Cognition 129 (2013) 264–272

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/COGNIT

Perspective tracking in progress: Do not disturb

Paula Rubio-Fernández *

University College London, United Kingdom, CSMN, Oslo, Norway

ARTICLE INFO

Article history:
Received 18 December 2012
Revised 6 July 2013
Accepted 8 July 2013
Available online 14 August 2013

Keywords:
Theory of Mind
Indirect and direct false-belief tests
Perspective tracking

ABSTRACT

Two experiments tested the hypothesis that indirect false-belief tests allow participants to track a protagonist's perspective uninterruptedly, whereas direct false-belief tests disrupt the process of perspective tracking in various ways. For this purpose, adults' performance was compared on indirect and direct false-belief tests by means of continuous eye-tracking. Experiment 1 confirmed that the false-belief question used in direct tests disrupts perspective tracking relative to what is observed in an indirect test. Experiment 2 confirmed that perspective tracking is a continuous process that can be easily disrupted in adults by a subtle visual manipulation in both indirect and direct tests. These results call for a closer analysis of the demands of the false-belief tasks that have been used in developmental research.

1. Introduction

Thirty years of intense experimental research on Theory of Mind have revealed a paradoxical pattern of results: whereas infants as young as 7 months succeed on indirect measures of false-belief understanding (Kovács, Téglás, & Endress, 2010; Onishi & Baillargeon, 2005), children under 4 years fail all manner of direct false-belief tests (Wellman, Cross, & Watson, 2001; Wimmer & Perner, 1983). This so called 'developmental paradox' has received a great deal of attention recently, with a special issue of the *British Journal of Developmental Psychology* dedicated to this topic (see Low & Perner, 2012 and references therein; see also Perner & Roessler, 2012). The various theoretical models that have been proposed to account for the Theory of Mind paradox all focus on different developmental changes that are argued to affect performance in false-belief tests. Thus, Clements and Perner (1994) argue that young children have an implicit understanding of belief that only develops into an explicit understanding after age 4 years (see also Low, 2010; Perner & Roessler, 2012). Likewise, Apperly and Butterfill (2009) and de Bruin and Newen (2012) propose that indirect and direct false-belief tests tap distinct Theory of Mind systems with different developmental trajectories. Also, Baillargeon, Scott, and He (2010) claim that passing a direct test requires more complex false-belief reasoning than passing an indirect test and that the extra processes involved in direct tests are too demanding for young children (see also He, Bolz, & Baillargeon, 2012; Scott, He, Baillargeon, & Cummins, 2012).

While the various developmental changes that have been argued to underlie the Theory of Mind paradox are very different in nature (not all being specific to Theory of Mind development, for example), all theoretical models share the general assumption that indirect and direct false-belief tests are fundamentally comparable as these tasks have been appropriately adapted for infants and children respectively. In the present study I investigated an alternative view on the Theory of Mind paradox, which does not rest on such an assumption: the differential performance of infants and children may simply be an artifact of the different false-belief tasks that have been used with these two groups, irrespective of their cognitive development. If this were the case, adults might also perform differently in the two types of false-belief test despite their cognitive maturity.

* Tel.: +44 609 356 2227; fax: +44 609 258 1113.
E-mail address: p.rubio-fernandez@ucl.ac.uk

0010-0277/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.cognition.2013.07.005

In a direct false-belief test, the child witnesses an agent storing an object in location A. Next, the agent leaves the scene and a second agent transfers the object to location B. Finally, the child is asked where the first agent will look for the object when he or she returns. Unlike standard false-belief tasks, the indirect tests that have been used with infants are generally non-verbal and feature a single agent (or at least the role of the other agent is clearly secondary). Furthermore, unlike direct tests, indirect false-belief tests do not use a wh-question that requires making an overt choice between two possible locations and focuses the participant's attention on the target object (e.g., 'Where will Sally look for her doll?'; Baron-Cohen, Leslie, & Frith, 1985). Instead, indirect tests measure implicit indicators of false-belief understanding; e.g., spontaneous responses revealing anticipation of the outcome (Clements & Perner, 1994), surprise that the mistaken agent looks for the object in the correct location (Onishi & Baillargeon, 2005), or an attempt to help the mistaken agent find the object (Buttelmann, Carpenter, & Tomasello, 2009).

In a recent study we hypothesized that these differences in task design were likely to have a critical effect on young children's performance (Rubio-Fernández & Geurts, 2013). Specifically, we suggested that infants perform better in indirect false-belief tests because they are allowed to focus on the protagonist and track his or her perspective uninterruptedly throughout the task. Conversely, young children's perspective-tracking is disrupted by task manipulations in direct false-belief tests (e.g., by having to track the perspective of a second character or answer a question about the object). We tested this hypothesis with a direct false-belief test that was designed to allow young children to track the protagonist's perspective throughout the story and respond to an overt question that did not interrupt their perspective tracking. Contrary to all previous studies, 3 year-old children performed reliably above chance in this new false-belief task, with a success rate of 80%.

