

This article was downloaded by:[University of Aberdeen]
On: 1 August 2007
Access Details: [subscription number 773500140]
Publisher: Psychology Press
Informa Ltd Registered in England and Wales Registered Number: 1072954
Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



European Journal of Developmental Psychology

Publication details, including instructions for authors and subscription information:
<http://www.informaworld.com/smpp/title~content=t713684600>

Can task modifications influence children's performance on false belief tasks?

Online Publication Date: 01 September 2007

To cite this Article: Brown, Emma L. and Bull, Rebecca (2007) 'Can task modifications influence children's performance on false belief tasks?', European Journal of Developmental Psychology, 4:3, 273 - 292

To link to this article: DOI: 10.1080/17405620600780951

URL: <http://dx.doi.org/10.1080/17405620600780951>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article maybe used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

© Taylor and Francis 2007

Can task modifications influence children's performance on false belief tasks?

Emma L. Brown and Rebecca Bull
University of Aberdeen, UK

Theory of mind (ToM) is commonly measured in young children by the false belief task. Many researchers claim that 3-year-olds' failure on false belief tasks is due to task constraints. Experiment 1 empirically tested the claim made by Wellman, Cross, and Watson (2001), that even when variables that enhance performance on the false belief task (deception; participation in transformation; salience; and item not real and present) are included in one task, 3-year-olds will still fail to perform above chance level. On the standard task 4-year-olds did perform significantly better than 3-year-olds, although there was no significant difference on the adapted task. Compared to the standard task, 3-year-olds performed better on the adapted task, although this improvement was not significant and performance was still not above chance level. Experiment 2 examined the influence of two additional variables; presence of protagonist at the time when the false belief question is asked, and the order in which test questions are asked (memory preceding false belief or vice versa). Results found no significant effects of either presence of the protagonist or question order on the performance of participants. Results are discussed with reference to reasons behind the failure of some children on false belief tasks.

Premack and Woodruff (1978) claimed that, "In saying that an individual has a theory of mind, we mean that the individual imputes mental states to himself and to others" (p. 515), and that this ability leads one to be able to predict the actions of another in a given situation. A *mental state* relates to beliefs, wishes, thoughts, wants, fears and ideas (e.g., Bartsch & Wellman, 1989). When an individual understands that another person has a particular mental state, and believes that they will act in a certain way as a result of that mental state, they are exercising a theory of mind (ToM). One of the

Correspondence should be addressed to Rebecca Bull, School of Psychology, William Guild Building, University of Aberdeen, King's College, Aberdeen AB24 2UB, UK. E-mail: r.bull@abdn.ac.uk

most common areas of investigation in ToM assessment is the child's understanding of belief, and in particular, their understanding of false belief. A test of false belief is one that distinguishes beliefs from reality.

False belief is often assessed using an unexpected transfer task. For example, in Wimmer and Perner's (1983) Maxi's Chocolate Task, children hear a story about Maxi who has a bar of chocolate, which he puts in cupboard X and leaves to play. In his absence Maxi's Mother uses some of the chocolate and puts the remainder of the chocolate in cupboard Y. Maxi returns and the child is told that Maxi is hungry and wants some chocolate. The child is asked where Maxi will look for his chocolate. To pass this test of false belief, the child must understand that another person (Maxi) can hold a belief, or representation of the world, which is false. Wimmer and Perner found that when asked where Maxi would look for his chocolate, no 3- to 4-year-olds, and only 57% of 4- to 6-year-olds pointed to location X, thus attributing a false belief, whereas almost all 6- to 9-year-olds pointed to this location. Those children who got the answer wrong were not failing because of a memory problem, as 80% of those who failed passed a memory test asking where Maxi put his chocolate in the beginning.

Many authors claim that when certain test adaptations are made, 3-year-olds can pass a test of false belief (e.g., Chandler & Hala, 1994; Lewis & Osborne, 1990; Mitchell & Lacohee, 1991; Siegal & Beattie, 1991), whereas others retain the claim that a conceptual deficit prevents 3-year-olds from passing under any circumstance (e.g., Perner, Leekam, & Wimmer, 1987; Wimmer & Perner, 1983). Wellman, Cross, and Watson (2001) conducted a meta-analysis to clarify the reasons why young children consistently fail tests of false belief. Over 500 conditions were coded for the dependent variable (proportion of false-belief questions answered correctly), and 13 independent variables were assessed. A clear developmental trend was observed in the data, with children in the younger age group (3 years 5 months and younger) consistently erring on the task, and older children (4 years and above) consistently making correct judgements. Significant main effects for a number of independent variables were also found, affecting the performance of all age groups. They were: motive for transformation; participation in the transformation; salience of the protagonist's mental state; and real presence of the target object.

