*, *22*, 215–245. doi:10.1111/j.1468-0017.2007.00307.x
Wellman, H. M.. (2011). Developing a theory of mind. In U. Goswami (Ed.), *The Wiley-Blackwell handbook of childhood cognitive development* (2nd ed., pp. 258–284). Oxford, UK: Wiley-Blackwell.
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

Appendix A

The Procedure of the Intensionality Tasks

1. General familiarization with the two experimenters

The first experimenter (E1) who led the child through the sessions and asked the test questions and the second experimenter (E2) who operated a puppet (called "Susi") that was the protagonist in all tasks introduced themselves to the child and played with here until she felt comfortable.

2. Familiarization with the objects (t_1)

The puppet then left the room and E1 successively introduced the child to all the dual-aspect objects to be used in the target tasks. For example, in the sortal contrasts one object was a soft toy that could be turned inside out and was a bunny on one side and a carrot on the other side. The objects were the exact same objects that had been used in a previous study with infants on sortal object individuation (Cacchione, Schaub, & Rakoczy, 2013). This study has shown that infants (14-month-olds) understand the dual nature of these objects: When they saw a bunny enter into a box and then after a while saw an experimenter remove a carrot from the box, those infants who did not know the objects expected there to be another object (the bunny) inside the box, whereas infants who knew the objects did not (as indicated in their searching behavior in the box). In the property contrast, for example, one item was a sock that could be turned inside out and was blue on one side and red on the other. E1 showed the child the two aspects of each object, how to transform it (by turning it inside out) and encouraged her to engage in transforming the objects herself. In the course of this, she emphasized several times that Susi the puppet was not present and did not know about these objects. After the introduction of each object, E1 stressed that Susi had not been present and asked the child *Control Question 1* whether Susi knew that the A was also a B (e.g., "Does she know that the bunny is also at the same time a carrot?").

3. Placing of the object in Box 1 (t_2)

The puppet then entered the room and the first trial began. E1 showed an object under its A-aspect (e.g., bunny or red sock) to the child and the puppet and put into Box 1. The object was very small (approximately 4 cm in diameter) in relation to the box (approximately 40 × 40 × 40 cm) so that easily several objects of this kind would have fit into the box (this is important because the child has to think that it is possible for the protagonist to assume that there was more than one object in the box).

4. Transformation of the object (t_3)

E1 then retrieved the object from Box 1, transformed it (turned it inside out) to its B-aspect (e.g., carrot or blue sock) and put it back into Box 1. In the intensional and extensional conditions, the puppet was not present in this phase, and E1 stressed toward the child that they were going to play a trick on "Susi." After the transformation, E1 asked the child whether Susi knew that the B was the A (e.g., "Does she know that the bunny is now a carrot?"). In the true belief control condition the protagonist was present and witnessed the transformation.

5. Transfer of the object (t_4)

In the presence of Susi (who returned in the intensional and extensional conditions and remained present in the true belief condition),

E1 then retrieved the object under its B-aspect from Box 1 and transferred it to Box 2.

6. Control and test questions

E1 then turned to the child and whispered in a very low voice (so that Susi could not hear it) two control questions and the test question into the child's ear:
Control Question 2: Where we did we put the A in the beginning?

Control Question 3: Where is the A now?

Test Questions: Where will she look for her . . . [A]/[B]? (e.g., Where will Susi look for her bunny/red sock?)

Appendix B

Objects Used in the Sortal and Property Contrasts in Study 1

	Object	Aspect A	Aspect B
Sortal			
1	Winnie the Pooh/honeypot	Winnie the Pooh	Honeypot
2	Bunny/carrot	Bunny	Carrot
Property			
1	Ball	Smooth	Pointy
2	Sock	Red	Blue

Appendix C

Objects Used in the Sortal and Property Contrasts in Study 3

	Object	Aspect A	Aspect B
Dual function			
1	Pen/rattle	Pen	Rattle
2	Pencil/pocket lamp	Pencil	Pocket lamp[a]
Dual identity			
1	Peter/firefighter	Peter	Firefighter (firefighter uniform)[b]
2	Anna/doctor	Anna	Doctor (white collar, mouth protection, etc.)[b]

[a]The pen was also a pocket lamp (this was not visible but could only be found out by using it). In the crucial step, when the object was moved under its B-aspect from Box 1 to Box 2, E1 held the object in her hand such that the object itself was not visible, but it was visible that it was a pocket lamp due to the light beam emerging from the hand. [b]Crucially, by only looking at her, the person wearing her B-aspect costume was not recognizable as the A-aspect person.