Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Cognition xxx (2007) xxx–xxx
COGNITION
www.elsevier.com/locate/COGNIT

The cost of thinking about false beliefs: Evidence from adults' performance on a non-inferential theory of mind task

Ian A. Apperly *, Elisa Back, Dana Samson, Lisa France

School of Psychology, University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK

Received 7 February 2007; accepted 12 May 2007

Abstract

Much of what we know about other people's beliefs comes non-inferentially from what people tell us. Developmental research suggests that 3-year-olds have difficulty processing such information: they suffer interference from their own knowledge of reality when told about someone's false belief (e.g., [Wellman, H. M., & Bartsch, K. (1988). Young children's reasoning about beliefs. *Cognition, 30*, 239–277.]). The current studies examined for the first time whether similar interference occurs in adult participants. In two experiments participants read sentences describing the real colour of an object and a man's false belief about the colour of the object, then judged the accuracy of a picture probe depicting either reality or the man's belief. Processing costs for picture probes depicting reality were consistently greater in this false belief condition than in a matched control condition in which the sentences described the real colour of one object and a man's unrelated belief about the colour of another object. A similar pattern was observed for picture probes depicting the man's belief in most cases. Processing costs were not sensitive to the time available for encoding the information presented in the sentences: costs were observed when participants read the sentences at their own pace (Experiment 1) or at a faster or a slower pace (Experiment 2). This suggests that adults' difficulty was not with encoding information about reality and a conflicting false belief, but with holding this information in mind and using it to inform a subsequent judgement.

* Corresponding author. Tel.: +44 121 414 3339.
E-mail address: i.a.apperly@bham.ac.uk (I.A. Apperly).

0010-0277/$ - see front matter © 2007 Published by Elsevier B.V.
doi:10.1016/j.cognition.2007.05.005

Keywords: False belief; Theory of mind; Adult

1. Introduction

Theory of mind (ToM) describes a set of abilities used to explain or predict behaviour in terms of mental states such as beliefs, desires and intentions. False belief tasks are a very widely used test of these abilities (Wellman, Cross, & Watson, 2001; Wimmer & Perner, 1983). In one typical false belief task a story character, Sally, places her marble in a basket, then goes outside to play. In her absence, a second character, Anne, moves the marble from the basket to a box, with the result that Sally has a false belief about the marble's location. Participants are then asked test questions that require them to infer Sally's false belief in order to say where Sally thinks the marble is located, or to predict where Sally will first look to find her marble (Baron-Cohen, Leslie, & Frith, 1985). Three-year-old children commonly fail such tasks by judging from their own knowledge of the object's location, rather than Sally's false belief. Four-year-olds typically answer correctly, but the cognitive processes responsible for such success in children or in adults are not well-understood. Only quite recently have researchers begun to investigate the cognitive basis of ToM in typical adults, and surprising gaps remain in our understanding. For example, current studies of adults systematically confound the process of inferring a mental state with any processes involved in simply representing this information. This is an important confound for two reasons. First, much of what we know about other people's mental states comes non-inferentially when they, or some other person, simply tell us what they know, think, want or intend. Thus, it is important to know how such information is encoded and held in mind. Second, in developmental research it is known that ToM inferences are not children's only problem: even when simply told about someone's false belief, three-year-olds make errors when asked to judge what the person will think or do, apparently suffering severe interference from their own knowledge of reality (de Villiers & Pyers, 2002; Flavell, Flavell, Green, & Moses, 1990; Wellman & Bartsch, 1988). The current study is the first investigation of adults' mental representation of ToM information that de-confounds the need to hold ToM information in mind from the need to make a ToM inference.

1.1. Existing studies of ToM in adults

Behavioural studies of ToM in adults have tested participants' ability to make inferences about mental states. Adults often make errors when asked to evaluate statements that require inferences about mental states with multiple embeddings (e.g., ''Bob thinks that John knew that Mary wanted to go to the shop''; Kinderman, Dunbar, & Bentall, 1998; Rutherford, 2004). Adults typically make few errors when they only have to infer one person's belief, or one person's belief about another's belief (e.g., Fletcher et al., 1995; Stone, Baron-Cohen, & Knight, 1998). However, the likelihood of error is increased if adults perform a concurrent task designed to tax working memory or other components of executive function (e.g., Bull, Phillips, & Conway, submitted for publication; McKinnon & Moscovitch, 2007). Relatedly, German and Hehman (2006) suggest that adults are slower and more error-prone when inferring combinations of mental states that place relatively high demands on inhibitory control (e.g., false belief plus negative desire) compared with combinations that make lower demands (e.g., true belief plus positive desire). Interestingly, even in the case of very simple false beliefs where adults would be unlikely to infer the belief incorrectly, adults may nonetheless show biases in the probabilities they attach to the likely behaviour of the person with the false belief (Birch & Bloom, 2007). Overcoming such a "curse of knowledge" may require executive control.

The above findings indicate a significant role for memory and executive resources in adults' performance of ToM tasks. However, the conclusions warranted by these findings are limited for two important reasons. First, it is widely agreed that ToM tasks make significant demands upon executive processes that have nothing to do with ToM per se (e.g., Apperly, Samson, & Humphreys, 2005; Bloom & German, 2000). Increasing task complexity, or placing adults under cognitive load, may lead to errors on ToM tasks because adults struggle to meet these incidental demands. Firm conclusions about a necessary role for executive processes are only warranted if these incidental demands are adequately controlled for, and the kind of check questions most commonly employed in ToM tasks fall short of achieving this (Apperly et al., 2005). Second, even when tasks do include well-matched comparison conditions or control trials (e.g., German & Hehman, 2006) it is unclear whether executive resources are required for inferring mental states, for holding this information in mind during the task, for formulating an answer to the test questions using the relevant ToM information, or for all of these processes. Separating these ToM processes in tasks with appropriate control trials is essential if the role of memory and executive processes in ToM is to be understood. The current paper speaks to this question by investigating how adults hold ToM information in mind.