Motive and participation. The variable "motive" refers to the reason given to the child for the transformation occurring. Performance was optimal across all age groups when the motive was deceptive, and explicitly stated (e.g., the child is told that the chocolate was moved to trick the protagonist). The variable "participation" assessed whether or not the child actively participated in the moving of the object, or whether they simply saw the movement take place. Wellman et al. (2001) found enhanced

performance across all age groups when the child actively made, or helped to make, the crucial transformation.

An influential study on deception and false belief understanding was conducted by Chandler and Hala (1994). Initially they found that when the child actively participated in deceiving another person by moving the object, the performance of 3-year-olds on the false belief question was near ceiling. Chandler and Hala then investigated the effect of the child simply watching another person move an object in an effort to trick another person—this had no effect on performance compared to a standard non-deceptive version of the task. Finally, they investigated whether merely participating in planning a deception was sufficient to enhance performance (with the experimenter carrying out the actual transfer). Results indicated that simply participating in the planning of the deceptive transfer was sufficient to render performance of 3-year-olds above chance level on the false belief task. Therefore, it can be concluded from this study that provided the 3-year-old participates in either the planning of a deceptive act, or actively participates in it, their performance on the false belief task should be enhanced.

Salience and presence. Salience of the protagonist's mental state relates to whether the belief of the protagonist is clearly stated, or pictured, before the false belief question is asked. For example, in Maxi's chocolate task, the experimenter may say, "Maxi thinks his chocolate is in cupboard X" before asking, "Where will Maxi look for his chocolate?" Wellman et al. (2001) concluded that when the protagonist's mental state is highlighted in such a way, overall performance is enhanced (e.g., Mitchell & Laco  e, 1991; Sullivan & Winner, 1991; Woolley, 1995).

Woolley (1995) conducted a study with the deceptive box task (Hogrefe, Wimmer, & Perner, 1986) where participants are shown a typical container (e.g., a matchbox) and are asked what they think is inside. Participants generally state the typical contents. They are then shown that they were holding a false belief, and that the box actually contains another object (e.g., chocolate). The box is then closed and the child is asked, "What will (name of friend/protagonist) say is in the box?" Woolley (1995) manipulated this task by having a protagonist present at the beginning of the study and having them clearly state their belief prior to the false belief question being asked. Results showed that 3-year-olds were correct 48% of the time when this enhancement was made to the standard deceptive box task. Likewise, when children posted (in a pretend post box) a picture of their belief about the contents of the deceptive box when they made their initial judgement, their performance was enhanced on the subsequent false belief question compared to a standard version of the task (Mitchell & Laco  e, 1991). Thus, when the protagonist's mental state is highlighted or when the salience of the child's own previous mental state was increased (by posting

the picture), 3-year-olds performance on the false belief question was enhanced.

Real presence. The variable real presence relates to whether a real object is present at the time the false belief question is asked. This variable is also a measure of salience, but relates to the salience of reality, rather than the salience of the protagonist's mental state. For example, in Maxi's chocolate task, the object is real (chocolate) and present (in the cupboard) when the question is asked. However, if Maxi's chocolate was eaten and the question asked, the object would no longer be present leaving no alternative location for the child to consider, reducing the salience of reality.

Wimmer and Perner (1983) found that while performance of 4-year-olds was enhanced when Maxi's Mother used all of the chocolate for her cake, 3-year-olds performance remained below chance level. However, Fritz (1990, cf. Woolley, 1995), found contrasting results. Performance of 3-year-olds was greatly enhanced in the deceptive box task when the box was emptied of its contents prior to the false belief question being asked. Koos, Gergely, Csibra, and Biro (1997) found that children younger than 3 years and 5 months performed at around chance level on an unexpected transfer task where the chocolate was eaten, whereas older children performed above chance. The performance of the younger children was better than the typical performance of this age group on the standard unexpected transfer task, and therefore these findings add further support to the claim that when the object is not present at the time the false belief question is asked, children's performance is enhanced.

