Executive function plays a role in coordinating different perspectives, particularly when one's own perspective is involved

CrossMark

Ella Fizke *, Dana Barthel, Thomas Peters, Hannes Rakoczy

Institute of Psychology & Courant Research Centre "Evolution of Social Behaviour", University of Göttingen, Waldweg 26, D-37073 Göttingen, Germany

ARTICLE INFO

ABSTRACT

While developmental experiments with children and elderly subjects, work with neuro-psychological patients and adult experimental studies have consistently found close relations between executive function and theory of mind, the foundation of this relation still remains somewhat unclear. One prominent account holds that executive function is specifically involved in ascribing such mental states, paradigmatically beliefs, that aim at representing the world truly because ascribing such states requires inhibition of normative defaults (beliefs being true) (e.g. Sabbagh, Moses, & Shiverick, 2006). The present studies systematically tested for the role of executive function in different forms of mental state ascription as a function of the type of state ascribed (beliefs or desires) and the first person involvement of the ascriber (whether she herself has an attitude conflicting with one to be ascribed to someone else) in young children. The results reveal that (i) executive function is related not only to belief ascription but equally to desire ascription when both are matched in terms of logical complexity (such that two subjective attitudes have to be ascribed to two agents that are incompatible with each other). (ii) Both for desires and for beliefs, these relations are strongest in such tasks where the ascriber herself is one of the two agents, i.e. has a belief or desire herself that stands in contrast to that to be ascribed to someone else. All in all, these findings suggest that executive function figures in coordinating perspectives more generally, not only epistemic ones, and in particular in coordinating others' and one's own conflicting perspectives.

1. Introduction

Theory of mind research investigates how children come to explain and predict rational action by ascribing mental states, such as beliefs, desires and intentions (e.g. Perner, 1991). This perspective taking ability – in its explicit forms – emerges during the preschool years and depends crucially on the development of domain-general cognitive capacities. In addition to linguistic abilities, one such factor that has been found to be crucial for theory of mind (ToM) across the lifespan is executive function (EF). EF and ToM have been shown to be closely related in child development (e.g. Carlson & Moses, 2001; Perner & Lang, 1999), in adults (e.g. Qureshi, Apperly, & Samson, 2010), in aging (e.g. Bailey & Henry, 2008; Rakoczy, Harder-Kasten & Sturm, 2012) and in neuropsychological patients (e.g. Samson, Apperly, Kathirgamanathan, & Humphreys, 2005).

Most extensively this ToM-EF relation has been studied in early child development, where it has been documented to be very robust, both in cross-sectional and longitudinal designs (where EF at time 1 predicts ToM at time 2), holding across difference cultures, and even when extraneous factors such as chronological and mental age are controlled for (e.g. Carlson & Moses, 2001; Carlson, Moses, & Hix,

* Corresponding author. Tel.: +49 551 39 9289; fax: +49 551 39 9322.
  E-mail address: ella.fizke@psych.uni-goettingen.de (E. Fizke).

1998; Rakoczy, 2010; Sabbagh, Xu, Carlson, Moses, & Lee, 2006). Importantly, the associations between ToM and EF appear to be rather specific: EF correlates with a variety of superficially different ToM tasks, but not with structurally similar tasks that do not require ToM reasoning (e.g. Sabbagh, Moses, & Shiverick, 2006).

But what accounts for this relation? What role might EF play in ToM reasoning? One prominent account views EF as fundamental to what has been considered the litmus test of ToM reasoning, namely the ascription of false beliefs (FB): "Executive functioning is required to reason only about representations that are intended to reflect a true state of affairs" (Sabbagh, Moses, et al., 2006, p. 1034). The idea behind this proposal is the following: beliefs, in contrast to other types of propositional attitudes such as desires, wishes or intentions, have so-called "mind-to-world direction of fit" (Searle, 1983). That is, they aim at truth, at representing the world as it is. They thus have a normative default (beliefs ought to be true), and ascribing false beliefs requires deviation from and inhibition of this default (Russell, 1996; Sabbagh, Moses, et al., 2006). This proposal gets empirical support from several studies showing that EF is related to false belief tasks, but not to closely matched tasks without the involvement of truth-aiming attitudes. First, so-called "conflict inhibition" EF tasks (incorporating both inhibition and working memory demands) correlate with false belief tasks and with tasks requiring an understanding of false signs (that also aim at truth), but not with structurally analogous "false photo" tasks (Sabbagh, Moses, et al., 2006). The crucial difference in terms of logical structure of the tasks is that the latter do not involve truth-aiming false representations (outdated photos aren't wrong. . . – see below and Appendix D).

Second, EF has been found to correlate with belief ascription specifically, as compared to other forms of attitude ascription: FB tasks, but not closely matched tasks involving ascribing simple desires or pretend attitudes, have been found to correlate with EF tasks (Moses, Carlson, Stieglitz, & Claxton, 2003; cited in Moses et al., 2005).

On a theoretical level, this account has been disputed by a prominent theory that views the role EF plays in ToM reasoning as much more general and pervasive. According to Leslie and colleagues, EF is involved in combining, coordinating and inhibiting meta-representations in ToM inferences very generally, both regarding beliefs and regarding desires (e.g. Friedman & Leslie, 2004; Leslie, German, & Polizzi, 2005; Leslie & Polizzi, 1998).[1]

On an empirical level, the studies showing specific correlations of EF only with false belief and false sign tasks involved some fundamental confounds. In contrast to the tasks that did not correlate with EF, the FB tasks (a) involved beliefs (or other representations aiming at truth), (b) presented so-called "perspective problems", and (c) did so in a way that there was a conflict between the ascriber's own self-perspective and the perspective to be ascribed to someone else.

Regarding (b), one more general theory about ToM development claims that what emerges around age 4 is a more general capacity to coordinate multiple subjective perspectives and to solve "perspective problems" (Perner, Brandl, & Garnham, 2003; Perner & Roessler, 2012; Perner, Stummer, Sprung, & Doherty, 2002; Perner, Zauner, & Sprung, 2005; see also Moll, Meltzoff, Merzsch, & Tomasello, 2013): These are formally defined as tasks requiring the coordination of the content of two representations (e.g. mental attitudes, pictures, sentences, etc.) whose content (e.g. p, q) cannot be joined by a simple conjunction (p and q) without being relativized to different perspectives. A good example of perspectives embodied in external (non-mental) representations are the visuo-spatial perspectives in pictorial depictions. Think of two objects A and B depicted from opposite sides of a room. The two contents (of the pictures) "A is in front of B" and "B is in front of A" cannot be combined by a simple conjunction ("A is in front of B and B is in front of A" is inconsistent). Only relativizing them to standpoints allows a conjunction: "A is in front of B as depicted in picture 1, but B is in front of A as depicted in picture 2". In similar fashion, the false sign test (in which a sign post, supposed to indicate the location of an object O which is in B, in fact points to A) presents a perspective problem: one has to understand the perspective clash between what the sign "says" (O is in A) and what is the case (O is in B) (Perner & Leekam, 2008; Sabbagh, Moses, et al., 2006; Sabbagh, Xu, et al., 2006). The superficially similar false photo test (Zaitchik, 1990), in contrast, does not present a perspective problem. In this task, a photo is taken of an object O in A at time 1. At time 2, O is moved to B, and the crucial question is what the photo will show about O's location. This does not present a perspective problem because the photo does *not* refer to the current real situation and misrepresents it in the way the false sign does. Rather it represents a past situation. Whereas the false sign is literally false, the photo is merely outdated[2] (Perner & Leekam, 2008; Sabbagh, Moses, et al., 2006; Sabbagh, Xu, et al., 2006) (see Appendix D for details).

Regarding mental representations, understanding conflicting visual perspectives ("A is front of B" as seen from standpoint 1 and "B is in front of A" as seen from standpoint 2) and ascribing false beliefs ("object O is in B" in reality, but the protagonist believes that "object O is in A") present paradigm cases of perspective problems. In contrast, ascribing simple desires (e.g. Repacholi & Gopnik, 1997; Wellman & Bartsch, 1988), pretence attitudes, etc. does not (Perner et al., 2005, see Appendix D for details). Crucially, however, there are perspective problems that

---

[1] This account receives empirical support from a series of studies showing that reasoning about a false belief combined with a negative desire (the aim is to avoid an object) is much harder than reasoning about a true belief combined with a negative desire or about a false belief combined with a positive one (the aim is to find an object) (Leslie & Polizzi, 1998). These results are interpreted as revealing multiple summative inhibitory demands in terms of target shifting in the process of selection, both when inhibiting the representations of the true states of affairs in ascribing beliefs deviating from truth, and in suppressing representations of locations where objects are when ascribing desires to avoid such objects (Leslie et al., 2005).

[2] While we find this general theoretical approach and its specific analysis of the false sign and the false photo tasks very convincing and useful due to its formal precision and thus use it as a conceptual basis for the current studies, it should be noted that it has not met with universal acceptance. For critique see Cohen and German (2013).

do not involve truth-aiming attitudes, such as incompatible desires held by different persons (Perner et al., 2005; Rakoczy, Warneken, & Tomasello, 2007) or even within one person (Choe, Keil, & Bloom, 2005). Such desires (say to the effect that p vs. non-p) are structured in a "mutually exclusive" (Moll et al., 2013) way, in the sense that only one can be fulfilled at the same time and that understanding them requires a relativization to standpoints ("p and non-p" is inconsistent; what is required is "person 1 desires p whereas person 2 desires non-p"). Work with such tasks suggests that EF might in fact be involved in such perspective problem tasks that do not involve belief ascription in the same way as in FB tasks (Rakoczy, 2010).

Regarding (c), one limitation of previous work is that all the studies suggesting a specific FB-EF relation have used standard FB tasks (Perner, Leekam, & Wimmer, 1987; Wimmer & Perner, 1983). In such tasks, the child has to ascribe a belief to a protagonist about a state of affairs regarding which she herself already has a contrasting belief – creating a situation with a strong "pull of the real" or "curse of knowledge" from the child's perspective (e.g. Birch & Bloom, 2007). It is well possible that EF plays a crucial role particularly in such situations where the ascriber's own perspective needs to be disregarded in judging another's perspective (see, e.g. Apperly, Samson, & Humphreys, 2005). Two lines of evidence – on belief and desire reasoning – might be taken as suggestive of such a possibility:

First, support comes from studies with so called reality unknown false-belief tasks (Call & Tomasello, 1999). In these modified FB tasks (details see below) the subject herself is ignorant about the real location of the target at the moment when the character's false belief has to be ascribed. The subject can only infer the real location when taking into account the false belief of a protagonist who (in contrast to the subject) has seen the hiding in one of two identical containers at time 1, has not witnessed the subsequent swapping of the containers at time 2, and indicates his (false) belief at time 3. That is, the subject initially has no belief of her own as to the location of the object but only reaches such a belief by taking into account and reasoning from the belief of the other person. From a theoretical point of view, two characteristics of these tasks are fundamental: In contrast to previous modifications of belief tasks (e.g. Wellman & Bartsch, 1988), (i) the subject initially does not have an own perspective and (ii) there is a conflict between reality and the protagonist's subjective perspective (i.e. the task represents a true perspective problem). Empirically, this type of FB task has been found to be as difficult (or even more difficult) as standard FB tasks for children (Call & Tomasello, 1999; Figueras-Costa & Harris, 2001), quite likely because it poses other kinds of task demands in terms of working memory and inferential complexity (reasoning backwards from the other's belief to reality; see also Robinson & Mitchell, 1995; Wertz & German, 2007). What matters for current theoretical purposes, however, is not so much the question of absolute levels of performance on standard vs. reality unknown FB tasks (which might be comparable but due to different reasons), but rather whether they involve different cognitive processes, in particular whether they relate differentially to EF.