1.2. Non-inferential ToM processing

Young children have difficulty processing ToM information even when they do not have to make a ToM inference. Wellman and Bartsch (1988) simply told children about a story character's false belief and the corresponding reality (e.g., "Sam thinks the puppy is in the garage/The puppy is really on the porch"). When asked where Sam would look for his puppy, most 3-year-olds and young 4-year-olds judged that he would look on the porch, making the same "reality" error observed in more standard tasks where it is necessary to infer a false belief (e.g., Wellman et al., 2001; see also Flavell et al., 1990). In a task designed to assess children's processing of embedded complement syntax de Villiers and Pyers (2002) told children short stories then summarized the false belief of a story character e.g., "He thought he found a ring, but really, it was a bottle cap". The experimenter then pointed to a picture of the character and asked "What did he think?" Many 3-year-olds judged incorrectly that

Please cite this article in press as: Apperly, I. A. et al., The cost of thinking about false beliefs: ..., *Cognition* (2007), doi:10.1016/j.cognition.2007.05.005

he thought he found a bottle cap. Clearly, children have difficulty processing ToM information even when no inference is necessary. Although children's non-inferential ToM reasoning has received relatively little theoretical attention, the literature on children's general ToM development offers a number of potential interpretations for children's difficulties, and these make different predictions about the pattern that might be observed in adults.

It is commonly suggested that children fail false belief tasks because they lack a concept of belief (e.g., Gopnik & Wellman, 1992; Perner, 1991; Wellman et al., 2001), and without the necessary concept it would be natural for children to have difficulty with non-inferential processing of information about beliefs. de Villiers and Pyers (2002) suggest that younger children make errors because they incorrectly process the embedded complement clauses of belief statements such as "He thought he found a ring" (though see e.g., Lillard & Flavell, 1992; Perner, Sprung, Zauner, & Haider, 2003; Smith, Apperly, & White, 2003, for data that seem inconsistent with this view). Importantly, neither the conceptual nor the syntactic explanation for children's errors would predict that adults would have any difficulty processing reports of false beliefs, since adults have a mature concept of belief and mature processing of the syntax of belief statements.

Many authors have proposed that children fail false belief tasks because they lack the necessary executive control (e.g., Carlson & Moses, 2001; Leslie, Friedman, & German, 2004; Mitchell, 1996; Russell, 1996; Zelazo, Muller, Frye, & Marcovitch, 2003). One suggestion is that executive control is necessary for the emergence of ToM concepts such as belief, perhaps by enabling children to disengage from the immediate objects of their attention (e.g., Carlson & Moses, 2001; Russell, 1996). If a concept of belief has not yet emerged in children's thinking then it is unsurprising that they have difficulty on non-inferential false belief tasks. However, emergence accounts could not explain any difficulty that adults might have on non-inferential false belief tasks since the requisite ToM concepts would already have emerged.

Another suggestion is that executive control is necessary for the expression of a belief concept, perhaps by enabling children to overcome default ascription of true beliefs (e.g., Leslie et al., 2004) or to resist any tendency to respond on the basis of their own knowledge rather than what the other person believes (e.g., Carlson & Moses, 2001; Russell, 1996). This role might exist whether the concept of belief is innate or emergent. It has also been suggested that any reasoning about false beliefs necessarily requires a certain level of executive control, without which children lack a proper concept of beliefs (e.g., Russell, 1996). Either of these accounts could explain young children's difficulties on non-inferential tasks, and on either account it is possible that ToM in adults also makes demands on executive control.

In sum, accounts of ToM development differ in their ability to explain any processing costs observed in adults' non-inferential ToM processing. Thus, as well as being informative about how adults process ToM information, data from adults can also play a valuable role in constraining interpretation of the development of ToM.

Please cite this article in press as: Apperly, I. A. et al., The cost of thinking about false beliefs: ..., *Cognition* (2007), doi:10.1016/j.cognition.2007.05.005

1.3. The current studies

The developmental literature suggests that children find it hard to resist interference from knowledge of reality when they are told about a false belief and must simply hold this information briefly in mind. The current studies tested whether adults would show analogous difficulties on a non-inferential false belief task. To test this we presented adults with information about a situation, e.g., "Really, the ball on the table is yellow", and information about someone's false belief about the situation, e.g., "He thinks that the ball on the table is red". Note that these sentences are in no way contradictory – it can be simultaneously true that the ball on the table is yellow and that someone thinks (falsely) that it is red. However, as with any false belief, the propositional content of the man's belief is clearly in conflict with reality, raising the possibility of processing costs for such information. Subsequently, we tested adults' ability to formulate judgements about this information by asking them to judge the accuracy of a picture probe that either depicted reality, or the man's false belief. We used pictures rather than sentences as probes to ensure that participants could not judge the probes on the basis of superficial similarity with the initial sentences in which information about belief and reality was presented.