We view perspective tracking as an early ability that allows infants to form expectations about another person's actions based on observations of his or her behavior, even when the expected actions are based on false information (see, e.g., Senju, Southgate, Snape, Leonard, & Csibra, 2011). What kinds of mental processes and representations underwrite perspective tracking is largely irrelevant to our project. For example, it is immaterial whether this capacity involves the mental representation of beliefs (cf., e.g., Perner, 2010), whether it undergoes a conceptual change around age 4 years (cf., e.g., Wellman et al., 2001), or whether it requires one system for mind-reading or several (cf., e.g., Apperly & Butterfill, 2009). The only assumption that is critical for us is that perspective tracking is a continuous process that requires focusing on another's perspective and is therefore susceptible to disruption. This is what I will call 'the tracking hypothesis.'

The fact that children under 4 years can pass a suitably streamlined version of the standard false-belief task suggests that the tracking hypothesis might account for the alleged developmental paradox. Given that our account does not focus on Theory of Mind development between infancy and childhood, the aim of the present study was to further test the tracking hypothesis by investigating adults' perspective-tracking in indirect and direct false-belief tests by means of continuous eye-tracking.

Recent studies on adult Theory of Mind have shown that adults can sometimes track other people's perspectives even when doing so is not required by the task at hand (Cohen & German, 2009; Samson, Apperly, Braithwaite, & Andrews, 2010; Schneider, Bayliss, Becker, & Dux, 2012a). However, other studies with adults suggest that Theory of Mind processing is not automatic. Studies by Apperly and colleagues have shown that belief ascription is not an automatic process (Apperly, Riggs, Simpson, Samson, & Chiavarino, 2006; Back & Apperly, 2010) and that attributing false beliefs is more costly than attributing beliefs about reality (Apperly, Back, Samson, & France, 2008). Moreover, a recent study by Schneider et al. has shown that spontaneous perspective-taking can be interrupted under heavy cognitive load, casting further doubts on the automaticity of Theory of Mind processing (2012b; cf. Apperly & Butterfill, 2009; see German & Cohen, 2012 for an alternative view on this debate).

More relevant for the present research is an eye-tracking study by Ferguson, Scheepers, and Sanford (2010), in which adult participants were able to anticipate the outcome of a false-belief narrative as accurately as that of a true-belief narrative. A further question, which was addressed in the present study, is whether adult participants would perform comparably in a direct false-belief test in which they have to answer the standard question 'Where will Sally look for her doll?' Like our developmental study, the present study focused on (a) the role of the false-belief question in direct tests and (b) the susceptibility of perspective-tracking to disruption by task manipulations.

2. The role of the false-belief question in direct tests

The argument that direct false-belief tests pose greater cognitive demands than indirect tests has been a recurrent theme in the developmental literature (e.g., Clements & Perner, 1994; Scott et al., 2012). However, the underlying assumption in these studies has been that, because of the inherent difficulty in producing an overt response to a direct question, passing a direct false-belief test requires more sophisticated false-belief reasoning than passing implicit versions of the task, with the more complex Theory of Mind or general cognitive abilities not emerging before age 4 years (see Apperly & Butterfill, 2009; Baillargeon et al., 2010; de Bruin & Newen, 2012; Perner & Roessler, 2012).

Because of their focus on development, current Theory of Mind accounts do not make predictions as to whether adults would perform differently in indirect and direct false-belief tests. Perner and Roessler, for example, argue that "to get [the false-belief question] right, children need intentionally to switch to [the protagonist]'s perspective to ascertain his subjective reasons. This they cannot do until about 4 years of age" (2012: 524). Given that adults are well over 4 years, no interesting prediction follows from this account about the adult Theory of Mind.

Our analysis of the false-belief question in direct tests is different from Perner and Roessler's and all other accounts of Theory of Mind development. In our view, infants and children are in principle tracking the protagonist's perspective during indirect and direct false-belief tests alike. Assuming that perspective tracking has not been interrupted by other task manipulations (e.g., the introduction of a second character in the story, who may hold the floor for a considerable time; e.g., Maxi's mother in Wimmer & Perner, 1983), the false-belief question in direct tests does not necessarily require an intentional switch to the protagonist's perspective. On the contrary, what the false-belief question does is *interfere* with the children's tracking of the protagonist's perspective.

In our analysis, there are two reasons why the standard false-belief question 'Where will Sally look for her doll?' may interfere with perspective tracking. The first reason is the pragmatics of wh-questions: in the context of a false-belief narrative, answering the test question requires choosing between two possible locations, thus making the participant consider alternative responses. Moreover, by mentioning the target object in the question, children's attention is directed towards the wrong answer (i.e., the actual location of the object). This explains why, in our developmental study, 3 year-old children succeeded in a direct false-belief test in which they were encouraged to produce an overt response by use of open prompts (i.e., 'What happens next? What is the girl going to do now?'; Rubio-Fernández, submitted for publication; Rubio-Fernández & Geurts, 2013).

Because our claims are based on a pragmatic analysis of wh-questions, we argue that children and adults face the same challenge in a direct false-belief test. That is, in order to answer the false-belief question, both children and adults must make an overt choice between two locations, which may disrupt the process of perspective tracking to varying degrees. Thus, whereas young children normally give the wrong response and fail to recover from the disruption of their perspective tracking (probably because the pull of the real is too strong after the question has focused their attention on the object; see Rubio-Fernández, submitted for publication), adults might momentarily consider the two possible responses to the question before giving the right answer. However, relative to an indirect version of the task in which participants are simply told the whole false-belief story, processing the wh-question in a direct test might interfere with the ability of adult participants to anticipate the outcome (see Ferguson et al., 2010).