Evidence for such differential involvement of EF in standard vs. reality unknown FB tasks comes from a neuropsychological case study by Samson, Apperly, Kathirgamanathan & Humphreys (2005) which suggests that ascribing beliefs to another person and inhibiting one's own perspective are functionally dissociable in adults to some degree. A patient with severe EF deficits due to a right fronto-temporal lesion was unable to ascribe false beliefs (on standard FB tasks) to someone else only when he himself knew about the current state of affairs in question (where the target object actually was located) at that moment. Crucially, however, he succeeded on the modified reality unknown FB task. Using the same two types of FB tasks, Bailey and Henry (2008) showed that decline in EF in older age is related to declining performance in standard FB tasks, but not in the modified tasks without a self-perspective that needs to be inhibited.

A second line of evidence comes from a developmental study investigating the role of EF in ascribing incompatible desires to two persons (Rakoczy, 2010): In this study, EF was generally correlated with such desire ascription, but this correlation tended to be higher for ascribing desires when one of the persons was the ascriber herself than for ascribing such desires to two third persons.

In contrast to these data with elderly and neuropsychological populations, to date there are no developmental data on the role of self-perspective inhibition in the relation of EF to belief ascription in children. All in all, existing work thus remains inconclusive as to the exact developmental role EF plays in ToM reasoning – whether it relates to ascribing truth-aiming attitudes, perspective problems more generally, or specifically to such perspective problems with a self-perspective to be inhibited. The rationale of the present work therefore was to investigate the relation of EF to different forms of ToM reasoning by systematically varying both - the type of attitude to be ascribed (belief vs. desire) and the involvement of a self-perspective to be inhibited. To this aim, we used tasks of ascribing beliefs vs. incompatible desires in versions in which the ascriber did or did not have a belief/desire that contrasted with the one to be ascribed to a protagonist.

The logic is the following: If EF is indeed specifically involved in ascribing truth-aiming representations, there should be specific correlations of EF with FB tasks, but not with desire ascription tasks. If on the other hand, EF were involved more generally in solving perspective problems, there should be correlations of EF with both – ascribing false beliefs and incompatible desires. And finally, if EF was specifically involved in inhibiting self-perspectives in solving such perspective problems, these correlations should be most pronounced in cases where the ascriber herself has a belief or desire conflicting with that of the protagonist.

Study 1 tested this with a full 2 (kind of attitude to be ascribed: belief vs. desire) × 2 (self-perspective inhibition demands: high vs. low) design of ToM tasks whose relation to EF was tested. As this study yielded clear evidence, converging with previous findings, that EF was related to ascribing incompatible desires, particularly when there was a strong self-perspective to be inhibited, but ambiguous evidence regarding the ascription of beliefs, Study 2

Table 1
Factors and corresponding ToM-tasks under systematic investigation in study 1. The design of study 2 followed study 1, focusing only on belief ascription and the role of self-perspective involvement.

Self perspective inhibition	Attitude		
	Desire	Belief	
		Study 1	Study 2
High	Conflicting desires task; first-person version	Reality known Search FB task Reality known Unexpected content FB task	Reality known Search FB task Reality known Location change FB task
Low	Conflicting desires task; third-person version	Reality unknown Search FB task	Reality unknown Search FB task Reality unknown Location change FB task

investigated belief ascription with and without self-perspective more systematically in a variety of different FB task formats – including also a standard FB test format with high and low requirements of self-perspective inhibition (see Table 1).

On the basis of the theoretical background and the available empirical evidence we expected (1) positive correlations among the conflicting-desires and among the false-belief sub-tasks and (2) positive raw and partial correlations between the different ToM tasks and executive functioning tasks (3) in particular for those ToM tasks with high demands of self-perspective inhibition.

2. Study 1

2.1. Method

2.1.1. Participants

45 Native German speaking children were included in the final sample (20 females). Children's mean age was 47;4 months (SD 4.67; range 40–55). Children in both studies were recruited through day-care centers in the city of Göttingen. Socio-economic background was mostly middle to high.

2.1.2. Design

The basic design of both studies is depicted in Table 1. For study 1 each child was tested in an interactive play setting with two experimenters in two sessions (average interval = 4 days). Children received three types of tasks: In one session (30–45 min), children were administered a conflicting-desires (CD) task in a first- and a third-person version. In the other session (25–35 min) children received three false-belief (FB) tasks. (Three additional tasks were administered, which focused on another research question and are therefore not reported here.) FB tasks comprised a standard unexpected content (UC) task ('Smarties') and the search FB task. This consisted of two versions, one in which the child knew about the real location of the target (reality known condition; RK) and another in which the child herself was ignorant about the location of the target (reality unknown condition; RU) but could infer it based on the false belief of a protagonist. In both conflicting-desires versions and in the three FB tasks the child consecutively

completed two trials and thus received a total score of 0–2 for each of the five tasks. The order of sessions was counterbalanced, as was the order of 1st and 3rd person version of the CD task, the order of search- and UC task and the order of RK and RU condition within the search task. An executive function (EF) task (the bear-dragon task) was always administered in the first session after the first FB task/the CD task. To control for verbal ability and working memory, both studies included a vocabulary and a working memory test.

2.1.3. Materials and procedure

For both studies sessions were videotaped and a single observer transcribed and coded all tasks life/on-line, making use of the videotape if necessary.

2.1.3.1. Conflicting-desires task. This task contained a first- and a third-person version (see Appendix A) as developed in the study by Rakoczy (2010; an extended version of a task by Moore et al., 1995) and thus enabled us to distinguish between high and low first-person-involvement in desire understanding. In the two 1st person trials, the child and another character (puppet 'Rudi' enacted by E) played a game in which they had mutually incompatible desires, only one of which could be satisfied. In the two 3rd person trials two third characters (toy figures enacted by E), who expressed their mutually incompatible desires, played against each other. Across children it was counterbalanced, who won the first trial. The setup of the game was as follows: the child/toy character 'Peter' and another character (puppet 'Rudi'/toy character Susi) shared a booklet. For each trial E pinned two stickers (an interesting and a boring one in the 1st person version/two interesting ones in the 3rd person version) on a 'chance machine' that determined, which of them would be put into the book (see Fig. 1). Each version included first a series of warm-up trials (see Appendix A for all trials). For the test trials, the child/Peter was asked which sticker should go in the book (children virtually always chose the interesting one). Then Rudi/Susi expressed the opposite desire. This was repeated a second time (if a child on the second turn changed her desire and agreed with Rudi, the task was terminated and excluded from analysis). For each trial two questions about each character's desire (Q1) and two questions about
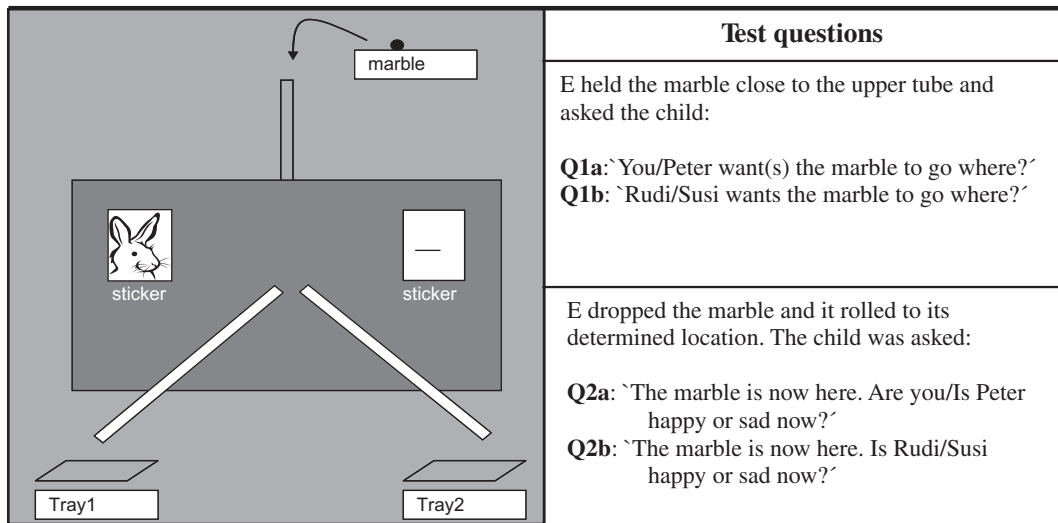
Fig. 1. 'Chance machine' in the conflicting-desires task. E dropped a marble into an inverted Y-shaped tube that disappeared behind a board and re-appeared in one of two tube ends from where it dropped in a seemingly random fashion (in fact, E could control the outcome). The sticker above the box in which the marble went was put in the booklet.
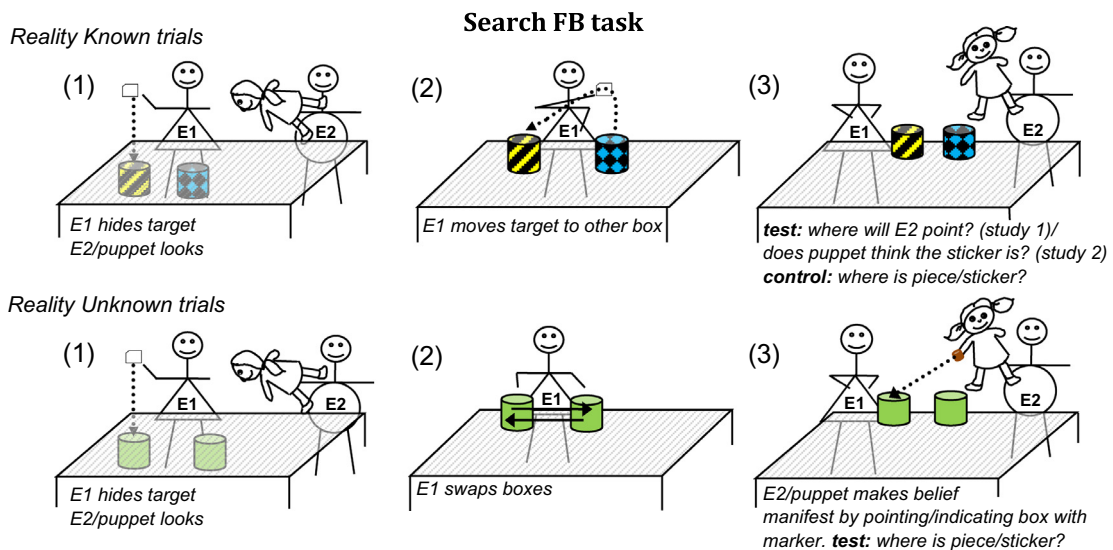

Fig. 2. Schematic event sequences of the search FB task used in studies 1 and 2. *Major changes in study 2.* E2 animated a puppet as communicator instead of playing this role herself. Children were given additional verbal hints in step (1) and (2) (see Section 2.2). In step 3 of the reality unknown trials the belief is made manifest by the puppet placing a marker on the box instead of pointing of E2 in study 1.

the character's desire-dependent emotions (Q2) followed (see Fig. 1). For each trial, children got a score of '1' for Q1 (and Q2 respectively) only if they answered both sub-questions (concerning Rudi and themselves/Peter and Susi) correctly (otherwise a '0') and thus received a sum score of 0–2.