Whereas the developmental literature suggests that young children may actually be unable to process such information correctly, adults would generally be expected to succeed. Thus, any processing costs for adults are likely to be evident in the time it takes to formulate a judgement, with only occasional errors. Importantly, a number of factors that are irrelevant to processing false beliefs per se will also contribute to these processing costs. Besides any problem with holding in mind someone's false belief, the participant must remember two sets of information about objects, locations and colours, and assign one set to the man and one set to reality (Object on the table is yellow/He thinks object on the table is red). Thus, evaluating the specific processing costs of our False Belief/Reality (False B/R) condition requires comparison against a baseline condition that also poses these incidental processing demands but in which there is no conflict between belief and reality.

In the developmental literature it is often noted that children pass true belief tasks before false belief tasks (e.g., Leslie et al., 2004; Wellman et al., 2001), and a true belief condition might, at first sight, appear a suitable baseline in the current study (i.e., "Really the ball on the table is yellow/He thinks the ball on the table is yellow"). However, a true belief condition presents the participant with a simple strategy for reducing processing costs by reducing the information they must hold in mind: all they need remember is a single set of information about object location and colour, and a single fact about the man (i.e., "Ball on the table is yellow/Man is right"). Therefore, a true belief condition might be easier to process than a False B/R condition merely because participants had to remember less information, not because there was no interference between belief and reality information. Thus, although we included true belief trials as filler items (see below) we did not think them a suitable experimental baseline.

Instead of a true belief baseline condition, we presented participants with information about reality and an unrelated belief, e.g., "Really the ball on the table is yellow''/''He thinks the ball on the chair is red''. Like the False B/R condition, the Unrelated B/R condition did not allow participants an obvious shortcut to remembering the necessary information: participants had to remember two sets of information about objects, locations and colours, and assign one set to the man and one set to reality (Ball on the table is yellow/He thinks the ball on the chair is red). Indeed, since there are different locations in the Unrelated B/R condition (table and chair) there are more distinct facts to remember than in the False B/R condition. Critically however, whereas the content of a false belief is in conflict with reality (this, recall, is often thought to be the critical source of difficulty for children), in the Unrelated B/R condition there is no conflict between reality and the content of the man's belief.

Different accounts of how adults might mentally represent the information that they have been given about belief and reality yield different predictions about the relative processing costs of False B/R and Unrelated B/R conditions. One possibility is that adults do not integrate the information presented in the two sentences, with the result that the ''falseness'' (and thus conflict between belief and reality) of the False B/R condition is not represented. If this were the case then we would expect the processing costs to be determined solely by the number of distinct pieces of information that participants had to hold in mind and, if anything, we should observe higher costs for the Unrelated B/R condition (where the objects are described in different locations) than for the False B/R condition (where there is only one location). On the other hand, if adults do integrate the information presented in the two sentences into a coherent representation then the ''falseness'' (and thus conflict between belief and reality) in the False B/R condition would be represented, and we might expect higher processing costs in the False B/R condition than in the Unrelated B/R condition. Moreover, if processing costs in adults correspond directly to the ''reality bias'' pattern of difficulty observed in young children, or the ''curse of knowledge'' observed in adults (e.g., Birch & Bloom, 2007) then in the False B/R condition we would expect the processing cost to be highest for judgements about beliefs, and lower for judgements about reality.

2. Experiment 1: Self-paced reading

2.1. Method

2.1.1. Participants

Sixteen undergraduate students (13 female, all right handed) participated for course credits. Participants ranged between 18 and 28 years (mean = 20 years, $SD = 2.28$ years).

2.1.2. Design and procedure

Each trial consisted of three events: Sentence 1, Sentence 2 and a picture probe (see Fig. 1). N.B., picture probes appeared in colour, so the colour of the object in the box could be directly observed. Participants judged whether the picture probe

Please cite this article in press as: Apperly, I. A. et al., The cost of thinking about false beliefs: ..., *Cognition* (2007), doi:10.1016/j.cognition.2007.05.005

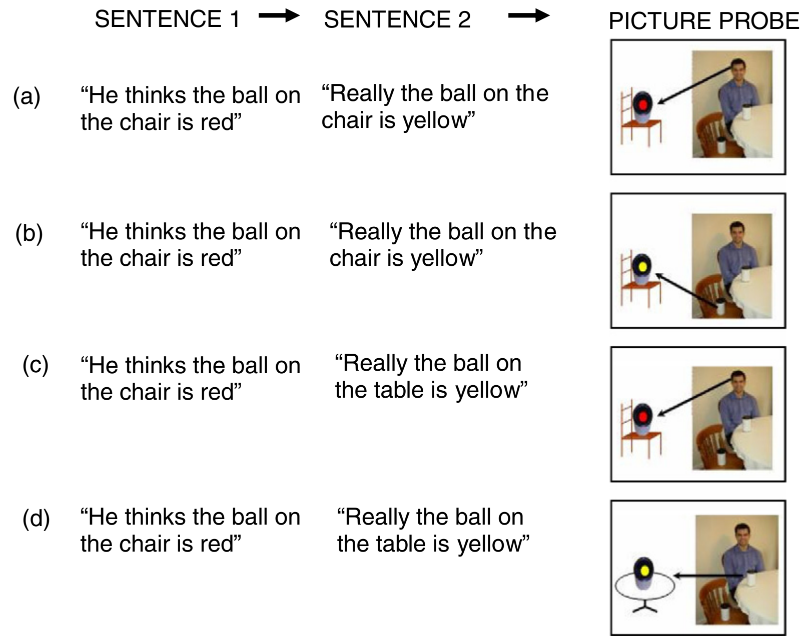

Fig. 1. Schematic diagrams of the event sequence for each experimental trial (N.B., actual probes were in colour, so the colour of the object in the box could be directly observed). Sentence 1 and Sentence 2 from a False Belief/Reality trial are followed by (a) a Belief probe (red object in box), or (b) a Reality probe (yellow object in box). Sentence 1 and Sentence 2 from an Unrelated Belief/Reality trial are followed by (c) a Belief probe (red object in box), or (d) a Reality probe (yellow object in box). Each probe in these examples has the correct answer "yes". (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

accurately represented the situation described in the sentences. The critical trial types differed only by a single word in one of the two sentences.