Previous studies have not directly compared adults' performance on indirect and direct false-belief tests. However, given that adults have greater memory capacity and executive control than young children, they may be able to compensate for the extra cognitive demands of having to answer the false-belief question in a direct test, performing comparably to an indirect test. Nonetheless, if eye movements reveal that adults can more easily anticipate the protagonist's actions in an indirect false-belief test than in a direct one, these results would offer further support to our view that the false-belief question can disrupt the online process of perspective tracking.

2.1. Experiment 1

2.1.1. Method

Fifty-two undergraduates from Princeton University, all native speakers of English, participated for monetary compensation. Twenty-six participants performed an indirect false-belief test and the other 26 performed a direct false-belief test.[1] Three participants in the indirect test were eliminated because of calibration problems.

The false-belief task used in the study was a computer version of the classic Sally–Anne task (Baron-Cohen et al., 1985). Participants were presented with a short cartoon in which two kindergarten characters, Sally and Anne, interacted (see Rubio-Fernández & Glucksberg, 2012). Participants were familiarized with the setting of the story in two warm-up trials before they were presented with one of the two false-belief tests. The warm-up trials introduced the two characters, their toys and their containers, and were the same in both versions of the test, other than for two filler questions about each girl's container in the direct test.[2] In both versions of the test, Sally is away when Anne moves Sally's doll from one container to the other and leaves the scene. At that point participants were presented with one of the following continuations:

Indirect false-belief test

> When Sally comes back [Sally re-appears] the next day, she goes to look for her doll in the basket.

Direct false-belief test

> When Sally comes back [Sally re-appears] the next day, where will she look for her doll?

Eye movements were monitored in both conditions, and response times were recorded in the direct test (although RTs will not be reported until Experiment 2).

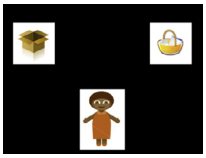
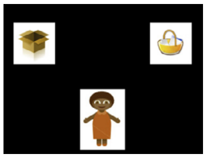
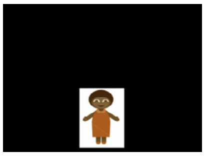
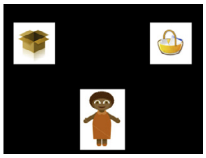
Participants were given written instructions that described their role as a control group in a developmental study investigating what children pay attention to during a story. In the direct test, the response keys were parallel to the two containers on the screen and participants had to press them with their dominant hand.

Eye movements were recorded with an infrared eye-tracking system (504 Pan/Tilt; Applied Science Laboratories Inc., Bedford, MA, USA) that measured eye position at a rate of 60 Hz. The eye-tracking system had a resolution of 0.14° and could detect differences in relative eye posi-

[1] The two tests were run between participants because a preliminary study revealed that adult participants were able to anticipate the false-belief question by the second trial (Rubio-Fernández, 2008). In order to avoid rapid learning effects that might mask differential performances in the two tests, adult participants in the present study were therefore tested on a single false-belief trial (as young children normally are).

[2] The location of the target object and the containers was fixed in this study. However, these locations were changed in another study that revealed the same pattern of results (Rubio-Fernández, Butterfill, & Richardson, 2011).

Table 1
Cartoon slides corresponding with the critical segments in Experiments 1 and 2.

Experiment 1			
Experiment 2			
Indirect test	*When Sally comes back* [1309 ms]	*the next day, she goes to* [2397 ms]	*look for her doll in her basket* [1904 ms]
Direct test	*When Sally comes back* [1105 ms]	*the next day, where will she* [1309 ms]	*look for her doll?* [1020 ms]

tion of ~0.25°. Participants were seated in a comfortable chair and their heads were secured in a chin rest for the duration of the experiment (approx. 4 min).

2.1.2. Results

EyeAnal software (Applied Science Laboratories Inc.) was used to analyze the eye-movement data. Data were first processed to automatically detect and remove eye-blinks. The duration of the critical segment in the narrative was delimited with the use of event markers in the script. Two reference points in the narrative were used to further divide the critical segment: the offset of the verb 'comes back' (corresponding with the point when Sally re-appears in the lower center of the screen) and the onset of the verb 'look for' in the main clause (see Table 1).

Relative to these two points in the narrative, our account makes two predictions: first, participants in both the indirect and direct test will show a reliable preference for the empty location before Sally re-appears, revealing that they are anticipating the correct outcome in both tests. However, while participants in the indirect test will continue to track Sally's perspective uninterruptedly, processing the wh-question in the direct test may result in participants momentarily considering the other possible response to the question (i.e., the actual location of the doll). This hesitation can obviously be overcome by employing extra cognitive resources in the processing of the wh-question (e.g., selective attention). However, assuming the processing the wh-question has an effect on adults' performance, we predict that this effect will be one of momentary hesitation, with a lesser proportion of fixations on the empty location and a greater proportion of fixations on the doll's location being observed in the direct test than in the indirect tests.