On the basis of these results, Wellman et al. (2001) concluded that the performance of young 3-year-olds can be enhanced on the false belief task only when at least one of the significant test manipulations is included (deception, participation, salience and absence of object). Although the enhancements are said to improve performance, to date they have only improved the performance of young 3-year-olds from below chance to at-chance levels. Wellman et al. (2001) hypothesized, based on statistical analysis, that even if all four variables were employed in one false belief test, young 3-year-olds would still only perform at chance level. However, this has never been tested empirically. Although Wellman et al. make no exact definition of what they class as "chance level", they appear to take it to mean the level at which 50% of participants would pass, and 50% would fail. The aim of the current study was to replicate previous results, which demonstrated that while 4-year-olds can pass a standard false belief test, 3-year-olds cannot, and to test the above hypothesis in order to examine whether young 3-year-olds could systematically pass tests of false belief if all four enhancements were made within one modification of the task.

EXPERIMENT 1

Method

Participants

Forty-nine children participated in this study. Children were grouped according to age. Group 1 consisted of 25 children aged between 3 years 0 months and 3 years 11 months ($M_{\text{age}} = 3$ years 6 months, $SD = 2.58$ months); group 2 consisted of 24 children aged between 4 years 0 months and 4 years 11 months ($M_{\text{age}} = 4$ years 7 months, $SD = 3.52$ months). Groups included equal numbers of boys and girls. Children were recruited from various nurseries following parental consent. Verbal IQ was assessed for 42 of the children using the Wechsler Pre-school and Primary Scale of Intelligence – Revised (WPPSI-R, Wechsler, 1990). In terms of age standardized vocabulary scores, the sample was representative of their age group in the general population (group 1: $M = 12.21$, $SD = 2.22$; group 2: $M = 11.59$, $SD = 2.75$).

Materials

Participants took part in one of two tests. In the standard Maxi's chocolate task, the children viewed a short play where two confederates acted out the original test using a red and green box, and two sweets. In the adapted version of Maxi's chocolate task, the child participated in the removal and consumption of a confederate's sweet from the red/green box. The two boxes were always present although only one box was used as a hiding place for the purpose of the adapted study, therefore allowing two possible options for the child to select from in both tests.

Design

The dependent variable was the response given by the participant (i.e., correct or incorrect location to search for the object). The design was between subjects with comparisons being made between different age groups, and between performance on the standard and adapted tasks. The experiment employed a counterbalancing procedure where half of the children saw the sweet being left in the red box by the confederate and the other half saw the sweet being left in the green box.

Procedure common to both standard and adapted tasks

All children were tested individually in a quiet corner of a classroom or a separate room. Participants were first introduced to the experimenter and

confederate(s) by their teacher. The experimenter checked that the child knew the colours (red and green) of the two boxes in front of them. The experimenter reminded the child about the confederate and pointed out that (s)he had something in her/his hand. This phrase prompted the confederate to show the child their sweets, saying, "Yes, look I've got some lovely fruit pastilles here, but I'm all full up now, so I'll just put them in this red/green box while I go and see my friend". The confederate then placed the sweets in the red/green box, replaced the lid and left the room.

Procedure for adapted task

After the confederate had left the room, the adapted version proceeded as follows. The experimenter glanced at the door and back to the child in a conspiring manner, before saying, "Good, (s)he's gone, did you see the colour of those sweets (s)he put in the red/green box?" The experimenter then opened the box, took out the sweets and showed the child, before asking him/her to name the colours. This was done to engage the child in the experiment and ensure that they were paying attention. The experimenter then said to the child, "I know, why don't we play a really good trick (deception) on [confederate's name], why don't you take the sweets out of the box (participation) and we eat them (item no longer present), you could eat one and I'll eat the other". All children agreed to engage in the deception. The sweets were unwrapped and eaten by the experimenter and participant before the memory and test question were asked. The memory question ("Where did [confederate's name] leave his/her sweets?) was used to check that the child could remember the original hiding location. If the child failed the memory question, their results were not included in the final analysis. The experimenter then said, "Right, so [confederate's name] thinks that his/her sweets are in the red box (enhancing salience), but we played a trick on him/her", before asking the false belief question, "So where will [confederate's name] look for his/her sweets when (s)he comes back?"