On 48 *False Belief/Reality* (False B/R) trials one of the two sentences described the real colour and location (table/chair) of an object (e.g., "Really the ball on the chair is yellow") and the other sentence described someone's false belief about the same object (e.g., "He thinks the ball on the chair is red"). Following Sentence 1 and Sentence 2 a picture probe appeared, showing a photograph in which a man sat next to a chair and a table with boxes on them, and an arrow pointed from the photograph to an open box on a cartoon chair or table (see Fig. 1). On half of the trials the probe depicted the man's *belief* by having the arrow point from the man's head to the cartoon box on the cartoon chair or table (see Fig. 1a). On half of the trials the probe depicted *reality* by having the arrow point from the chair or table in the photograph to the cartoon box on the cartoon chair or table (see Fig. 1b). The meaning of each probe type was explained to participants at the beginning of the experiment.

On 48 *Unrelated Belief/Reality* (Unrelated B/R) trials one of the two sentences described the real colour and location of an object (e.g., "Really the ball on the table

Please cite this article in press as: Apperly, I. A. et al., The cost of thinking about false beliefs: ..., *Cognition* (2007), doi:10.1016/j.cognition.2007.05.005

is yellow'') and the other sentence described someone's unrelated belief about an object in a second location (e.g., ''He thinks the ball on the chair is red''). Sentences 1 and 2 were followed by exactly the same type of picture probes used for the False Belief/Reality condition (see Fig. 1c and d).

In both trial types Sentences 1 and 2 described belief and reality equally often. Half of the probes required a ''yes'' response because the colour of the ball in the cartoon box was the same colour described in the corresponding belief or reality sentence (e.g., in Fig. 1a, the probe depicts the man's belief and is consistent with the corresponding belief sentence, so requires a ''yes'' response). Half of the probes required a ''no'' response because the colour of the ball in the cartoon box was the colour described in the non-corresponding sentence (if the probe depicted the man's belief then the non-corresponding sentence would be the reality sentence).

Experimental trials were mixed with anti-strategy filler trials designed to ensure that participants had to attend to all of the information in Sentence 1 and Sentence 2 in order to respond accurately. Thirty two filler trials described a true belief (e.g., ''He thinks the ball on the table is blue/Really the ball on the table is blue''), thirty two trials described two aspects of the man's belief (e.g., ''He thinks the ball is on the table/He thinks the ball is blue'') and thirty two trials described two aspects of reality (e.g., ''Really the ball on the table is red/Really the ball on the chair is blue''). Picture probes for these filler trials either required a ''yes'' response because they accurately depicted the sentences or a ''no'' response because the colour of the ball in the picture did not match the colours described in the sentences. A further set of filler trials was paired with probes that required a ''no'' response because the location of the ball in the picture probe did not correspond to the locations described in the sentences. These filler trials were comprised of a further sixteen trials of each kind described above (true belief; two aspects of belief; two aspects of reality), plus forty eight trials based on False B/R or Unrelated B/R sentences. Thus, to respond correctly on all trial types participants needed to attend to both the colour and the location of the balls in the sentences about reality and sentences about the man's belief.

The 96 experimental trials and 192 fillers were distributed in a pseudo-random manner into 4 equivalent blocks, avoiding more than three consecutive presentations of experimental trials or of yes or no responses. Participants completed all 4 blocks in an order that was rotated between participants. Seven further trials (including experimental and filler trials) were presented as a short warm-up block at the beginning of the experiment. Participants paced their own reading of the sentences using the space key, and responded to the picture probe by pressing keys corresponding to ''yes'' and ''no''. The experiment was presented on a standard Pentium-based desktop computer using eprime (http://www.pstnet.com/products/eprime/).

2.2. Results and discussion

Error rates ranged from 4% to 20% incorrect across different conditions, and both errors and response times showed very similar patterns of greater processing costs on False B/R trials than on Unrelated B/R trials (see Table 1). Since our hypothesis concerned overall processing cost rather than speed or accuracy per se, we computed

Table 1
Mean reaction time and proportion correct for experimental trials of Experiments 1 and 2

	Probe type	Sentence 1 information				Sentence 2 information			
		False B/R		Unrelated B/R		False B/R		Unrelated B/R	
		Belief	Reality	Belief	Reality	Belief	Reality	Belief	Reality
E1: Self-paced reading	RT (ms)	1479	1536	1254	1327	1401	1447	1224	1317
	Proportion correct	.876	.798	.896	.876	.896	.927	.917	.964
E2: 1500 ms reading time	RT (ms)	1375	1408	1322	1277	1201	1317	1227	1296
	Proportion correct	.850	.781	.875	.849	.913	.901	.923	.949
E2: 2500 ms reading time	RT (ms)	1207	1244	1208	1174	1118	1255	1199	1179
	Proportion correct	.897	.886	.928	.943	.928	.908	.948	.938

N.B., processing costs were computed individually for each subject, not from overall condition means.

processing costs by dividing participants' mean correct RT by the proportion of correct responses for each condition (see Fig. 2). Because the nature and kind of interference might differ when the probe assessed memory for information presented in Sentence 2 (i.e., the more recently presented information) compared with Sentence 1 (that had to be held in mind while processing Sentence 2), we included information position as a factor.