The proportion of fixations on each container was established for each 60 Hz (i.e., 17 ms) sample from the tracker for the duration of the critical segment. In order to test the first prediction, statistical analyses focused on the 408 ms prior to Sally re-appearing in the scene (approximately corresponding with the phrase 'comes back' in both recordings). For the second prediction, analyses focused on the 1020 ms following the onset of the main verb (approximately corresponding with the phrase 'look for her doll' in both recordings). For more accurate statistical analyses, the early time-window was further divided into two consecutive 204 ms time-bins and the late time-window into five consecutive 204 ms time-bins.

Figs. 1 and 2 plot the proportion of fixations on each container during the critical segment in the indirect and direct tests.[3]

In the early time-window, two-choice binomial tests (two tailed) revealed that the proportion of fixations on the empty location relative to the total number of fixations on both locations was reliably above chance in the indirect test ($p < .001$ in both time-bins) and the direct test ($p < .007$ in both time-bins). The same comparison in the late time-window revealed a continuous reliable preference for the empty location in the indirect test ($p < .001$ in all five time-bins), whereas in the direct test, the preference for the empty location was significant only in the first, second and fifth time-bins ($p < .04$ in the three time-bins). Full statistical details are reported in Table 2.

In the late time-window, paired comparisons between the indirect and direct tests revealed a reliably larger proportion of fixations on the empty location in the indirect than in the direct test ($p < .001$ in all five t-tests) and the reverse pattern for the proportion of fixations on the dolĺs location ($p < .001$ in all five t-tests). Full statistical details are reported in Table 3.

The results of the first experiment suggest that processing the false-belief question in a direct test can momentarily disrupt adult participants' ability to anticipate the protagonist's actions, as shown in the indirect false-belief test. The results of Experiment 1 therefore offer support to the tracking hypothesis.

2.1.2.1. Disrupting perspective-tracking in adults

Previous studies with adults have shown that false-belief reasoning can be disrupted when performing a secondary task that is cognitively demanding (e.g., Dungan & Saxe, 2012; Newton & de Villiers, 2007; Schneider, Lam,

[3] Given that planning and launching a saccade takes approximately 200 ms, the earliest eye-movements shown in the graphs were planned prior to the start of the critical segment.

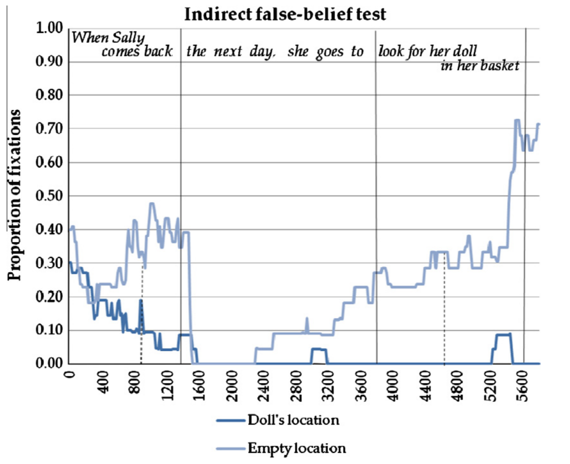


Fig. 1. Proportion of fixations on each container during the critical segment in the indirect test. Continuous vertical lines mark the reference points in the narrative (with the first line marking Sally's re-appearing in the lower center of the scene) and dashed vertical lines mark the analysis regions.

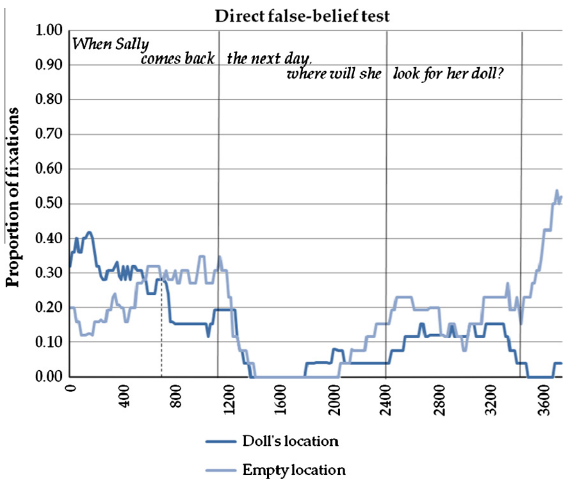


Fig. 2. Proportion of fixations on each container during the critical segment in the direct test.

Bayliss, & Dux, 2012b). The purpose of Experiment 2, however, was to try to disrupt adults' perspective-tracking by means of a subtle task-manipulation, parallel to that used in our developmental study (Rubio-Fernández & Geurts, 2013).

A task manipulation that had a disruptive effect on the performance of the 3 year-olds in our study was the sudden disappearance of the protagonist from the scene when the experimenter dropped the puppet in a bag of toys and continued the story. As evidenced by the results of Experiment 1, adults are not affected by the protagonist's disappearance from the scene and are able to anticipate the right outcome when she returns, even prior to the test phrase/question. However, a similar manipulation that might have a disruptive effect in adults is a brief disappearance of the containers, as they visually represent the two perspectives on the object's location: the participant's and the protagonist's (see Altmann & Kamide, 2009). Thus, in Experiment 2 the containers briefly disappeared from the scene just before the mistaken protagonist returned. The only difference with the first experiment was that the containers were missing in two of the critical slides (see Table 1).