2.1.3.2. False-belief tasks. (1) Unexpected content false-belief task. A standard version of this FB measure with a choco-late box containing a pen was used (Perner et al., 1987). After children had been shown the content, they received

the control question ('now what is in here?') and were given negative feedback and a second chance if they an-swered incorrectly. Two test trials followed, including a first-person ('initially, before you looked into the box, what did you think was in the box?') and a third-person ('what will you friend (name of friend) think is in the box?') test question.

(2) Search false-belief task (see Fig. 2). A modified version of the Call and Tomasello's (1999) procedure was used, that measures children's understanding of false belief through a search task. The version used in the first study

(see Appendix B), consisted of three to eight familiarization trials (a minimum of three consecutively successful trials had to be reached), four blocks of one to two control trials, and then two reality known (RK) and two reality unknown (RU) trials in counterbalanced order. Children had to pass at least one of two trials in each control block. In all trials, the child observed E1 hiding a desired target (a puzzle piece in study 1) in one of two boxes (but not in which one). During the trials, a communicator was present, who could observe the hiding procedure. In the test and control trials, the communicator left the room and was not present while E1 changed the location of the target. Only in the RK trials the child then could see the target being moved by E1 from one box to the other. In the RU trials, however, E1 swapped two identical boxes. As the target was not taken out of the box, the child did not know in which box it was located. By her return E2 pointed to the box where she had seen the target being hidden. The child could deduce the real location if she was aware of the false belief of the communicator.

2.1.3.3. Bear-dragon (see Appendix C). This task developed by Reed, Pien, and Rothbart (1984) and adapted by Kochanska, Murray, Jacques, Koenig, and Vandegeest (1996), measures children's conflict inhibition. In the task, E1 animated a nice bear and a mean dragon that both gave commands. Children were instructed to perform the movements of the bear but not those commanded by the dragon. They received up to six dragon practice trials followed by five bear and five dragon test trials in alternating order. For coding, only dragon trials were taken into account; children were given a score of '3' if they did not carry out the action; a score of '2' if they carried out a different action instead; a score of '1' if they partially carried out the target act; and a score of '0' if they fully carried it out. The sum score thus ranged from '0' to '15'. A second, independent observer coded a 20% sample of ten randomly selected tapes for reliability; weighted κ was .82 for dragon trials.

2.1.3.4. Working memory: Counting and labeling. In this task, developed by Gordon and Olson (1998) E1 demonstrated with a set of three toys the steps (1) labeling, (2) counting and (3) counting and labeling of the toys. Children received two test trials each with their own set of toys, in which they were asked to repeat the steps. In each trial they received a score of '2' if they mastered step (3) on the first attempt and a score of '1' if they mastered it, after E had modeled step (3) one more time.

2.2. Results

Data from 45 children were used in the final analysis. In the unexpected content task, two children did not answer to a test question and two failed to answer the control question. In the search task, one child failed to answer the reality known control question, one child failed in the pretest, 32 children failed in both trials of trust, invisible displacement or ignore communicator control blocks, resulting in only 16 valid scores for this task. In the 1st person version of the conflicting-desires task three children

adapted their desire to the desire of Rudi and one child did not answer to all Q2 questions, resulting in 41 valid values in this version. In the 3rd person version one child did not answer to one Q1 question. In the counting and labeling task, four children rejected the task and owing to experimental error one child did not provide a value in the vocabulary test.

In the following analyses, one-tailed tests were conducted when testing directed a priori hypotheses derived on theoretical and empirical grounds as described above: First, we expected positive correlations among the conflicting-desires and among the different FB sub-tasks. Second, we expected positive raw and partial correlations between the different ToM tasks and the executive functioning tasks in particular for those ToM tasks with high demands of self-perspective inhibition.

2.2.1. Performance on false-belief and conflicting-desires tasks

The proportions of successfully mastered false-belief and conflicting-desires tasks are depicted in Fig. 3. Children's performance across the different CD sub-tasks was uniform. The exception was Q2 in the 1st person version when the other player's desire was fulfilled ("child looses") which was answered significantly less proficiently than Q2 after the child's desire was fulfilled ("child wins") (McNemar's test, $\chi^2(1, N = 41) = 12.07$, $p < .0001$). In this case, children often answered incorrectly "The other payer is happy, and I'm happy as well". This very same answer pattern that has been found previously in this task (Rakoczy, 2010) might reflect false negatives either due to demand characteristics (children do not want to admit they are disappointed) or because children are really happy because the other player is happy. It is thus questionable whether Q2 in the 1st person plural version constitutes a valid indicator of the ascription of desires and desire-dependent emotions. Given this doubt, following previous work (Rakoczy, 2010), for statistical analyses regarding desire ascription in the 1st person plural, an adapted sum score (0–3) ranging over the two Q1 data points and Q2 only after the child has won was computed and used for control analyses.

In the 3rd person version of the conflicting-desires task, Q1 and Q2 sum scores were correlated ($r(44) = .58$, $p < .001$, *one tailed*). For the 1st person version there was a trend in the same direction ($r(41) = .26$, $p = .053$, *one tailed*). However, given the doubts about the validity of Q2 when the child lost (see above), more important was that the Q1 score was correlated with Q2 when the child wins (Spearman's $r(42) = .30$, $p < .03$, *one tailed*). To compare children's performance with other tasks, proportion scores (0–1) were computed across Q1 and Q2 scores for the 1st ($M = .85$, $SD = .21$) and for the 3rd person version ($M = .92$, $SD = .21$) respectively, and an adapted proportion score for the 1st person version out of the two Q1 sub-tasks and Q2 when the child won ($M = .94$, $SD = .16$). The 1st and 3rd person score differed significantly (Wilcoxon test, $Z(41) = −2.22$, $p < .03$), but importantly this difference was no longer found when comparing the 3rd person to the adapted (and arguably more valid) 1st person score (Wilcoxon test, $Z(42) = .36$, $p > .71$).
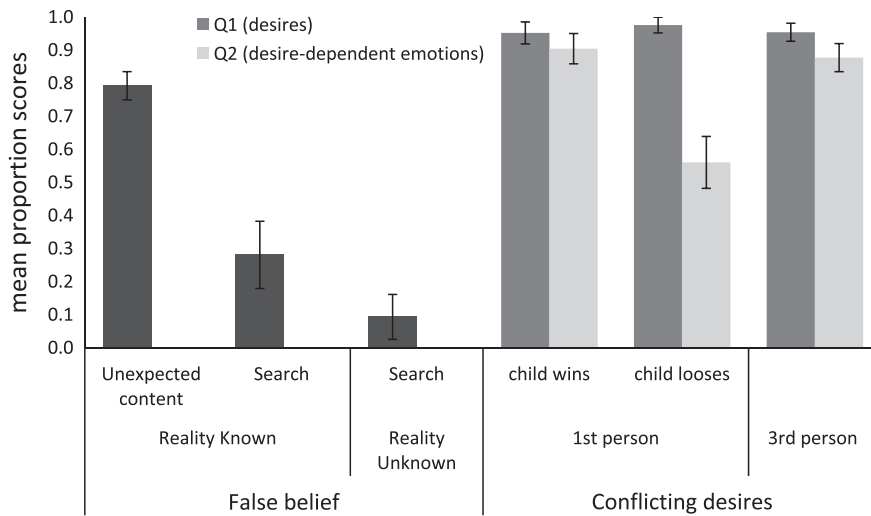
Fig. 3. Mean sum scores of the false-belief and conflicting-desires tasks (error bars represent SEM).

A repeated measurement ANOVA revealed significant differences between the false-belief tasks ($F(2,14) = 23.4$, $p < .0001$). Specific comparisons showed that the unexpected content task (mean proportion score, $M = 0.79$) was easier for children than both the reality known ($M = .28$) and the reality unknown search task ($M = .09$) (ts > 4.39, ps < .001). Performance across search tasks did not differ significantly. Unexpectedly, the unexpected content and the reality known task did not correlate significantly ($r(16) = .24$, $p = .19$, *one tailed*) nor did the two search tasks ($r(16) = .20$, $p = .23$, *one tailed*). The comparison of conflicting-desires with false-belief tasks revealed significant differences across tasks ($F(4,10) = 32.34$, $p < .0001$). Specific comparisons showed that children performed significantly better in the CD 1st person version than in both search tasks (ts > 5.55, ps < .0001). The adapted 1st person CD score and the 3rd person CD score were higher than the unexpected content score and both search scores (ts > 4.01, ps < .005).

2.2.2. Conflict inhibition and working memory

The two dependent measures in the bear-dragon (BD) task, the number of practice dragon trials children needed until they succeeded ($M = 1.40$, $SD = .75$) and the sum score

Table 2
Raw/partial correlations (controlling for age and vocabulary) (with valid N) between false-belief (FB) and conflicting-desires tasks and the bear-dragon and working memory measures.

	Vocabulary	WM	BD composite score	Unexpected content FB	Search RK	Search RU	Conflicting-desires 1st person	Conflicting-desires adapted 1st person score	Conflicting-desires 3rd person
Age	.35** (44)	44** (41)	.31** (45)	.54** (41)	.35 (16)	.14 (16)	.19 (41)	.12 (42)	.05 (44)
Vocabulary		.38** (41)	.43** (44)	.41** (40)	.29 (16)	.14 (16)	.25+ (40)	.28* (41)	.16 (43)
WM			.47** (41)/ 32* (37)	.18 (38)	.28 (15)/ .11 (11)	.28 (15)/ .23 (11)	.38** (37)/ .30* (33)	.28* (38)/ .20 (34)	.19 (40)/ .16 (36)
BD composite score				.57** (41)/ .45** (36)	−.11 (16)/ .35 (12)	−.18 (16)/ .12 (12)	.60** (41)/ .55** (36)	.69** (42)/ .67** (36)	.37** (44)/ .35* (39)
Unexpected content FB					.24 (16)	.36 (16)	.31* (37)/ .20 (33)	.46** (38)/ .43** (34)	.41** (40)/ .43** (36)
Search RK						.20 (16)	.00 (14)/ .12 (10)	−.06 (15)/ −.17 (11)	−.05 (15)/ .01 (11)
Search RU							.15 (14)/ .11 (10)	.10 (15)/ .06 (11)	.10 (15)/ .01 (11)
Conflicting-desires 1st person								.83** (41)/ .82** (36)	.30* (41)/ .27* (36)
Conflicting-desires adapted 1st person score									.45** (42)/ .43** (36)

Note. All *p* values are one tailed: **$p < .01$; *$p < .05$; +$p < .10$. BD, bear-dragon; WM, working memory.