Procedure for standard task

Following the departure of the first confederate in the standard task the child saw a second confederate enter the room and ask if the experimenter had seen confederate 1. The experimenter replied that she had, and that (s)he had been here, but had left his/her sweets in the red/green box before going out to see his/her friend. Confederate 2 then opened the box to have a look. (S)he said, "Ooh, I like fruit pastilles, I think I'll have one of these", before taking the two sweets out of the box, replacing the lid, eating one of the sweets and putting the remaining sweet in the green/red box (not the box in which the sweets were originally) before saying (s)he was going shopping.

Once the second confederate had left, the memory and false belief questions were asked. The memory question was identical to that asked in the adapted task, but in the standard task, the false belief question was, "Where will [confederate's name] look for his/her sweets?"

Results

In order to answer the question correctly in either task, the child had to state that confederate 1 would look in the box where they left the sweets (red/green). An incorrect response was recorded if the child said she would look in the other box, gave no response, or gave an unrelated response (e.g., "my tummy", "the cupboard"). Data from one 4-year-old and four 3-year-old children were excluded because they failed the memory question.

Association between age and task success

An initial chi-square analysis of age group (3.0–3.11 or 4.0–4.11) by success (passing/failing) on the standard condition found, as predicted, a significant association between age and success on the task, with significantly more 4-year-olds passing the test than 3-year-olds, $\chi^2(1, n = 21) = 5.84$, $p < .05$. In the adapted condition, however, there was no significant association between age and success on the task, with both 3-year-olds and 4-year-olds more likely to fail than pass the task, $\chi^2(1, n = 23) = 0.006$, $p > .05$. These points are clarified further in Figure 1, where the results show

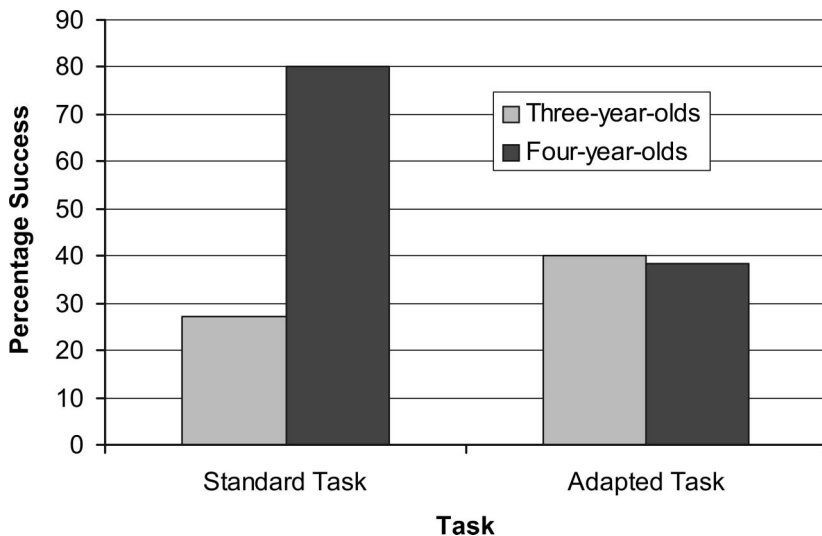


Figure 1. Percentage of children passing the standard and adapted false belief tasks.

that a higher percentage of 4-year-olds passed the standard test than did 3-year-olds, whereas a lower percentage of 4-year-olds passed the adapted test than did 3-year-olds. The 3-year-olds failed to perform at chance level on either the standard or adapted task. Binomial tests indicated that when chance level was set that 50% of participants would be expected to pass and 50% fail, participants in neither age group, for neither task type, differed significantly from chance level performance (all $ps > .05$). However, 80% of 4-year-olds passed the standard task, and so it is possible that sample size was a contributing factor behind the lack of significant effects.

To assess whether there were differences in performance dependent on task type, several Fisher Exact Probability tests were conducted between task type (standard vs. adapted) and performance, separated by age group. For 3-year-olds, while 27% passed the standard task and 40% passed the adapted task, there was no significant effect of task type on performance ($p > .05$), indicating similar levels of performance in both tasks. For 4-year-olds there was a marginally significant effect of task type on performance, with 80% of participants passing the standard task and 39% passing the adapted task ($p = .09$).