2.2. Experiment 2

2.2.1. Method

Forty-nine undergraduates from Princeton University, all native speakers of English, participated for monetary compensation. Twenty-three participants performed an indirect test and 26 a direct test. The eye-tracking data from three of the latter participants were discarded because of calibration problems.

The materials and procedure in the second experiment were the same as in the first one (including the warm-up trials), except that, after Anne had left, the containers momentarily disappeared from the scene. Thus, in both the indirect and the direct test, when the critical segment started, the scene was empty although the narrative continued as in Experiment 1 (see Table 1).

Participants in Experiment 2 were presented with an extra trial in which Sally moved her own doll from one container to the other. This True-Belief condition was used as a control for the accuracy of first fixation measure (i.e., where participants first fixated when the containers reappeared).[4]

2.2.2. Results

If the brief disappearance of the containers made participants momentarily lose track of Sally's perspective, their eye movements should reveal an initial preference for the actual location of the doll (as representing their own perspective of the object's location) when the containers re-appear. This would result in an increase in the proportion of fixations on the doll's location during the late time-window relative to Experiment 1.

Figs. 3 and 4 plot the proportion of fixations on each container during the critical segment in the indirect and the direct test in the disrupted condition.

Two-choice binomial tests (two-tailed) revealed different results from those observed in the late time-window of Experiment 1: the proportion of fixations on the doll's location relative to the total number of fixations on both locations was reliably above chance in the second and third time-bins in the indirect test ($p < .006$ in both tests) and in the direct test ($p < .003$ in both tests; for full statistical details, see Table 2).

Paired comparisons between the indirect and direct tests revealed a reliably larger proportion of fixations on the empty location in the indirect than in the direct test

[4] The true-belief trial was always presented after the false-belief trial. However, a previous study showed that order of presentation did not affect performance in the true-belief trial (Rubio-Fernández, 2008).

Table 2
Proportion of fixations on the preferred location (empty vs. doll's) relative to the total number of fixations in each time-bin and reliability of the preference relative to chance level (p values from two-tailed two-choice binomial tests in parentheses). Missing data are indicated by a hyphen.

Experiment/Test	Preferred location	Early window		Late window				
		1st	2nd	1st	2nd	3rd	4th	5th
1/Indirect	Empty	.83 (.000)	.89 (.000)	1.00 (.000)	1.00 (.000)	1.00 (.000)	1.00 (.000)	1.00 (.000)
1/Direct	Empty	.62 (.006)	.66 (.000)	.70 (.000)	.61 (.033)	.54 (.630)	.59 (.082)	.67 (.001)
2/Indirect	Empty	–	–	–	–	–	.60 (.020)	.73 (.000)
	Doll's	–	–	–	.72 (.005)	.64 (.001)	–	–
2/Direct	Empty	–	–	–	–	–	–	.65 (.001)
	Doll's	–	–	–	.83 (.002)	.70 (.000)	.58 (.088)	–

Table 3
Mean proportion of fixations on the two locations (empty vs. doll's) in the late time-window of Experiments 1 and 2 (standard deviation in parentheses) and paired t-tests for each time-bin and location.

Experiment/time window	Time bin	Empty location				Doll's location			
		Indirect test	Direct test	t	p	Indirect test	Direct test	t	p
1/Late	1st	.24 (.047)	.21 (.031)	5.180	.000	.00 (.000)	.08 (.028)	10.457	.000
	2nd	.24 (.020)	.20 (.004)	7.252	.000	.00 (.000)	.12 (.014)	37.081	.000
	3rd	.23 (.003)	.12 (.026)	15.326	.000	.00 (.000)	.12 (.011)	40.000	.000
	4th	.27 (.036)	.18 (.044)	11.839	.000	.00 (.000)	.14 (.020)	30.834	.000
	5th	.32 (.022)	.22 (.024)	11.726	.000	.00 (.000)	.11 (.046)	8.719	.000
2/Late	1st	.00 (.000)	.00 (.000)	–	–	.00 (.000)	.00 (.000)	–	–
	2nd	.05 (.066)	.02 (.035)	2.327	.040	.13 (.126)	.08 (.099)	3.307	.007
	3rd	.20 (.065)	.13 (.036)	6.476	.000	.37 (.057)	.30 (.042)	4.067	.002
	4th	.37 (.036)	.22 (.052)	6.078	.000	.26 (.043)	.30 (.030)	5.046	.000
	5th	.43 (.040)	.37 (.082)	4.524	.001	.17 (.032)	.20 (.066)	2.504	.029

($p < .05$ in all five t-tests), replicating the pattern of results observed in the late time-window of Experiment 1. In contrast, the proportion of fixations on the doll's location was larger in the indirect than in the direct test in the second and third time-bins of the late time-window ($p < .008$ in both t-tests), while the reverse pattern was found in the fourth and fifth time-bins ($p < .03$ in both t-tests; for full statistical details, see Table 3).