Table 3
Multiple regressions for predicting conflicting-desires scores.

	Conflicting desires 1st person sum score				Conflicting desires 1st person adapted sum score				Conflicting desires 3rd person sum score			
	B	SEB	Beta	T	B	SEB	Beta	T	B	SEB	Beta	T
Age	−.01	.03	−.08	−.49	−.01	.68	−.14	−.10	−.02	.03	−.11	−.63
Vocabulary	−.02	.04	−.07	−.44	.00	.02	.01	.09	.02	.04	.06	.35
BD composite score	.38	.11	.56	3.46**	.30	.06	.72	4.95**	.37	.13	.50	2.92**
WM	.09	.08	.19	1.14	.00	.04	.01	.06	−.01	.09	−.01	−.06
	$R = .62$, $R^2 = .38$, adjusted $R^2 = .30$				$R = .69$, $R^2 = .48$, adjusted $R^2 = .42$				$R = .49$, $R^2 = .24$, adjusted $R^2 = .16$			

Note. All p values are one tailed: **$p < .01$; *$p < .05$; +$p < .10$. BD, bear-dragon; WM, working memory.

Table 4
Multiple regressions for predicting false-belief scores.

	Unexpected content FB task				Reality known FB score				Reality unknown FB score			
	B	SEB	Beta	T	B	SEB	Beta	T	B	SEB	Beta	T
Age	.05	.02	.45	3.37**	.04	.05	.29	.82	.02	.04	.15	.37
Vocabulary	.03	.02	.16	1.23	.04	.09	.16	.40	−.02	.08	−.10	−.21
BD composite score	.24	.06	.52	3.76**	−.19	.20	−.33	−.10	.02	.17	.04	.10
WM	−.10	.05	−.31	−2.17*	.15	.14	.35	1,04	.09	.12	.28	.74
	$R = .74$, $R^2 = .54$, adjusted $R^2 = .49$				$R = .64$, $R^2 = .29$, adjusted $R^2 = .007$				$R = .31$, $R^2 = .09$, adjusted $R^2 = −.27$			

Note. All p values are one tailed: **$p < .01$; *$p < .05$; +$p < .10$. BD, bear-dragon; WM, working memory.

for the dragon trials ($M = 12.96$, $SD = 4.33$), were significantly negatively correlated ($r(45) = −.67$, $p < .001$). Following previous studies (e.g. Sabbagh, Moses, et al., 2006) these two scores were standardized (with reverse scoring for practice trials) and summed to form a composite BD score. The mean score in working memory (counting and labeling task) was 2.32 ($SD = 1.81$). WM was correlated with the composite BD score ($r(41) = .47$, $p < .001$; see Table 2).

2.2.3. Relations between conflict inhibition, working memory, false-belief and conflicting-desires tasks

The raw and partial (age- and vocabulary-controlled) correlations between the BD composite score, WM and the FB and CD scores are presented in Table 2.

Both conflicting-desires versions were correlated with the unexpected content task, this correlation remained significant for the 3rd person and for the adapted 1st person score, when age and vocabulary were controlled for. Regarding conflict inhibition and it's relation to belief and desire understanding, the BD composite score was correlated (in raw and partial correlations controlling for age and vocabulary) with both conflicting-desires versions, with this pattern being clearest for the 1st person version. This was also confirmed for the adapted 1st person proportion score. The BD composite score was also correlated with performance in the unexpected content task. WM correlated (in raw and partial correlations) with the 1st person CD score. Concerning the search tasks, as expected, the reality unknown FB score did not correlate with the BD composite score but unexpectedly, also the reality known search task showed no correlation with the BD score. Note

that part of the reason for that might be the high rate of children, which did not pass all control trials and the very low performance in both search tasks, which was close to floor. Taking into account the missing correlation between the two RK tasks (the search RK task and the unexpected content task) the validity of the search task seems questionable.

Results concerning conflict inhibition were confirmed by separate multiple regression analysis with conflicting-desires and false-belief scores as dependent variables (Tables 3 and 4). Age, vocabulary, the BD composite score and WM were all entered simultaneously. Only the BD composite score predicted the CD scores in both versions. Both, the BD score and WM predicted the unexpected content FB score. For the search FB scores, none of the predictors turned out to be significant (which might again result from the high dropout rate and floor effect).[3]

2.3. Discussion

First, this study confirmed the well-established finding of a significant role of executive function in ToM reasoning. Concerning the nature of this relation, the study suggests – in line with other recent evidence – that EF plays a significant role not only in belief, but also in desire-reasoning if the task presents a true perspective problem as in the case of incompatible desires. Moreover, the study supports the

---

[3] Note, however, that given the multiple testing and the use of one tailed tests (though based on theoretically motivated a priori hypotheses), future replications of the present results are necessary to consolidate the correlational and regression patterns found here.

Table 5
Raw/partial correlations (controlling for age and vocabulary) (with valid N) between proportion false-belief scores and the executive function and working memory measures.

	Vocabulary	WM	Day-night	Grass-snow	BD composite score	Reality known score	Reality unknown score
Age	.20[+] (60)	.42** (61)	.22[+] (39)	.32** (58)	.33** (52)	.19[+] (62)	.32** (62)
Vocabulary		.34** (59)	.16 (38)	.15 (57)	.34** (52)	.38** (60)	.17[+] (60)
WM			.14 (39)	.35** (58)	.39** (51)	.26* (61)/.06 (47)	.36** (61)/.22[+] (47)
Day-night				.24[+] (39)	.03 (33)	.14 (39)	.05 (39)
Grass-snow					.11 (49)	−.00 (58)	.05 (58)
BD composite score						.40** (52)/.28 * (48)	.21 (52)/.10 (48)
RK score							.37** (62)

Note. All p values are one tailed: **$p < .01$; *$p < .05$; +$p < .10$. BD, bear-dragon; WM, working memory.

assumption that self-perspective inhibition constitutes an important factor in the nature of the ToM-EF relation, as EF was related to desire reasoning in particular when there was a strong self-perspective to be inhibited. However, evidence regarding the ascription of beliefs remained somewhat ambiguous: as expected, the standard (unexpected content) FB task revealed a strong correlation with EF, but a novel reality unknown FB measure, in which the child herself is ignorant about the real location of the target did not correlate with EF. However, this missing correlation of EF with the RU task is difficult to interpret:

(1) Missing correlations between the standard unexpected content task and the search task (both the reality known and reality unknown versions) suggest that there might have been a problem with the validity of the search task; (2) the high dropout rates and floor effects (i.e. due to many trials, long sessions, attention deficits, the nonverbal structure of the task or systematic competence problems regarding belief reasoning) in the search task make its interpretation difficult; (3) there were several structural differences between the tasks that make it difficult to tell whether they tapped at comparable cognitive capacities. Ideally, however, a comparison of high and low first-person-involvement would involve structurally identical tasks. In addition, if self-perspective played a significant role in belief reasoning, this should be demonstrated for different FB test formats as superficially quite diverse FB tasks have been shown to correlate with EF in comparable ways. Study 2 therefore investigated the relation of belief-reasoning and EF more systematically using two different false belief test formats. For each format – the interactive search task and a standard location change FB test – structurally identical versions with high and low first-person-involvement were used.

3. Study 2

The variation of self-perspective applied to the two false belief test formats is illustrated in Table 1: First, the search task from study 1 was modified to overcome the problems of high dropout rates and floor effects. Secondly, a standard false-belief task was developed, that equally comprised of a reality known (RK) and a reality unknown (RU) condition but that was comparable to traditional FB measures regarding its surface structure. Moreover, to investigate to what extend previous results generalize to other executive functioning measures, three different EF measures were used.

3.1. Method

3.1.1. Participants

In the second study, 62 native German speaking children were included (33 females), aged 49;6 months (*SD* 4,987; range 42–59).

3.1.2. Design

Each child was tested by two experimenters in an interactive play setting. Children received three executive function tasks in a fixed order, alternating with four false-belief tasks (for an overview of all tasks see Table 5). In addition to the search FB task from study 1, two versions of a traditional location-change task (standard FB task) were administered (reality known condition and reality unknown condition). In each condition the child consecutively completed two trials (with different protagonists and targets in the traditional change of location task) and thus received a total score of 0–2 for each of the FB tasks. The order of FB tasks (standard and search) was counterbalanced across children, as was the order of RK and RU condition within both tasks. To control for verbal ability and working memory, children received a vocabulary test at the beginning and a working memory task at the end of the session.

3.1.3. Materials and procedure

3.1.3.1. False-belief tasks. (1) *Traditional location-change false-belief task* (see Fig. 4). The standard FB task consisted of two versions (with two trials each):

Reality Known trials represented 'standard versions' of location change FB stories, acted out with plastic figures. In these versions the child observed the hiding of a target in one box and the change of location into another box and therefore had a representation about the real location of the target. The child's own perspective in this case differed from the false belief of the protagonist that had to be deduced.

For *Reality Unknown* trials the task was modified in such a way that the two boxes were hidden behind drapes and the child therefore was ignorant about the real location of the target while observing the switch of the two boxes. Not until the protagonist went to one of the two boxes to get his target the child was able to deduce the real location by taking into account the false belief of the protagonist.
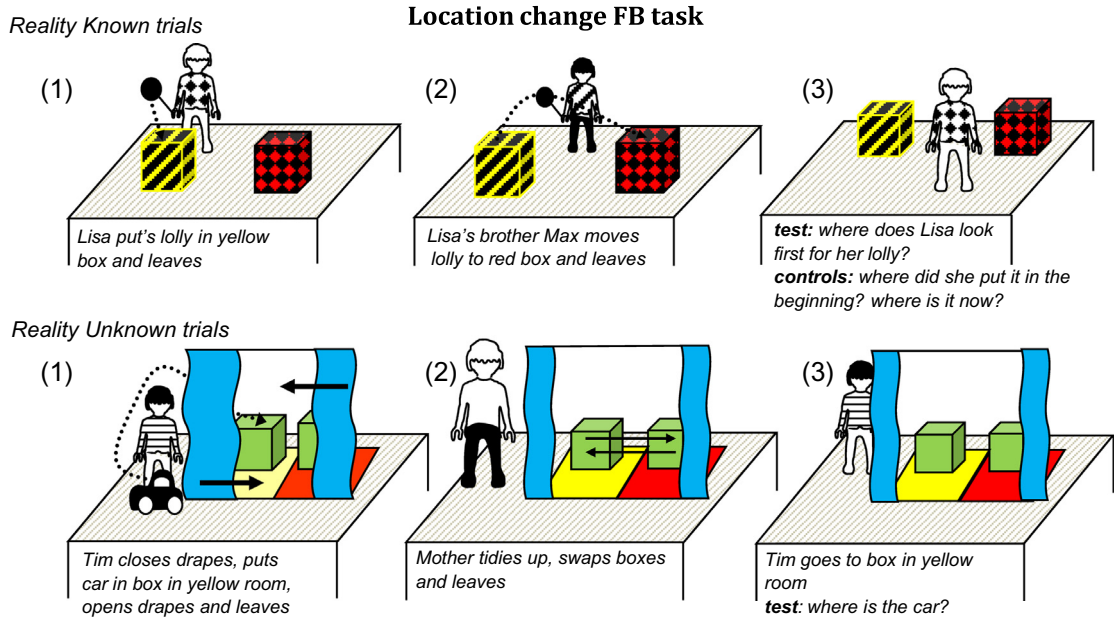
Location change FB task

Reality Known trials

(1) Lisa put's lolly in yellow box and leaves

(2) Lisa's brother Max moves lolly to red box and leaves

(3) **test:** *where does Lisa look first for her lolly?* **controls:** *where did she put it in the beginning? where is it now?*

Reality Unknown trials

(1) Tim closes drapes, puts car in box in yellow room, opens drapes and leaves

(2) Mother tidies up, swaps boxes and leaves

(3) Tim goes to box in yellow room **test**: *where is the car?*

Fig. 4. Schematic event sequences of the standard Location change FB task used in study 2.