We computed a three-way ANOVA with trial type (False B/R, Unrelated B/R), probe type (Belief, Reality) and information position (Sentence 1, Sentence 2) as repeated measures and processing cost as the dependent variable. There were significant main effects of trial type, $F(1,15) = 8.91$, $p = .009$, $n_p^2 = .373$, (False B/R > Unrelated B/R), and information position, $F(1,15) = 11.5$, $p = .004$, $n_p^2 = .435$, (Sentence 1 > Sentence 2). There was a significant interaction between information position and probe type, $F(1,15) = 13.1$, $p = .003$, $n_p^2 = .465$. The mean processing costs for Belief and Reality probes were 1585 and 1771 (respectively) for Sentence 1 information, and 1495 and1475 for Sentence 2 information, suggesting that the interaction was due to disproportionately higher processing costs for Reality probes about Sentence 1 information. No other effects were significant (all Fs < .292, all ps > .11).

Mean processing costs (and standard deviations) for each type of filler trial are displayed in Table 2. The function of these fillers was to prevent participants from adopting superficial strategies on the experimental trials, and we did not intend to enter them into statistical analyses with the experimental trials. However, it is noteworthy that the true belief trials (e.g., He thinks that the ball on the table is blue/Really the ball on the table is blue) incurred the lowest processing cost. This is consistent with our expectation that participants could adopt a strategy to solve these trials easily, and supports our selection of an Unrelated B/R condition rather than true belief condition as the appropriate comparison for the False B/R condition.

Please cite this article in press as: Apperly, I. A. et al., The cost of thinking about false beliefs: ..., *Cognition* (2007), doi:10.1016/j.cognition.2007.05.005

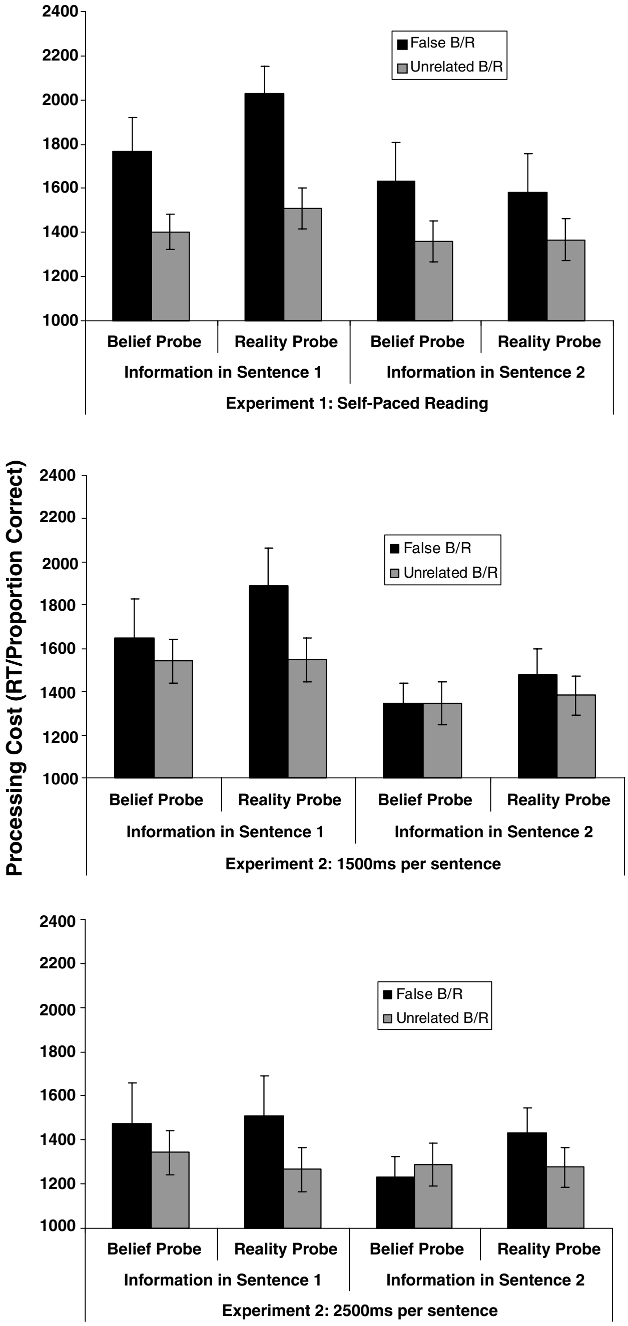

Fig. 2. Mean (and Standard Error) processing cost for each condition of Experiment 1 (top panel) and Experiment 2 (lower two panels).

Table 2
Mean processing costs and standard errors for each trial type of Experiments 1 and 2

	Experimental trials		Filler trials			
	Related	Unrelated	True Belief/ Reality	Reality/ Reality	Belief / Belief	Probe shows wrong location
E1: Self-paced reading mean (*SD*)	1665 (533)	1407 (305)	1000 (187)	1438 (261)	1034 (179)	1331(220)
E2: 1500 mean (*SD*)	1590 (134)	1454 (90)	1045 (185)	1655 (465)	1053 (145)	1467 (367)
E2: 2500 mean (*SD*)	1411 (134)	1293 (90)	994 (437)	1372 (527)	1062 (357)	1254 (444)

Participants showed a greater processing cost on False B/R trials than on Unrelated B/R trials, for both types of probe and regardless of whether the probed information was presented in Sentence 1 or Sentence 2. This effect could arise because participants were still processing the incoming information from Sentences 1 and 2 when they reached the picture probe, so that the greater processing cost for probes in the False B/R condition reflected the greater costs of encoding this information. Another possibility is that interference occurred *after* encoding, while participants held the information in mind. If differences in the time needed for encoding were responsible for the difference in processing cost between False B/R and Unrelated B/R conditions in Experiment 1, we would expect this effect to be modulated by varying the time available for encoding. In Experiment 2, Sentences 1 and 2 were either presented for a period that was less than the mean self-paced reading time from Experiment 1, or substantially greater than the self-paced reading time.