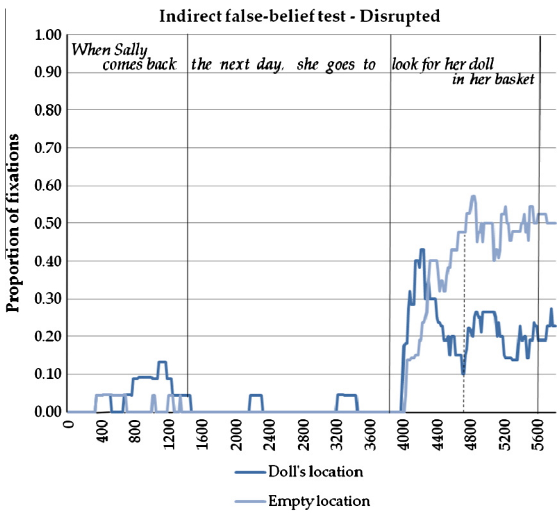


Fig. 3. Proportion of fixations on each container during the critical segment in the indirect test – disrupted condition. The first vertical line marks Sally's re-appearing in the lower center of the scene and the second one the containers' re-appearing in their usual place.

Paired comparisons across the two experiments revealed a reliably larger proportion of fixations on the doll's location in Experiment 2 than in Experiment 1 in the late time-window of the indirect test ($p < .006$ in the last four time-bins) and the direct test ($p < .001$ in the last three time-bins). The full statistical details are reported in Table 4.

2.2.2.1. Accuracy of first fixation

When the containers re-appeared in the scene in the indirect test, 11 participants first fixated on the correct

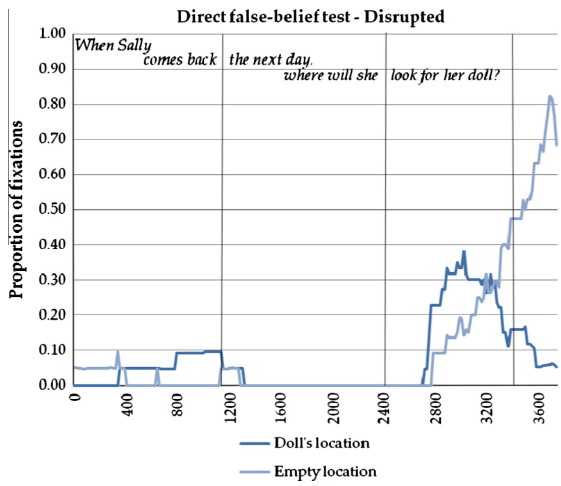


Fig. 4. Proportion of fixations on each container during the critical segment in the direct test – disrupted condition.

Table 4
Mean proportion of fixations on the doll's location in the late time-window of the indirect and direct tests across Experiments 1 and 2 (standard deviation in parentheses) and paired *t*-tests for each time-bin and test.

Late time window	Doll's location							
	Indirect test				Direct test			
Time bin	Exp 1	Exp 2	*t*	*p*	Exp 1	Exp 2	*t*	*p*
1st	.00 (.000)	.00 (.000)	–	–	.08 (.028)	.00 (.000)	10.457	.000
2nd	.00 (.000)	.13 (.126)	3.446	.005	.12 (.014)	.08 (.099)	1.562	.147
3rd	.00 (.000)	.37 (.057)	22.894	.000	.12 (.011)	.30 (.042)	15.304	.000
4th	.00 (.000)	.26 (.043)	20.634	.000	.14 (.020)	.30 (.030)	14.281	.000
5th	.00 (.000)	.17 (.032)	18.091	.000	.11 (.046)	.20 (.066)	7.768	.000

container and 11 on the wrong container (or the doll's location) in the false-belief condition, whereas 19 first fixated on the correct container and 4 on the wrong container in the true-belief control. A chi-square test with Yate's correction revealed a reliable difference between the two conditions, $\chi^2(1, N = 45) = 4.013$, $p = .045$.

In the direct test, 12 participants first fixated on the correct container and 10 on the wrong container in the false-belief condition, whereas 17 first fixated on the correct container and 2 on the wrong container in the true-belief control, also resulting in a reliable difference, $\chi^2(1, N = 41) = 4.439$, $p = .035$ (Yate's correction applied).[5]

2.2.2.2. Response times

All participants responded to the tests question with 100% accuracy. Response times were measured from the offset of the false-belief question. The mean RT was 661 ms (SD 339) in Experiment 1 and 1011 ms (SD 425) in Experiment 2. This difference was reliable, $t(50) = 3.276$, $p = .002$.

The results of Experiment 2 offer support to the tracking hypothesis, as evidenced by the disruptive effect of a subtle visual manipulation on adults' perspective-tracking, both in an indirect and a direct false-belief test.