(2) *Search false-belief task.* The search task from study 1 was modified in several important aspects (see Fig. 2 and Appendix B): Familiarization and control trials were reduced, resulting in two to four familiarization trials, two reality known and two reality unknown trials and three blocks of one or two control trials. As in study 1, children had to pass at least one of two trials in each control block. The order of trials was changed in such a way that RK and RU trials were intermitted by one EF task and invisible displacement and ignore communicator control trials were conducted at the end of the task, as not to confuse the child with the many different trials and not derogate the trust in the communicator's signal. In contrast to study 1, a by E2 animated puppet ('Susi') was used as communicator and instead of pointing to the box, the puppet placed a marker on it (both was assumed to be easier for children to ignore[4]). To simplify the task, additional verbal hints were given: When Susi looked under the table, she stated: "I know where the sticker is! I've seen it!" and when swapping the boxes/moving the sticker, E1 stated: "Look! Susi now can't see anything!" The RK test question was changed by asking: "Where does Susi think the sticker is?" and unlike study 1, if a child did not answer to this question at all or answered "I don't know", she was given a second chance by asking: "Where is Susi going to point in a moment?"

3.1.3.2. Executive functioning tasks. EF tasks consisted of two stroop-like tasks and the bear-dragon task as a measure of conflict inhibition.

(1) *Day-Night* (Gerstadt, Hong, & Diamond, 1994). Children were explained they had to say "day" if shown a

white card depicting a sun by E1 and "night" if shown a black card depicting moon and stars. They received two practice trials for each card in alternating order (with feedback) and 16 test trials in fixed random order without feedback. To carry out the task, children had to perform at least one "day" and one "night" practice trial correctly.

(2) *Grass-Snow* (Carlson & Moses, 2001). E1 introduced a green and a white field attached on a Din A4 cardboard in front of them. Children were instructed to touch the green field when E1 said "snow" and the white field when she said "grass". Children received two practice trials for each word in alternating order (with feedback) and 16 test trials in fixed random order without feedback. To administer the task, children had to perform one "grass" and one "snow" practice trial correctly.

(3) *Bear-dragon* (see Appendix C). The same bear-dragon task as in study 1 was used. Reliability was calculated for a 30% sample of randomly selected tapes coded by a second, independent observer. Weighted κ was .98 for dragon trials.

3.1.3.3. Working memory. The same counting and labeling task as in study 1 was used.

3.2. Results and discussion

Over all, data from 62 children was used in the final analysis. One child did not answer either of the standard reality known test questions and three children failed to answer the standard RK control questions, resulting in 58 valid scores on the standard RK task. Four children were uncooperative in the search reality known trials, 19 children failed to answer one of the search RK control

---

[4] Hala and Russell (2001) provided evidence that a marker is easier for children to resist than a pointing gesture.
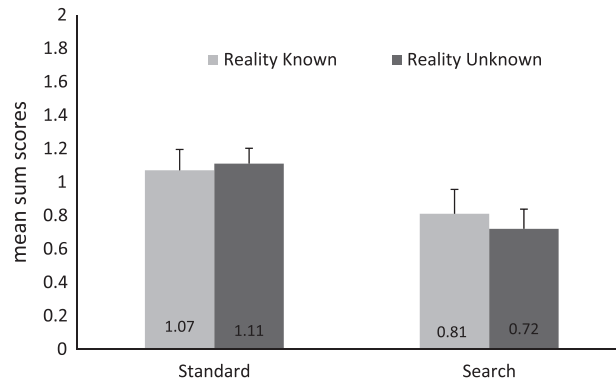
Fig. 5. Mean sum scores of the false-belief tasks (error bars represent SEM).

question or failed in a block of two control trials (trust, invisible displacement or ignore communicator), resulting in 42 valid scores on the search RK task and 50 on the search reality unknown task.

Again, one-tailed tests were conducted when testing theoretically and empirically derived directed a priori hypotheses as described above.

3.2.1. Performance on false-belief tasks

The mean sum scores for the four false-belief tasks are depicted in Fig. 5. A 2 × 2 repeated measurement ANOVA revealed a significant effect of task (standard vs. search) ($F(1,40) = 13.90$, $p < .001$), but no significant effect of condition (reality known vs. reality unknown) or interaction. Among all FB tasks only the standard RK and the search RU task did not correlate significantly ($r(47) = .17$, $p = .12$, one tailed), all other tasks were related ($rs > .30$; $ps < .05$, one tailed). For the purpose of comparing children's performance in RK and RU versions, a proportion correct score (0–1) was computed, calculating the proportion of correct responses in the four RK trials (two standard (0–2) and two search (0–2) trials) for a RK score and analogously in the four RU trials for a RU score. These two scores were correlated ($r(62) = .37$, $p < .01$, one tailed).

3.2.2. Performance on EF-tasks and working memory task

In the day-night task, 23 children failed to perform the necessary practice trial correctly, mostly because they could not remember the rules. In the grass-snow task, 4 children did not pass the required practice trials. Owing to experimental error, 3 children did not provide values in the bear-dragon task and data of 7 children was removed because children refused the task ($N = 5$) or could not clearly indicate that they understood the rules ($N = 2$). Mean sum scores were 13.31 ($SD = 4.10$) for Day-Night and $M = 14.07$ ($SD = 4.20$) for Grass-Snow. The number of practice dragon trials children needed until they succeeded ($M = 2.06$, $SD = 1.81$) was significantly negatively correlated ($r(52) = -.53$, $p < .001$) with the sum score for the dragon trials ($M = 10.50$, $SD = 6.13$). As in study 1 a composite BD score was computed. The mean score in WM (counting and labeling task) was 2.48 ($SD = 1.84$). WM was correlated with the composite BD score and Grass-Snow. No relations were found between EF tasks (see Table 5).

3.2.3. Relations between executive functioning tasks, working memory and reality known and reality unknown false-belief proportion scores

The raw and partial (age- and vocabulary-controlled) correlations between EF-tasks and RK and RU proportion correct scores are presented in Table 5. No significant correlations were found with Day-Night or Grass-Snow. These results, in line with several other studies, suggest that conflict inhibition indeed plays a special role in ToM reasoning. Note, however, that this result in our study might be explained by the high dropout rates in the Day-Night task and performance close to ceiling in the Grass-Snow task.

WM correlated with both FB scores, this correlation remained marginally significant only for the reality unknown score when age and vocabulary were controlled for.

Regarding conflict inhibition (BD) and it's relation to FB, the reality known score was clearly and significantly correlated with the composite BD score ($r(52) = .40$, $p < .005$). This correlation remained significant even when age and vocabulary was controlled for. This was not the case for the RU score, which was not associated with the BD measure in raw ($r(52) = .21$, ns) or partial correlation ($r(48) = .10$, ns). These results concerning conflict inhibition were confirmed by separate multiple regression analyses with FB scores as dependent measures (see Table 6). Age, vocabulary, BD composite score and WM were all entered simultaneously. The BD composite score and vocabulary predicted the RK score. For the RU score only WM turned out to be a marginally significant predictor.

Control analyses. One alternative factor that could account for a low correlational association between the reality unknown proportion score and the BD composite score would be a higher random variance and thus more noisy data in the reality unknown tasks compared to the reality known tasks. To test this possibility, we analyzed the contingency in performance on both trials for each task. Interestingly, both RK tasks showed a homogeneous

Table 6
Multiple regressions for predicting false belief scores.

	Reality known score				Reality unknown score			
	B	SEB	Beta	T	B	SEB	Beta	T
Age	.003	.01	.04	.26	.01	.01	.16	1.02
Vocabulary	.04	.02	.29	2.06[*]	.01	.02	.07	.50
BD composite score	.06	.03	.28	1.86[*]	.01	.03	.05	.32
WM	−.002	.04	−.01	−.05	.04	.03	.22	1.41[+]
	$R = .47, R^2 = .23$, adjusted $R^2 = .16$				$R = .37, R^2 = .14$, adjusted $R^2 = .06$			

Note. All p values are one tailed: [**]$p < .01$; [*]$p < .05$; [+]$p < .10$. BD, bear-dragon; WM, working memory.

pattern for the first and second trial (standard: $r(61) = .81$, $p < .001$; search: $r(58) = .71$, $p < .001$), as did the search reality unknown FB task ($r(62) = .45$, $p < .001$). In the standard RU FB task, in contrast, performance across trials did not correlate ($r(62) = .05$, ns). However, when eliminating this task and analyzing the EF-ToM correlations only for the search tasks, the same pattern of results emerged: The BD composite score significantly correlated only with the RK search score ($r(38) = .37$, $p < .01$) but not with the RU search score ($r(42) = .25$, $p = .06$). When age and vocabulary was controlled for, only the correlation with the RK search score persisted ($r(34) = .31$, $p < .05$), but disappeared for the RU search score ($r(34) = .03$, $p = .87$).

Also, when changing the coding for the standard reality unknown task such that children receive a score of 1 only if they succeed in both trials, the resulting pattern of correlations was the same as for the search task: no significant correlation is found with the composite BD score ($r(52) = .16$, $p = .25$) and the relation completely disappears when controlling for age and vocabulary ($r(48) = .08$, $p = .56$).

All in all, study 2 thus eliminated crucial methodological problems of study 1 and showed that self-perspective involvement plays a crucial role in the relation between EF and belief reasoning: the proportion score of those tasks in which the child herself had an own perspective about the state of affairs at question, conflicting with that of the protagonist, was clearly correlated with conflict inhibition as a measure of EF. In contrasts, the proportion score of those tasks in which the child herself was ignorant about the location of the target at the moment when the character's false belief had to be ascribed did not do so in the same way.