3. Experiment 2: Short versus long encoding time

3.1. Method

3.1.1. Participants

Thirty two undergraduate students (27 female, 29 right handed) participated for course credits. Participants ranged between 18 and 40 years (mean = 21 years, *SD* = 3.9 years).

3.1.2. Design and procedure

The method was the same as for Experiment 1 in all respects except that participants no longer determined for themselves the time available for reading Sentences 1 and 2. Instead, half of the participants saw each sentence for just 1500 ms (>300 ms faster than the 1838 ms average reading time in Experiment 1, and judged "fast" by pilot participants). Other participants saw each sentence for 2500 ms (>600 ms slower than the average reading time in Experiment 1, and judged "comfortable" by pilot participants).

3.2. Results and discussion

Processing costs for responses to the picture probes were calculated by dividing participants' mean RT by their proportion correct in each condition (see Table 1 for mean RT and proportion correct and Fig. 2 for a graph of mean processing cost for each condition). We computed a four-way ANOVA with trial type (False B/R or Unrelated B/R), probe type (Belief, Reality) and information position (Sentence 1, Sentence 2) as within-subject variables and reading time (1500 ms or 2500 ms) as a between-subject variable and processing cost as the dependent variable. There were significant main effects of trial type, $F(1,30) = 9.24$, $p = .005$, $n_p^2 = .235$, (False B/R > Unrelated B/R), probe type, $F(1,30) = 6.80$, $p = .014$, $n_p^2 = .185$, (Reality > Belief) and information position, $F(1,30) = 22.6$, $p < .001$, $n_p^2 = .430$, (Sentence 1 > Sentence 2). Information position interacted with trial type, $F(1,30) = 5.42$, $p = .027$, $n_p^2 = .153$ and reading time, $F(1,30) = 5.34$, $p = .028$, $n_p^2 = .151$. Trial type interacted with probe type, $F(1,30) = 10.0$, $p = .004$, $n_p^2 = .250$. No other effects were significant.

Because information position interacted with both trial type and reading time we conducted two separate three-way ANOVAs for information presented in Sentences 1 and 2, with trial type (False B/R or Unrelated B/R) and probe type (Belief, Reality) as within-subject variables and reading time (1500 ms or 2500 ms) as a between-subject variable.

For information presented in Sentence 1, there was a significant main effect of trial type, $F(1,30) = 9.19$, $p = .005$, $n_p^2 = .235$ (False B/R > Unrelated B/R). Trial type interacted significantly with probe type, $F(1,30) = 4.73$, $p = .038$, $n_p^2 = .136$. Paired t-tests showed a significantly greater processing cost in the False B/R condition than in the Unrelated B/R condition for reality probes, $t(31) = 3.91$, $p < .001$, and a trend for a difference in the same direction for belief probes $t(31) = 1.48$, $p = .149$. There was a non-significant interaction between probe type and reading time, $F(1,30) = 3.12$, $p = .088$, $n_p^2 = .094$, with proportionately more processing cost on reality probes than on belief probes in the 1500 ms reading condition. No other effects approached significance (all Fs < .19 all ps > .18).

For information presented in Sentence 2 there was no significant effect of trial type. However, there was a significant main effect of probe type, $F(1,30) = 6.70$, $p = .015$, $n_p^2 = .183$ (Reality > Belief), and a significant interaction between trial type and probe type, $F(1,30) = 4.59$, $p = .040$, $n_p^2 = .133$. Paired t-tests showed a significantly greater processing cost in the False B/R condition than in the Unrelated B/R condition for reality probes, $t(31) = 2.57$, $p = .015$, but no difference for belief probes ($p = .58$). No other effects in the main analysis were significant (all Fs < .20 all ps > .17).

Mean processing costs (and standard deviations) for each type of filler trial are displayed in Table 2. The patterns were similar to those observed in Experiment 1.

Overall, Experiment 2 replicated the greater processing cost on False B/R trials than on Unrelated B/R trials found in Experiment 1 irrespective of reading time, suggesting that the cost arises during storage and retrieval rather than encoding of the relevant information. The processing costs for the False B/R condition seemed less severe in Experiment 2 than in Experiment 1. It is possible that the need to self-pace in Experiment 1 may have added a processing load that enhanced the difference between the more demanding False B/R condition and the less demanding Unrelated B/R condition.

It is also the case that the detailed pattern was somewhat different in Experiment 2. Belief probes only showed a trend for greater processing costs in the False B/R condition than the Unrelated B/R condition when they probed information presented in Sentence 1, and showed no such pattern for information presented in Sentence 2. We suggest that this pattern should be interpreted cautiously since two other effects apparent in both experiments indicate that it may be the consequence of a floor effect. As in Experiment 1, Experiment 2 found lower processing costs for belief probes than for reality probes, and lower processing costs for information presented in Sentence 2 than for information presented in Sentence 1. Thus, Experiment 2 showed smaller effects overall, and failed to find differences between False B/R and Unrelated B/R conditions in the trial types that both experiments show to have the lowest overall processing costs (i.e., belief probes, and especially for Sentence 2 information). It may be that the effects of conflict in the False B/R condition are less severe when overall processing costs are low.