3. General discussion

The aim of the present study was to test 'the tracking hypothesis': namely the view that perspective tracking is a continuous process that requires focusing on another's perspective and is therefore sensitive to disruption by task manipulations. In Experiment 1, two groups of adults performed an indirect and a direct false-belief test respectively. Early eye-movements revealed that both groups of participants correctly anticipated that Sally would go back to the empty container to look for her doll. However, the processing of the false-belief question ('Where will she look for her doll?') made participants in the direct test momentarily consider the other possible response, resulting in a greater proportion of fixations on the doll's location than in the indirect test. These results offer support to our claim that the false-belief question used in direct tests disrupts perspective tracking by making participants consider alternative responses (as well as focusing children's attention on the object unnecessarily), in contrast to indirect tests in which participants can simply anticipate the outcome (see also Ferguson et al., 2010).

The results of Experiment 2 also offer support to the tracking hypothesis. In the disrupted condition, participants showed an initial preference for the doll's location when the containers re-appeared in the scene – an inaccurate first response that was not observed in the true-belief condition. Past this fist response, participants continued to show a greater tendency to fixate on the doll's location in the disrupted condition, both in the indirect and the direct test. Moreover, the visual disruption delayed participants' response times in the direct test by 350 ms, on average. The disruptive effect of the containers' brief disappearance from the scene suggests, in support of Altmann and Kamide's view (2009), that the two containers in a standard false-belief task visually represent the two perspectives on the object's location (i.e., the participant's and the protagonist's), with the two representations competing for attention during false-belief reasoning (see also Hoover & Richardson, 2008).

Overall, the results of our study with adults confirm the importance of preserving the continuity of the perspective-tracking process in direct false-belief tests with young children, who may not have the necessary cognitive resources to recover from the disruption of their perspective tracking by task manipulations.

A further question for future research concerns the processing of the prompts that have been used in some indirect false-belief tests to elicit anticipatory looking in young children (e.g., Clements & Perner, 1994; see He et al., 2012 for a review). These prompts normally include indirect or subordinate wh-questions (e.g., 'I wonder where Sally will look for her doll'). Importantly, for young children to pass these indirect tests and look in anticipation to the correct location, they must interpret these indirect questions as genuine self-addressed utterances rather than as direct questions (see He et al., 2012). The results of these developmental studies therefore suggest that, as long as indirect false-belief questions are interpreted as such, their processing does not interrupt perspective tracking to the extent that direct questions do. In order to test this hypothesis with adults, however, a more naturalistic paradigm would need to be designed as it is unlikely that adult participants in the Sally-Anne task used in the present study, for example, would take a pre-recorded prompt 'I wonder where Sally will look for her doll' as a genuine self-addressed remark by the adult narrator.

[5] Six participants did not fixate on either container before the end of the critical segment and were therefore excluded from the analyses.

Contrary to the general assumption that the Theory of Mind paradox results from developmental changes between infancy and childhood which affect performance in false-belief tests (cf. Apperly & Butterfill, 2009; Baillargeon et al., 2010; Clements & Perner, 1994; de Bruin & Newen, 2012; Perner & Roessler, 2012), our approach has been to focus on a fundamental ability that infants, children, and adults are likely to have in common: namely the ability to form expectations about other people's actions on the basis of observations of their behavior (what we broadly refer to as 'perspective tracking'; Rubio-Fernández & Geurts, 2013). Our aim in taking this approach was to do an analysis of the cognitive demands of the false-belief tests that have been used with infants and children, which could explain the differential performances of these two groups in a more parsimonious way than by directly assuming a qualitative difference in their Theory of Mind abilities (see also Carruthers, 2013; Jacob, 2012). Even in the light of our positive results, the question remains, however, as to whether the many and obvious differences in the false-belief reasoning abilities of infants, children and adults are specific to Theory of Mind development, or can be accounted for by general development in other key areas of cognition (see, e.g., Low, 2010 for a study on the role of language proficiency in false-belief understanding). Future research will hopefully address this issue.

The fact that disrupting perspective-tracking in a false-belief task might result in an initial preference for the actual location of the target object suggests that the participant's own perspective can sometimes serve as the starting point in adopting the mistaken protagonist's perspective (cf. Kovács et al., 2010). However, if participants are allowed to track the protagonist's perspective uninterruptedly, their own knowledge of the situation need not interfere with the false-belief task.

Acknowledgements

This research was supported by an X-Prag Travel Grant, *Do you see what I'm thinking about? How adults' eye-movements are influenced by representations of mental states* (European Science Foundation). The author would like to thank Bart Geurts, Chris Cummins and two anonymous reviewers for their insightful comments and suggestions on earlier versions of the paper.