4. General discussion

Consistent with much previous work, the two present studies revealed significant relations of executive function and theory of mind. Regarding the exact nature of this relation – whether EF relates specifically to ascribing truth-aiming attitudes, to perspective problems more generally, or to such perspective problems with a self-perspective to be inhibited in particular – the present findings suggest the following:

First, the results of Study 1, converging with previous findings, are incompatible with the claim that EF is spe-

cifically involved only in ascribing truth-aiming attitudes due to the fact that such ascriptions have to deal with and inhibit normative defaults of truth: When two strongly subjective incompatible attitudes had to be ascribed, the EF-ToM relation held in much the same ways both when the attitudes were beliefs and when they were desires.

Second, the present findings are compatible with the possibility that EF is involved in the coordination of different perspectives more generally. In the present study and in previous work (Rakoczy, 2010), tasks presenting perspective problems, both involving incompatible beliefs and incompatible desires, were (mostly) related to EF. Similarly, in other work, only such tasks that embodied perspective problems (FB tasks and false sign tasks) were related to EF, but not structurally analogous and other types of ToM tasks that did not present perspective problems – such as the false photo task, or tasks of simple, non-perspectival desire or pretense ascription (Moses et al., 2005; Sabbagh, Moses, et al., 2006; Sabbagh, Xu, et al., 2006).

Third, these findings suggest that in fact EF might be particularly relevant to solving tasks that require the coordination of two incompatible perspectives one of which is one's own current perspective, be that a cognitive or a conative perspective. The present data thus complement existing neuropsychological (e.g. Samson et al., 2005) and gerontological (Bailey & Henry, 2008) findings by indicating that already early in development inhibition plays a central role in ascribing mental states conflicting with the ascriber's own perspective. The present two studies thus contribute to clarifying the nature of the relation between executive function and theory of mind. At the same time, however, some fundamental questions remain to be answered in future research:

First of all, the present findings, as those of much previous developmental work, remain correlational and therefore difficult to interpret in causal terms. Desirable for future search would be, for example, less ambiguous longitudinal correlational designs (testing whether EF at time 1 predicts ToM at time 2 but not vice versa; see Carlson et al., 1998), and crucially, experimental studies manipulating EF demands directly, either in dual task designs (e.g. Qureshi et al., 2010), or by varying the structure of the ToM tasks regarding their EF complexity.

A second question is whether the relation of EF to some cognitive and conative perspective problems found here

and elsewhere extends to *all* kinds of perspective problems. Remember that perspective problems are defined as cognitive problems requiring the coordination of representations the propositional contents of which cannot be conjoined without recourse to representational standpoints (Perner et al., 2002, 2003, 2005). Crucially, therefore, perspective problems are not confined to understanding mental representations, but apply equally to understanding non-mental representations such as pictures, language or maps. The question thus is whether EF is involved in such perspective problems in the same way. So far there is only one finding that suggests that this might indeed be the case, namely that EF was related to understanding false signs as much as to understanding false beliefs (Sabbagh, Moses, et al., 2006; Sabbagh, Xu, et al., 2006). More data with systematic contrasts of different kinds of perspective problems are thus needed. A more specific question in this context is whether EF is involved in understanding non-mental perspective problems only (or mostly) in cases where there is a first-person perspective involved that needs to be inhibited. This is the case, for example, in the false sign test where the child knows what is the case (e.g. an object O is at location L) and has to describe the sign as indicating something different (the signs says O is not at L, but at M). This is not necessarily the case, in contrast, in other types of non-mental perspective problems such as understanding conceptual or spatial perspectives embodied in different linguistic construals or in pictures (see Perner et al., 2002). For example, understanding two pictures depicting A being in front of B and B being in front of A, respectively, while having no view on A and B at the same time poses a perspective problem without any of the perspectives being one's current own perspective. Future research will need to explore to which degree EF is involved in such perspective coordination as well.

With this work we show that self-perspective inhibition is a crucial factor in the EF-ToM-relation. However, EF demands might be at place for various structural characteristics of the task (e.g. involving self-perspective, truth-aiming attitudes, mental state reasoning, etc.). Thus, EF involvement could be a relatively domain-general phenomenon but there might also be different kinds of domain-specific forms of inhibition involved in different kinds of perspective problems. To test this, in future research, more comprehensive designs will be required to isolate different forms of EF and different forms of reasoning about perspectives and to understand how they play together in complex belief–desire reasoning (for computational proposals regarding inhibitory combinations, see the work of Leslie and colleagues, e.g. Friedman & Leslie, 2004; Leslie & Polizzi, 1998; Leslie et al., 2005).

More generally, what we need to make progress with regard to these issues are more elaborated and detailed information processing models as to how different kinds of perspective problems relate to each other and to EF in terms of underlying cognitive processes. Such models will have to clarify if the very same cognitive processes that underlie theory of mind also underlie the understanding of the broader class of perspective problems,

and whether these processes in turn share a cognitive deep structure with processes underlying EF. Different classes of accounts in the literature on theory of mind and its relation to domain-general cognitive processes can serve as sources for the elaboration of such models (see, e.g. Moses & Tahiroglu, 2010; Perner & Lang, 1999 for review):

The first class of accounts pictures EF or related domain-general processes as substantial part of the very ToM competence and therefore as a prerequisite for the ontogenetic emergence of ToM: One family of such accounts in fact traces back both EF and perspective understanding to the same underlying cognitive deep structures such as handling certain forms of complex relational reasoning (Andrews, Halford, Bunch, Bowden, & Jones, 2003; Halford, Wilson, & Phillips, 1998) or hierarchical, recursive thinking (Frye, Zelazo, & Palfai, 1995). Other accounts in this class view EF processes such as inhibition and working memory themselves as the foundations of theory of mind and other forms of perspective understanding (e.g. Russell, 1996).

A competing class of accounts considers core ToM competence to be domain-specific, perhaps even modular and thus independent of domain-general phenomena such as EF (e.g. Leslie, 2005; Leslie et al., 2005). EF plays a role in such accounts only as a performance factor in explicit tasks posing high inhibitory and linguistic demands.

Finally, recent two-system-theories (e.g. Apperly & Butterfill, 2009) provide a more comprehensive and differentiated synthesis: such accounts distinguish between ontogenetically early developing systems of tracking some simple forms of mental states and later developing full blown ToM capacities that are flexible, but cognitively demanding. With regard to the role of EF, these models suggest that only the latter essentially involve domain-general cognitive capacities such as EF and language, whereas the early system operates rather fast, automatic and cognitively efficient.

These different classes of theoretical accounts, it is true, yield rather different views regarding the cognitive of structure of early and implicit forms of social cognition (incorporating full-blown modular perspective-taking competence not yet expressed in explicit tasks according to some accounts; presenting precursors to full-blown perspective-taking only according to others). But regarding later explicit forms of mentalizing and perspective understanding, the implications of these different accounts might show some broader convergence. Essentially, the accounts might converge towards the idea that coordinating different representations/perspectives in the service of explicitly reasoning about them, contrasting and ascribing them to others or oneself, requires domain-general representational capacities, in particular when such coordinating involves conflicts between representations entertained by oneself and those to be ascribed.

In fact, such a picture might bear interesting relations with broader theoretical accounts on the role of handling and coordinating perspectivity in our higher cognitive architecture. A recent approach, for example, identifies

multiperspectivity as a core feature and essential architectural requirement of any complex information-processing system that is faced with the challenge of integrating the outputs of different specialized subsystems (Mausfeld, 2011). Information-processing models drawing on such theoretical accounts might allow us in the future to better understand the cognitive processes involved in computing, coordinating, comparing, contrasting and ascribing perspectives and how they relate to executive functions and other domain-general processes.

One more specific question to be clarified in this context is whether different forms of EF are differentially involved in different sub-processes of perspective coordination. The present study, consistent with much previous research, found that a so-called conflict inhibition task, involving both working memory and inhibition, was most clearly related to ToM. It is possible, though that on closer inspection different forms of ToM and perspective understanding turn out to differentially involve different forms of EF such as working memory, inhibition, set shifting and planning.

Acknowledgements

Appendix A

Structure of the two versions (first and third person) of the conflicting-desires task, which differed only in Section 5.

1 **Introduction of the 'chance machine'**
E put a marble into the upper tube repeatedly and made it emerge in (seemingly) random fashion

2 **Introduction of the puppet**
'Rudi' is introduced and the booklet belonging to Rudi and the child together

3 **Explanation of the game**
E explained basic logic to the child: 'Look! Here is a sticker and there is one (pointing to left/right sticker). Now we put the marble in. The marble sometimes comes out here (points). Then this sticker here goes in the booklet'

4 **8 warm-up trials**
In trial 1+2 (without Rudi) E pinned 2 interesting sticker on the board, dropped the marble and showed the child which sticker it won
In trial 3+4 E pinned an interesting and a boring

sticker on the machine. Child won first the interesting, in the next trial the boring sticker. In each trial child is asked which sticker it won
In trial 5+7 child played alone, in trial 6+8 Rudi played. In one of the trials child/Rudi won, in the other child/Rudi lost. In each trial child/Rudi was asked, which sticker child/Rudi wants to be in the booklet, in which box the marble has to come for this result, (after E had put the marble in the machine) if the child/Rudi now is now happy or sad and which sticker now goes into the booklet. If child answered incorrectly, E gave feedback and provided the correct answer

5 **2 test trials**
E pinned an interesting and a boring sticker/two interesting sticker on the machine
Child/Peter was asked which sticker should be in the booklet and answered X. Rudi/Susi exclaimed "No!" and expresses the opposite desire. Child/Peter was asked again and Rudi/Susi repeated his/her "No!" and expressed the opposite desire again
E held marble over machine and asked child: 'Where do you/Peter want the marble to go?' (Q1a) and 'Where does Rudi/Susi want the marble to go?' (Q1b) E put marble in machine and pointed to resulting box: 'Marble came here. Are you/Is Peter now happy or sad?' (Q2a) and 'Is Rudi/Susi now happy or sad?' (Q2b)
The order of these two questions and which character won on the first trial were counterbalanced
Finally, the child could put sticker in the booklet

Appendix B

Sequence of the search FBT.

	Study 1	Study 2
1	**Introduction of the puzzle game** Children received a puzzle without pieces and were explained the rules of the game. If children found a puzzle piece, they might put it in the puzzle	**Introduction of the sticker game** Children received a sticker album and were explained the rules of the game. If children found a sticker, they might put it in their sticker album
2	**3–8 Familiarization trials** Consisted of (d) and (f) only with two different boxes	**2–4 Familiarization trials** Consisted of (d) and (f) only with two different boxes in the first trial
3	**1–2 Invisible displacement control**	**2 RK trials or 2 RU trials**

Appendix B (*continued*)

	Study 1	Study 2
	trials Consisted of (d), then (f) and then (e) and ensured children's ability for object permanence	**(counterbalanced)** – *Intermission by an EF task* –
4	**1–2 Ignore communicator control trials** Consisted of (a), (b) and (c) but with E2 pointing to the box in (c) and the test question of (f)	**1–2 Trust control trials** Consisted of (a), (b) but without location change and (f) and should enhance children's trust in Susi's pointing
5	**1–2 Trust control trials**[a] Consisted of (a), (b) but without location change and should enhance children's trust in E2's pointing RU test question: Where is the sticker?	**2 RK trials or 2 RU trials in (counterbalanced)** RK test question: Where does Susi think the sticker is? (if children didn't answer: Where is Susi going to point?)
6	**2 RK trials or 2 RU trials (counterbalanced)** RK test question: Where is (name of E2) going to point? RU test question: Where is the puzzle piece?	**1–2 Invisible displacement control trials** Consisted of (d), then (f) and then (e) and ensured children's ability for object permanence
7	**1–2 Trust control trials**	**1–2 Ignore communicator control trials** Consisted of (a), (b) and (c) but with Susi placing her marker in (c) and the test question of (f)
8	**2 RK trials or 2 RU trials (counterbalanced)**	

[a] Figueras-Costa and Harris (2001) introduced trust control trials to re-establish the communicator's reliability. As in their study, this was especially important in our first study, because during the control tests before, children had had some experience of the communicator no longer being helpful. Without trust controls it cannot be excluded that children learn the simple strategy to choose the box that was not marked by the communicator in all trials in which the communicator leaves the room. Even if in the second study, the ID and IC control trials were conducted at the end of the task, we included trust control trials because also in the first test trials, children may experience the communicator as not reliable.