4. General discussion

The non-inferential false belief task allows us to investigate how adults mentally represent ToM information independent of any need to make ToM inferences. Adult participants showed greater processing costs (were slower and/or more error-prone) when informed about a false belief, the content of which conflicted with reality, than when informed about a belief whose content did not conflict with reality. This pattern would not be expected if the information in each stimulus sentence was treated in isolation and processing costs were determined solely by the number of people, objects and properties that participants held in mind: there were more "bare facts" in Unrelated Belief/Reality sentences (one man, *two* objects and two properties) than in False Belief/Reality sentences (one man, *one* object and two properties). Instead, the results suggest that participants integrated the information in the two sentences and represented the relationship between the man's belief and reality in a way that allowed interference when the man's belief was false. This is the first direct evidence about how adults hold ToM information in mind and suggests that even when no ToM inference is necessary – as in the everyday circumstance of being told what someone thinks – thinking about false beliefs may entail resisting interference between the content of the belief and what is known about reality. It follows that individual differences in adults' executive function may contribute to the efficiency with which they can maintain an uncontaminated record of what other people think, and such effects might be particularly extreme following damage to systems that help participants resist interference from their own perspective (e.g., Samson, Apperly, Kathirgamanathan, & Humphreys, 2005).

Two additional features of the data are also worthy of note. First, the pattern of processing costs when participants held in mind information about false beliefs differed from that observed in children. Whereas children show an asymmetrical effect of knowledge of reality interfering with judgements about beliefs (and not vice versa), the current studies showed interference on both types of judgement, with larger interference on judgements about reality than judgements about beliefs. Thus, the current findings cannot be explained on the basis of a cognitive bias such as the "reality bias" (e.g., Mitchell & Taylor, 1999) or "curse of knowledge" (e.g., Birch & Bloom, 2007). Further work should determine if these differing patterns are a consequence of differences between the current method in which adults made a large number of similar judgements under time pressure, and those used with children, who would typically complete only a small number of trials and would not be placed under time pressure. It is noteworthy that, when tested on tasks without time pressure, similar asymmetrical patterns to those observed in children have been observed in adults with brain injury (e.g., Samson et al., 2005). This suggests that adults do not simply process belief and reality information in a different way from children.

Instead, we speculate that the observed pattern may be revealing something about how ToM information is processed on-line. In order to form a coherent representation of information about someone's belief that conflicts with reality, it may be necessary to inhibit or suppress the information about reality temporarily. In children, a failure to inhibit information about reality has often been suggested as a reason for failure on false belief tasks (e.g., Carlson & Moses, 2001; Leslie et al., 2004; Mitchell, 1996; Russell, 1996). We suggest than in adults, successful inhibition of reality information may result in greater processing costs for reality probes while the incoming information is being processed. If this inhibition is only temporary then we might expect the pattern to change towards that observed in children if belief and reality probes were presented after a longer interval from initial presentation of the belief and reality sentences, when information about reality was no longer being inhibited.

A second noteworthy pattern was that both experiments showed greater processing costs when probes concerned information presented in Sentence 1 than when they concerned information presented in Sentence 2. Again, we think this may reflect the dynamics of information processing on this task. It seems reasonable that participants might have been faster and more accurate at responding to probes concerning Sentence 2 information because this is the sentence they have read most recently. In contrast, participants had to hold Sentence 1 information in mind while they read Sentence 2. Thus, it seems reasonable that Sentence 1 information may have suffered interference from information in Sentence 2, and, as outlined above, may indeed have been actively inhibited, at least when it described reality. Again, this might be investigated by presenting probes at a longer interval after participants read Sentence 1 and Sentence 2. We might expect this increase in storage time to increase any general interference between Sentence 1 and Sentence 2 information, but to decrease any effects of temporarily inhibiting Sentence 1 while reading Sentence 2.

The current findings are also potentially informative about the nature of children's ToM development. We find that non-inferential false belief processing remains costly for adults who have a mature concept of belief and mature syntactic abilities. Thus, although it remains possible that 3-year-olds make errors on non-inferential tasks because of an immature concept (e.g., Wellman and Bartsch,

1988) or immature syntax (de Villiers & Pyers, 2002), these may not be the most parsimonious explanations, since they cannot explain why non-inferential false belief tasks remain costly for adults. However, whatever the role of concepts and language in development, the current data are clearly consistent with working memory and executive resources having a lasting role in ToM processing.

As described in Section 1, a number of proposals have been made about the developmental relationship between executive function and ToM. If executive function continues to have a role in adults' ToM processes, then this by no means rules out a role for executive function in the emergence of ToM in children (e.g., Carlson & Moses, 2001; Russell, 1996), but it does suggest that executive function is not only involved in the emergence of ToM but may also have specific roles in on-line ToM processes. What remains unclear is whether, in on-line processing, memory and executive processes are involved only in the efficient use of ToM concepts (e.g., Leslie et al., 2004), or whether memory and executive abilities are themselves constitutive of ToM competence (e.g., Russell, 1996).