Discussion

Overall, the results support the suggestion of Wellman et al. (2001); 3-year-old children did not perform above chance levels on the false belief task even when all performing enhancing variables were incorporated. Furthermore, the results support previous research on the standard unexpected transfer task, with 4-year-olds performing significantly better than 3-year-olds (e.g., Wimmer & Perner, 1983). However, several surprising results require further consideration.

First, as predicted by Wellman et al. (2001), performance of 3-year-olds was better in the adapted than the standard task, although this difference was not significant. However, Wellman et al. also proposed that the incorporation of the enhancing variables into the false belief task should also improve the performance of 4-year-olds. In the current study, performance of 4-year-olds was marginally significantly worse on the adapted compared to the standard version of the false belief task. The most obvious cause of this difference is the use of a between-subjects design, with different groups of children taking part in each of the standard and adapted tasks. However, care was taken to ensure that groups were as evenly matched as possible and they did not differ in verbal IQ. It is possible that the 4-year-olds were unnecessarily complicating the adapted task by looking for an alternative logical answer, as the correct answer appeared too simple. Clearly, when the child pointed to their mouth as a response in the adapted task, they were not thinking logically (Russell, 1996, also found evidence of

children responding in this way), however, on occasions where the child said, "The green/red box" (giving an incorrect response), they generally looked perplexed by the question and seemed to give a great deal of thought to their response. Perhaps they were thinking that because the experimenter had just said, "So [confederate's name] thinks that his/her sweets are in the red box", the answer could not possibly be as simple as repeating what the experimenter had just said. In hindsight it would have been interesting to note the percentage of participants who responded with an irrelevant response (e.g., tummy) to the false belief question in order to assess whether chance level could be set differently for this task. However this data is not available.

The majority of published literature to date suggests that 4-year-olds can pass a test of false belief, although Wimmer and Perner (1983) found that it was not until children were aged 6 years that they could systematically pass the false belief task. Furthermore studies not included in the meta-analysis due to their recent status as published articles have also found that not all 4-year-olds pass false belief tests, or that children of this age benefit from simplifications to the false belief tasks (e.g., Hughes, Adlam, Happé, Jackson, Taylor, & Caspi, 2000; Nadel & Melot, 2001). Such results demonstrate that the traditional claim that 4-year-olds systematically pass false belief tasks should be accepted with caution, and results from the current study are in keeping with these recent findings.

Overall, the lack of a significant change in performance on the adapted task runs contrary to findings from Chandler, Fritz, & Hala (1989), Hala, Chandler, and Fritz (1991), Sullivan and Winner (1991, 1993), Chandler and Hala (1994), Hala and Chandler (1996), Woolley (1995), Mitchell and Lacohee (1991), and Koos et al. (1997), who found that when one or more of the enhancing variables was included, performance on the false belief task was significantly improved. However, in these studies, enhancing variables were seldom studied on their own, most often being combined with at least one other enhancing variable, making it difficult to assess which specific variable resulted in improved performance.

To conclude, the performance of 3-year-olds was poor on both the standard and adapted tasks whereas the performance of 4-year-olds was marginally significantly better on the standard task than the adapted task. However, before the conclusion that 3-year-olds cannot pass a test of false belief is made, it is necessary to eliminate certain variables that Wellman et al. (2001) did not consider in their meta-analysis. One such variable relates to the presence of the protagonist at the time of questioning. It may be the case that when the protagonist is present in the unexpected transfer task, the participant feels more inclined to "tell" the protagonist where the object is and thus fail the task. A second variable not considered as a

mediating factor on false belief task performance is the order of presentation of the memory and false belief questions. The child may fail whichever question is asked first because of inhibitory control deficits (e.g., Carlson & Moses, 2001, Carlson, Moses, & Breton, 2002); their automatic tendency is to reveal the object's true location, and this may result in failure of either the false belief question ("Where will [confederate's name] look for his/her object?") or the memory question ("Where did [confederate's name] put his/her object?"), whichever is asked first. It is important to investigate these variables, as in some false belief tasks the protagonist (whether puppet, cartoon, video or real person) is present when the false belief question is asked (e.g., Nadel & Melot, 2001; Wimmer & Perner, 1983), and in other tasks the protagonist has been absent (e.g., Hala & Chandler, 1996). In addition, in some false belief tasks participants are asked the false belief question prior to any control questions (e.g., Hala et al., 1991; Siegal & Beattie, 1991, Wimmer & Perner, 1983) and in other tasks at least one control question is asked prior to the false belief question (e.g., Hala & Chandler, 1996). Experiment 2 investigated the influence of the presence of protagonist and question ordering upon performance on the false belief task.