Appendix C

Procedure of the bear-dragon task.

Children were first asked to perform 10 simple movements or gestures (such as 'touch your ears'), ensuring that they were able to carry those out. Then E introduced a 'good' and a 'mean' puppet in the following way: 'The bear is very nice; we do everything he says, because he's our friend. But this nasty dragon isn't our friend at all. We do not do what he says.' Two kinds of practice trials followed. First, E animated the bear, spoke on his behalf with a friendly voice and gives a command: 'touch your tummy'. Second, E animated the dragon, spoke with a low gruff voice and said: 'put your hand on your neck'. Children frequently failed this dragon practice trial. In such cases, E gave negative feedback and repeated the rules. If a child did not pass the fifths of those practice trials, E explained to the child that she will help her on the next trial by fixing the child's hands with hers on the table. Children then got a sixth dragon practice trial, in which they were detained by E to carry out the action. Before the test trials began, children's understanding of the rules was tested. Then 5 bear and 5 dragon test trials followed in alternating order. After the 5th trial, E reminded children of the rules.

Appendix D

What does it mean to understand a perspective problem and which tasks measure such understanding?

A perspective problem can be formally defined in the following way (Perner et al., 2003): a situation presents a perspective problem if (and only if) there are at least two representations involved (such as mental attitudes, pictures, and sentences) whose content (e.g., p, q) cannot be combined by simple conjunction (p AND q) to yield a consistent composite representation. Regarding visual perspectives, for example, imagine two objects A and B seen from opposite sides of a room: "A is in front of B" and "A is behind B". The simple conjunction "A is in front of and behind B" is inconsistent. Rather, to yield a consistent composite representation each content has to be relativized to a standpoint or perspective or marked as a (mental) representation ("A is in front of B seen from perspective 1, but behind B seen from the opposite perspective...").

Given this formally precise definition, one can see which tasks do and which don't necessarily measure an understanding of perspective problems (see also Moll et al., 2013; Sabbagh, Moses, et al., 2006; Sabbagh, Xu, et al., 2006).

Task	Setup	Results	Why does it involve a perspective problem?
1. Tasks that require an understanding of perspective problems to be solved			
Standard false belief tasks (Wimmer & Perner, 1983)	*Time 1*: Protagonist P sees that object O is in location L1	• Typically developing children succeed from around age 4	"Object O is in location L1" (reality) and "object O is in location L2" (content of protagonist's false belief) cannot be combined without recourse to the protagonist's standpoint ("O is in location L1 but protagonist believes it is in L2")
	Time 2: In the absence of the P, O is transferred to L2	• Robust correlations with EF	
	Time 3: P returns and wants to have O	• Deficits in autistic children	
	Test: "where will she look for O first?"		
Reality unknown false belief tasks (Call & Tomasello, 1999)	*Time 1*: Protagonist P sees (but subject does not see) in which of two identical containers C1 and C2 object O is placed	• Typically developing children succeed from around age 4	"Object O is in location C1" (content of the protagonist's communicative act) and "object O is in C2" (reality inferred from the protagonist's communicative act and the presumption that it is mistaken) cannot be combined without recourse to the protagonist's standpoint ("O is in C2 but protagonist believes it is in C1")
	Time 2: In the absence of the PC1 and C2 are swapped		
	Time 3: P returns and points to container C1 (saying to the child "here it is")		
	Test: where will subject look for O		
Conflicting desires tasks (e.g. Moore et al., 1995; Rakoczy et al., 2007)	A lottery device (marble that can go to locations L1 or L2) determines an outcome (L1 → O1/L2 → O2). Two players P1 and P2 express divergent and incompatible desires for O1 and O2	• Typically developing children succeed from around age 3	"The marble goes to location L1" and "The marble goes to location L2" can only be combined by relativizing the contents to conative perspectives: "A wants the marble to go to L1, but B wants it to go to L2"
	Test:	• Correlations with EF found	
	Q1: "Where does P1 want the marble to go, and where does P2 want it to go?"		
	Q2: after the marble has gone to one location: "Is P1 happy or sad now, and is P2 happy or sad now?"		
False sign test (Parkin, 1994; Leekam, Perner, Healey, & Sewell, 2008; Perner & Leekam, 2008; Sabbagh, Moses, et al., 2006; Sabbagh, Xu, et al., 2006)	Sign S is supposed to show location of object O	• Typically developing children succeed from around age 4 (saying "L1")	"O is in location L1" and "O is in location L2" cannot be combined without reference to the "standpoint" of the sign: "O is really in L2, but the sign says/signals (wrongly/misleadingly) that it is in L1"

Task	Setup	Results	Why does it involve a perspective problem?
	Time 1: O is in location L1 and initially S points towards L1 *Time 2*: Someone transfers O to L2 without changing the sign *Test*: "Where does the sign say O is?"	• Robust correlations with EF • Deficits in autistic children	
Visual perspective taking level II (e.g. de C. Hamilton, Brindley, & Frith, 2009; Flavell, Everett, Croft, & Flavell, 1981)	Child has to judge how a scene looks from 2 different perspectives (typically her own and someone else's)	• Typically developing children succeed from around age 4 • Deficits in autistic children	"A is in front of B" and "A is behind B" can only be combined by relativizing them to perspectives ("A is in front of B seen from perspective 1, but behind B seen from the opposite perspective…")

2. Tasks that do not necessarily require an understanding of perspective problems but can be solved in simpler ways

Task	Setup	Results	Why does it involve a perspective problem?
Simple belief tasks (e.g. Wellman & Bartsch, 1988)	"Protagonist wants to find object O which might be in location L1 or in location L2. He thinks it is in location L1 *Test*: "Where will he look for it?"	Children master this tasks well before passing standard FB tasks	This does not require an understanding of perspective problems since there is no conflict between any two contents at all – the only content brought into play is "object is in L1" given as the content of the protagonists belief
False photo tasks (Leekam & Perner, 1991; Sabbagh, Moses, et al., 2006; Sabbagh, Xu, et al., 2006; Zaitchik, 1990)	*Time 1*: object O is in location L1 → a Polaroid picture is taken and needs some time to develop	• Typically developing children succeed from around age 4	The reason why this task does not represent a perspective problem lies in the temporal structure of the scenario: Photos taken at one time do not aim at representing how the world is at some other time. In the current story, the photo at time 3 represents the state of the world at time 1 – it is not false but outdated (see Perner and Leekam, 2008; Sabbagh, Moses, et al., 2006; Sabbagh, Xu, et al., 2006). That means that "O is L1" and "O is in L2" can be combined without relativizing them to standpoints. Rather, all that is needed is relativizing them to points in time: "O was in L1 at time 1 (when the photo was taken), but is in L2 at time 3" (see also Moll et al., 2013)
	Time 2: object O is transferred from L1 to L2 *Time 3*: picture has now developed *Test*: "Where in the picture will O be?"	• No correlations with EF • No deficits in autistic children	

(*continued*)

Task	Setup	Results	Why does it involve a perspective problem?
Simple preference tasks (e.g. Repacholi and Gopnik, 1997)	The child is faced with a situation in which she and someone else have different desires or preferences: For example, there are broccoli and crackers, and whereas the child prefers broccoli, the other person expresses her preference for broccoli *Test*: what will the child give to the other person when requested to "give me what I like	Children from around 18 months have been found to solve this task	This task does not represent a perspective problem as differences in desires do not yet amount to a perspective clash. Rather, the situation can be understood in a quasi-objectivist framework: "Broccoli is good in this person's mouth" but "crackers are good in my mouth". Broccoli being in one mouth and crackers in another are just two independent events and can thus be independently described as good or bad without the need for subjective relativization. In contrast, this strategy does not work for conflicting desires tasks as used in the present study. Here the situation has two potential outcomes (marble in L1 or marble in L2) that are mutually exclusive. Consequently, the propositional contents of the two desires ("marble goes to L1" and "marble goes to L2") are strictly incompatible, i.e., cannot both be fulfilled. Thus combining them requires relativizing them to conative subjective standpoints ("A wants the marble to go to L1, but B wants it to go to L2")
Simple desire and desire-based emotion understanding tasks (e.g. Wellman and Bartsch, 1988)	"Protagonist desires that p. Either p or non-p turns out to be the case" *Test*: "How will the protagonist feel?"	Children from around 2–3 years have been found to solve this task	This task does not represent a perspective problem since the situation can be understood in quasi-objectivist terms: that the protagonist desires p is simply understood along the following lines: it is good (for the protagonist) that p, and thus, if p is the case the protagonist, experiencing the good, is happy, sad otherwise To rule out such simpler interpretations on the part of the children, scenarios are need in which there are several desires that stand in contrast to each other like in the conflicting desires tasks used in the present study

References

Andrews, G., Halford, G. S., Bunch, K. M., Bowden, D., & Jones, T. (2003). Theory of mind and relational complexity. *Child Development, 74*(5), 1476–1499. http://dx.doi.org/10.1111/1467-8624.00618.

Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review, 116*(4), 953–970. http://dx.doi.org/10.1037/a0016923.

Apperly, I. A., Samson, D., & Humphreys, G. W. (2005). Domain-specificity and theory of mind: Evaluating neuropsychological evidence. *Trends in Cognitive Sciences, 9*(12), 572–577. http://dx.doi.org/10.1016/j.tics.2005.10.004.

Bailey, P. E., & Henry, J. D. (2008). Growing less empathic with age: Disinhibition of the self-perspective. *The Journals of Gerontology: Series B: Psychological Sciences and Social Sciences, 63B*(4), 219–226. http://dx.doi.org/10.1093/geronb/63.4.P219.

Birch, S. A. J., & Bloom, P. (2007). The curse of knowledge in reasoning about false beliefs. *Psychological Science, 18*(5), 382–386. http://dx.doi.org/10.1111/j.1467-9280.2007.01909.x.

Call, J., & Tomasello, M. (1999). A nonverbal false belief task: The performance of children and great apes. *Child Development, 70*(2), 381–395. http://dx.doi.org/10.1111/1467-8624.00028.

Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development, 72*(4), 1032–1053. http://dx.doi.org/10.1111/1467-8624.00333.