Advancing this debate requires recognition that executive function is not a unitary construct, and that different aspects of executive function such as working memory, resolving cognitive conflict and inhibiting an on-going response may play different roles in development and in on-line processing. Likewise, it may be necessary to recognise that ToM – even particular components of ToM such as belief reasoning – consist of multiple processes, including inferring, storing and using ToM information. Different aspects of executive function may play distinctive roles in each of these processes. Like previous studies with children (e.g., de Villiers & Pyers, 2002; Flavell et al., 1990; Wellman et al., 2001) the current study of adults isolates the process of thinking about ToM information from the need to make ToM inferences. Preliminary evidence to isolate processing costs associated with ToM inferences per se comes from a study by Apperly, Simpson, Riggs, Samson, and Chiavarino (2006), who found slower reaction times when adults had to infer rather than merely recall the false belief of a character in a video. Such fractionation is a necessary step towards investigating the mental representations and cognitive and neural resources involved in ToM. Data from adults and from parallel research in children are necessary for a full account of ToM, and the relationships that ToM has with other cognitive processes such as language and executive function.

Acknowledgement

This work was supported by a grant from the ESRC: RES-000-23-1419.

References

Apperly, I. A., Samson, D., & Humphreys, G. W. (2005). Domain-specificity and theory of mind: Evaluating evidence from neuropsychology. *Trends in Cognitive Sciences, 9*(12), 572–577.

Apperly, I. A., Simpson, A., Riggs, K. J., Samson, D., & Chiavarino, C. (2006). Is belief reasoning automatic? *Psychological Science, 17*(10), 841–844.

Please cite this article in press as: Apperly, I. A. et al., The cost of thinking about false beliefs: ..., *Cognition* (2007), doi:10.1016/j.cognition.2007.05.005

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a ‘‘theory of mind’’? *Cognition, 21*, 37–46.

Birch, S. A. J., & Bloom, P. (2007). The curse of knowledge in reasoning about false beliefs. *Psychological Science, 18*(5), 382–386.

Bloom, P., & German, T. P. (2000). Two reasons to abandon the false belief task as a test of theory of mind. *Cognition, 77*, B25–B31.

Bull, R., Phillips, L. H., & Conway, C. (submitted for publication). The role of control functions in mentalising: Dual task studies of Theory of Mind and Executive Functioning.

Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children’s theory of mind. *Child Development, 72*(4), 1032–1053.

de Villiers, J. G., & Pyers, J. E. (2002). Complements to cognition: A longitudinal study of the relationship between complex syntax and false-belief-understanding. *Cognitive Development, 17*, 1037–1060.

Flavell, J. H., Flavell, E. R., Green, F. L., & Moses, L. J. (1990). Young children’s understanding of fact beliefs versus value beliefs. *Child Development, 61*, 915–928.

Fletcher, P. C., Happé, F., Frith, U., Baker, S. C., Dolan, R. J., Frackowiak, R. S. J., et al. (1995). Other minds in the brain: A functional imaging study of ‘‘theory of mind’’ in story comprehension. *Cognition, 57*, 109–128.

German, T. P., & Hehman, J. A. (2006). Representational and executive selection resources in ‘‘theory of mind’’: Evidence from compromised belief-desire reasoning in old age. *Cognition, 101*, 129–152.

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

Kinderman, P., Dunbar, R., & Bentall, R. P. (1998). Theory-of-mind deficits and causal attributions. *British Journal of Psychology, 89*, 191–204.

Leslie, A. M., Friedman, O., & German, T. P. (2004). Core mechanisms in ‘‘theory of mind’’. *Trends in Cognitive Sciences, 8*, 528–533.

Lillard, A. S., & Flavell, J. H. (1992). Young children’s understanding of different mental states. *Developmental Psychology, 28*, 626–634.

McKinnon, M. C., & Moscovitch, M. (2007). Domain-general contributions to social reasoning: Theory of mind and deontic reasoning re-explored. *Cognition, 102*(2), 179–218.

Mitchell, P. (1996). *Acquiring a conception of mind. A review of psychological research and theory*. Hove, E. Sussex: Psychology Press.

Mitchell, P., & Taylor, L. M. (1999). Shape constancy and theory of mind: Is there a link? *Cognition, 70*, 167–190.

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

Perner, J., Sprung, M., Zauner, P., & Haider, H. (2003). Want that is understood well before say that, think that, and false belief: A test of de Villers’s linguistic determinism on German-speaking children. *Child Development, 74*(1), 179–188.

Russell, J. (1996). *Agency: Its role in mental development* (1st ed.). UK: Erlbaum.

Rutherford, M. D. (2004). The effect of social role on theory of mind reasoning. *British Journal of Psychology, 95*, 91–103.

Samson, D., Apperly, I. A., Kathirgamanathan, U., & Humphreys, G. W. (2005). Seeing it my way: A case of selective deficit in inhibiting self-perspective. *Brain, 128*, 1102–1111.

Smith, M., Apperly, I. A., & White, V. (2003). False belief reasoning and the acquisition of relative clause sentences. *Child Development, 74*, 1709–1719.

Stone, V. E., Baron-Cohen, S., & Knight, R. T. (1998). Frontal lobe contributions to theory of mind. *Journal of Cognitive Neuroscience, 10*, 640–656.

Wellman, H. M., & Bartsch, K. (1988). Young children’s reasoning about beliefs. *Cognition, 30*, 239–277.

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.

Zelazo, P. D., Muller, U., Frye, D., & Marcovitch, S. (2003). The development of executive function in early childhood. *Monographs of the Society for Research in Child Development, 68*(3), [Serial No. 274.].

Please cite this article in press as: Apperly, I. A. et al., The cost of thinking about false beliefs: ..., *Cognition* (2007), doi:10.1016/j.cognition.2007.05.005