References

Altmann, G. T. M., & Kamide, Y. (2009). Discourse-mediation of the mapping between language and the visual world: Eye movements and mental representation. *Cognition, 111*, 55–71.
Apperly, I., Back, E., Samson, D., & France, L. (2008). The cost of thinking about false beliefs: Evidence from adults' performance on a noninferential theory of mind task. *Cognition, 106*, 1093–1108.
Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review, 116*, 953–970.
Apperly, I. A., Riggs, K. J., Simpson, A., Samson, D., & Chiavarino, C. (2006). Is belief reasoning automatic? *Psychological Science, 17*, 841–844.
Back, E., & Apperly, I. A. (2010). Two sources of evidence on the non-automaticity of true and false belief ascription. *Cognition, 115*, 54–70.
Baillargeon, R., Scott, R. M., & He, Z. (2010). False-belief understanding in infants. *Trends in Cognitive Sciences, 14*, 110–118.
Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a 'Theory of Mind?'. *Cognition, 21*, 37–46.
Buttelmann, D., Carpenter, M., & Tomasello, M. (2009). Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition, 112*, 337–342.
Carruthers, P. (2013). Mindreading in infancy. *Mind and Language, 28*, 141–172.
Clements, W. A., & Perner, J. (1994). Implicit understanding of belief. *Cognitive Development, 9*, 377–397.
Cohen, A. S., & German, T. C. (2009). Encoding of others' beliefs without overt instruction. *Cognition, 110*, 356–363.
de Bruin, L. C., & Newen, A. (2012). An association account of false-belief understanding. *Cognition, 123*, 240–259.
Dungan, J., & Saxe, R. (2012). Matched false-belief performance during verbal and nonverbal interference. *Cognitive Science, 36*, 1148–1156.
Ferguson, H. J., Scheepers, C., & Sanford, A. J. (2010). Expectations in counterfactual and theory of mind reasoning. *Language and Cognitive Processes, 25*, 297–346.
German, T. C., & Cohen, A. S. (2012). A cue-based approach to 'theory of mind': Re-examining the notion of automaticity. *British Journal of Developmental Psychology (Special Issue), 30*, 45–58.
He, Z., Bolz, M., & Baillargeon, R. (2012). 2.5-year-olds succeed at a verbal anticipatory-looking false-belief task. *British Journal of Developmental Psychology (Special Issue), 30*, 14–29.
Hoover, M. A., & Richardson, D. C. (2008). When facts go down the rabbit hole: Contrasting features and objecthood as indexes to memory. *Cognition, 108*, 533–542.
Jacob, P. (2012). A puzzle about belief-ascription. In B. Kaldis (Ed.), *Mind and society: Cognitive science meets the philosophy of the social sciences. Synthese Philosophy Library*. Berlin: Springer.
Kovács, Á. M., Téglás, E., & Endress, A. D. (2010). The social sense: Susceptibly to others' beliefs in human infants and adults. *Science, 330*, 1830–1834.
Low, J. (2010). Preschoolers' implicit and explicit false-belief understanding: Relations with complex syntactical mastery. *Child Development, 81*, 597–615.
Low, J., & Perner, J. (2012). Implicit and explicit Theory of Mind: State of the art. *British Journal of Developmental Psychology (Editorial to Special Issue), 30*, 1–13.
Newton, A. M., & de Villiers, J. G. (2007). Thinking while talking: Adults fail nonverbal false-belief reasoning. *Psychological Science, 18*, 574–579.
Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science, 308*, 255–258.
Perner, J. (2010). Who took the cog out of cognitive science? Mentalism in an era of anti-cognitivism. In P. A. Frensch & R. Schwarzer (Eds.). *Cognition and neuropsychology: International perspectives on psychological science* (Vol. 1, pp. 241–261). Hove, England: Psychology Press.
Perner, J., & Roessler, J. (2012). From infants' to childrens' appreciation of belief. *Trends in Cognitive Sciences, 16*, 519–525.
Rubio-Fernández, P. (2008). What can Sally and Ann tell us about adult Theory of Mind? In Poster presented at the 20th Annual Convention of the Association for Psychological Science, Chicago.
Rubio-Fernández, P. (submitted for publication). Don't mention the marble! Why the false-belief question is too difficult for young children.
Rubio-Fernández, P., Butterfill, D., & Richardson, D. (2011). Do you see what I'm thinking about? Eye movements in FB reasoning. In Paper presented at the 2nd Experimental Pragmatics Workshop, Pisa.
Rubio-Fernández, P., & Geurts, B. (2013). How to pass the false-belief task before your 4th birthday. *Psychological Science, 24*, 27–33.
Rubio-Fernández, P., & Glucksberg, S. (2012). Reasoning about other people's beliefs: Bilinguals have an advantage. *Journal of Experimental Psychology: Learning, Memory and Cognition, 38*, 211–217.
Samson, D., Apperly, I. A., Braithwaite, J., & Andrews, B. (2010). Seeing it their way: Evidence for rapid and involuntary computation of what other people see. *Journal of Experimental Psychology: Human Perception and Performance, 36*, 1255–1266.
Schneider, D., Bayliss, A. P., Becker, S. I., & Dux, P. E. (2012a). Eye movements reveal sustained implicit processing of others' mental states. *Journal of Experimental Psychology: General, 141*, 433–438.
Schneider, D., Lam, R., Bayliss, A. P., & Dux, P. E. (2012b). Cognitive load disrupts implicit Theory-of-Mind processing. *Psychological Science, 23*, 842–847.
Scott, R. M., He, Z., Baillargeon, R., & Cummins, D. (2012). False-belief understanding in 2.5-year-olds: Evidence from two novel verbal spontaneous-response tasks. *Developmental Science, 15*, 181–193.

Senju, A., Southgate, V., Snape, C., Leonard, M., & Csibra, G. (2011). Do 18-month-olds really attribute mental states to others? A critical test. *Psychological Science, 22*, 878–880.

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

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