Carlson, S. M., Moses, L. J., & Hix, H. R. (1998). The role of inhibitory processes in young children's difficulties with deception and false belief. *Child Development, 69*(3), 672–691. http://dx.doi.org/10.2307/1132197.

Choe, K. S., Keil, F. C., & Bloom, P. (2005). Children's understanding of the Ulysses conflict. *Developmental Science, 8*(5), 387–392. http://dx.doi.org/10.1111/j.1467-7687.2005.00426.x.

Cohen, A., & German, T. (2013). Are mental representations special because they are mental or because they are representations? *Paper presented at the biennial meeting of the society for research in child development*, Seattle, USA.

Figueras-Costa, B., & Harris, P. (2001). Theory of mind development in deaf children: A nonverbal test of false-belief understanding. *Journal of Deaf Studies and Deaf Education, 16*(2), 92–102.

Flavell, J. H., Everett, B. A., Croft, K., & Flavell, E. R. (1981). Young children's knowledge about visual perception: Further evidence for the Level 1–Level 2 distinction. *Developmental Psychology, 17*(1), 99–103. http://dx.doi.org/10.1037/0012-1649.17.1.99.

Friedman, O., & Leslie, A. M. (2004). A developmental shift in processes underlying successful belief–desire reasoning. *Cognitive Science, 28*(6), 963–977. http://dx.doi.org/10.1207/s15516709cog2806_4.

Frye, D., Zelazo, P. D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cognitive Development, 10*(4), 483–527. http://dx.doi.org/10.1016/0885-2014(95)90068-x.

Gerstadt, C. L., Hong, Y. J., & Diamond, A. (1994). The relationship between cognition and action: Performance of children 312–7 years old on a stroop-like day-night test. *Cognition, 53*(2), 129–153. http://dx.doi.org/10.1016/0010-0277(94)90068-x.

Gordon, A. C. L., & Olson, D. R. (1998). The relation between acquisition of a theory of mind and the capacity to hold in mind. *Journal of Experimental Child Psychology, 68*(1), 70–83. http://dx.doi.org/10.1006/jecp.1997.2423.

Hala, S., & Russell, J. (2001). Executive control within strategic deception: A window on early cognitive development? *Journal of Experimental Child Psychology, 80*(2), 112–141. http://dx.doi.org/10.1006/jecp.2000.2627.

Halford, G. S., Wilson, W. H., & Phillips, S. (1998). Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. *Behavioral and Brain Sciences, 21*(6), 803–864. http://dx.doi.org/10.1017/s0140525x98001769.

de C. Hamilton, A. F., Brindley, R., & Frith, U. (2009). Visual perspective taking impairment in children with autistic spectrum disorder. *Cognition, 113*(1), 37–44. http://dx.doi.org/10.1016/j.cognition.2009.07.007.

Kochanska, G., Murray, K., Jacques, T. Y., Koenig, A. L., & Vandegeest, K. A. (1996). Inhibitory control in young children and its role in emerging internalization. *Child Development, 67*(2), 490–507. http://dx.doi.org/10.1111/j.1467-8624.1996.tb01747.x.

Leekam, S., Perner, J., Healey, L., & Sewell, C. (2008). False signs and the non-specificity of theory of mind: Evidence that preschoolers have general difficulties in understanding representations. *British Journal of Developmental Psychology, 26*(4), 485–497. http://dx.doi.org/10.1348/026151007x260154.

Leekam, S. R., & Perner, J. (1991). Does the autistic child have a metarepresentational deficit? *Cognition, 40*(3), 203–218.

Leslie, A. M. (2005). Developmental parallels in understanding minds and bodies. *Trends in Cognitive Sciences, 9*(10), 459–462. http://dx.doi.org/10.1016/j.tics.2005.08.002.

Leslie, A. M., & Polizzi, P. (1998). Inhibitory processing in the false belief task: Two conjectures. *Developmental Science, 1*(2), 247–253. http://dx.doi.org/10.1111/1467-7687.00038.

Leslie, A. M., German, T. P., & Polizzi, P. (2005). Belief–desire reasoning as a process of selection. *Cognitive Psychology, 50*(1), 45–85. http://dx.doi.org/10.1016/j.cogpsych.2004.06.002.

Mausfeld, R. (2011). Intrinsic multiperspectivity: Conceptual forms and the functional architecture of the perceptual system (PSYNDEXalert). In W. Welsch, W. Singer, & A. Wunder (Eds.), *Interdisciplinary anthropology. Continuing evolution of man* (pp. 19–54). Berlin: Springer.

Moll, H., Meltzoff, A. N., Merzsch, K., & Tomasello, M. (2013). Taking versus confronting visual perspectives in preschool children. *Developmental Psychology, 49*(4), 646–654. http://dx.doi.org/10.1037/a0028633.

Moore, C., Jarrold, C., Russell, J., Lumb, A., Sapp, F., & MacCallum, F. (1995). Conflicting desire and the child's theory of mind. *Cognitive Development, 10*(4), 467–482. http://dx.doi.org/10.1016/0885-2014(95)90023-3.

Moses, L. J., Carlson, S. M., Sabbagh, M. A., Schneider, W., Schumann-Hengsteler, R., & Sodian, B. (2005). *On the specificity of the relation between executive function and children's theories of mind. Young children's cognitive development: Interrelationships among executive functioning, working memory, verbal ability, and theory of mind* (pp. 131–145). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Moses, L.J., Carlson, S.M., Stieglitz, S., & Claxton, L.J. (2003). *Executive function, prepotency, and children's theory of mind*. Unpublished manuscript.

Moses, L. J., & Tahiroglu, D. (2010). Clarifying the relation between executive function and children's theories of mind. In B. W. Sokol, U. Müller, J. I. M. Carpendale, A. R. Young, & G. Iarocci (Eds.), *Self and social regulation: Social interaction and the development of social understanding and executive functions* (pp. 218–233). New York, NY, US: Oxford University Press.

Parkin, L. J. (1994). Children's understanding of misrepresentation, Unpublished D. Phil Thesis, University of Sussex.

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

Perner, J., Brandl, J. L., & Garnham, A. (2003). What is a perspective problem? Developmental issues in belief ascription and dual identity. [Was ist ein Problem der Perspektive? Entwicklungspsychologische Fragen bei Annahmenzuschreibung und dualer Identität]. *Facta Philosophica, 5*, 355–378.

Perner, J., & Lang, B. (1999). Development of theory of mind and executive control. *Trends in Cognitive Sciences, 3*(9), 337–344. http://dx.doi.org/10.1016/S1364-6613(99)01362-5.

Perner, J., & Leekam, S. (2008). The curious incident of the photo that was accused of being false: Issues of domain specificity in development, autism, and brain imaging. *The Quarterly Journal of Experimental Psychology, 61*(1), 76–89. http://dx.doi.org/10.1080/17470210701508756.

Perner, J., Leekam, S. R., & Wimmer, H. (1987). Three-year-olds' difficulty with false belief: The case for a conceptual deficit. *British Journal of Developmental Psychology, 5*(2), 125–137. http://dx.doi.org/10.1111/j.2044-835X.1987.tb01048.x.

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

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*(3–4), 1451–1472. http://dx.doi.org/10.1016/s0885-2014(02)00127-2.

Perner, J., Zauner, P., & Sprung, M. (2005). What does 'that' have to do with point of view? Conflicting desires and 'want' in German. In J. W. Astington & J. A. Baird (Eds.), *Why language matters for theory of mind* (pp. 220–244). New York, NY, US: Oxford University Press.

Qureshi, A. W., Apperly, I. A., & Samson, D. (2010). Executive function is necessary for perspective selection, not level-1 visual perspective calculation: Evidence from a dual-task study of adults. *Cognition, 117*(2), 230–236. http://dx.doi.org/10.1016/j.cognition.2010.08.003.

Rakoczy, H. (2010). Executive function and the development of belief–desire psychology. *Developmental Science, 13*(4), 648–661. http://dx.doi.org/10.1111/j.1467-7687.2009.00922.x.

Rakoczy, H., Harder-Kasten, A., & Sturm, L. (2012). The decline of theory of mind in old age is (partly) mediated by developmental changes in domain-general abilities. *British Journal of Psychology, 103*(1), 58–72. http://dx.doi.org/10.1111/j.2044-8295.2011.02040.x.

Rakoczy, H., Warneken, F., & Tomasello, M. (2007). "This way!", "No! That way!"—3-year olds know that two people can have mutually incompatible desires. *Cognitive Development, 22*(1), 47–68. http://dx.doi.org/10.1016/j.cogdev.2006.08.002.

Reed, M. A., Pien, D. L., & Rothbart, M. K. (1984). Inhibitory self-control in preschool children. *Merrill-Palmer Quarterly, 30*(2), 131–147.

Repacholi, B. M., & Gopnik, A. (1997). Early reasoning about desires: Evidence from 14- and 18-month-olds. *Developmental Psychology, 33*(1), 12–21. http://dx.doi.org/10.1037/0012-1649.33.1.12.

Robinson, E. J., & Mitchell, P. (1995). Masking of children's early understanding of the representational mind: Backwards explanation versus prediction. *Child Development, 66*(4), 1022–1039. http://dx.doi.org/10.1111/j.1467-8624.1995.tb00920.x.

Russell, J. (1996). *Agency: Its role in mental development*. Oxford, England: Erlbaum (UK), Taylor & Francis, Publ..

Sabbagh, M. A., Moses, L. J., & Shiverick, S. (2006). Executive functioning and preschoolers' understanding of false beliefs, false photographs, and false signs. *Child Development, 77*(4), 1034–1049. http://dx.doi.org/10.1111/j.1467-8624.2006.00917.x.

Sabbagh, M. A., Xu, F., Carlson, S. M., Moses, L. J., & Lee, K. (2006). The development of executive functioning and theory of mind: A comparison of Chinese and U.S. preschoolers. *Psychological Science, 17*(1), 74–81. http://dx.doi.org/10.1111/j.1467-9280.2005.01667.x.

Samson, D., Apperly, I. A., Kathirgamanathan, U., & Humphreys, G. W. (2005). Seeing it my way: A case of a selective deficit in inhibiting self-perspective. *Brain: A Journal of Neurology, 128*(5), 1102–1111. http://dx.doi.org/10.1093/brain/awh464.

Searle, J. R. (1983). *Intentionality: An essay in the philosophy of mind*. Cambridge Univ. Pr..

Wellman, H. M., & Bartsch, K. (1988). Young children's reasoning about beliefs. *Cognition, 30*(3), 239–277. http://dx.doi.org/10.1016/0010-0277(88)90021-2.

Wertz, A. E., & German, T. C. (2007). Belief–desire reasoning in the explanation of behavior: Do actions speak louder than words? *Cognition, 105*(1), 184–194. http://dx.doi.org/10.1016/j.cognition.2006.08.002.

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

Zaitchik, D. (1990). When representations conflict with reality: The preschooler's problem with false beliefs and "false" photographs. *Cognition, 35*(1), 41–68. http://dx.doi.org/10.1016/0010-0277(90)90036-J.