EXPERIMENT 2

Method

Participants

Seventy-three children participated in this study. The participants were aged between 3 years and 5 years 2 months and split into two groups—3-year-old participants ($n = 30$, $M_{\text{age}} = 3$ years 6 months, $SD = 3.4$ months), and 4-year-old participants ($n = 43$, $M_{\text{age}} = 4$ years 7 months, $SD = 4.6$ months). Children were recruited from various nurseries following parental consent. Verbal IQ was assessed for 69 of the children using the Vocabulary subtest of the WPPSI-R (Wechsler, 1990). In terms of vocabulary, the sample was representative of their age group in the general population (3-year-olds, $M = 12.74$, $SD = 2.12$; 4-year-olds, $M = 11.63$, $SD = 1.85$).

Design

The design was mixed with subjects variable of protagonist's presence and between subjects variable of question ordering (false belief question followed by memory question or vice versa). The dependent variable was the response given by the participant (i.e., correct or incorrect location to search for the object). The experiment controlled for a colour preference by employing a counterbalancing procedure where for one false belief task (the

chocolate task), the chocolate was always hidden in the red cupboard, and in the other false belief task (the book task), the book was always hidden in the green drawer.

Materials

Participants took part in four standard false belief tasks. In tasks one and two (chocolate task), children saw a bar of chocolate being placed in a red cupboard, before it was later transferred to the green cupboard in the protagonist's absence. In tasks three and four (book task), children saw a book being placed in a green drawer, before it was later transferred to the red drawer in the protagonist's absence. Participants also took part in the same true belief task on two occasions, where a picture of a dog was placed in a red box. A green box was also present throughout this task although it was never used in the task.

Procedure

All children were tested individually in a quiet corner of a classroom or a separate room. In the first session participants were given two false belief tasks (one book, one chocolate), one of which had a protagonist present when the test question was asked, and the other which had an absent protagonist. The order in which participants received the task (present vs. absent) and the task type (chocolate present vs. book present, chocolate absent vs. book absent) was counterbalanced across participants and testing sessions. In the second testing session participants received the two alternate false belief tasks. Participants were also given two true belief tasks, one in each test session, where the target object was replaced in the initial location. There were four weeks between testing sessions. Protocols for the different stories are given in the Appendix. True belief tasks were included to check that the child would be willing to respond to the same location twice and to eliminate the possibility that success on the false belief task was due to simple alternation between location selection. The independent variable of question order was a between-subjects variable, with approximately half of the participants always receiving the memory control question before the belief question, and the other half receiving the belief question prior to the memory control question.

Results

To pass the false belief question, the child had to state that the protagonist would look for the chocolate/book in the location where the protagonist left it. An incorrect response was recorded if the child stated that the protagonist

would search for the chocolate/book in the object's current location. In order to pass the true belief conditions the child had to respond that the protagonist would search for the object in the current location, as this was the same location as the original location for the object. Data from one 3-year-old and one 4-year-old were excluded due to failure on both true belief tasks.

Traditionally, if a child fails the memory question in a false belief task, their results are excluded from further analysis. On this occasion between 19.7% and 26.8% of children failed the memory question in one or more of the false belief conditions. Hughes et al. (2000) report a similar pass rate (72%) on memory questions. For analyses, Hughes et al. adopted a method whereby children who failed the memory question were classified as having failed the false belief task even if they answered the false belief question correctly. A similar classification procedure was employed in the current study.

Association between question order and performance

Figure 2 shows the percentage of participants passing the false belief question for condition 1 (chocolate, protagonist present), condition 2 (chocolate, protagonist absent), condition 3 (book, protagonist present), condition 4 (book, protagonist absent), condition 5a (true belief task 1) and 5b (true belief task 2) by age group and question order. The graph shows that none of the 3-year-olds who received the false belief question first correctly answered this question in conditions 1, 2 or 3. However, when the memory question was asked first, approximately 20% of children passed the subsequent false belief question in each of these conditions.

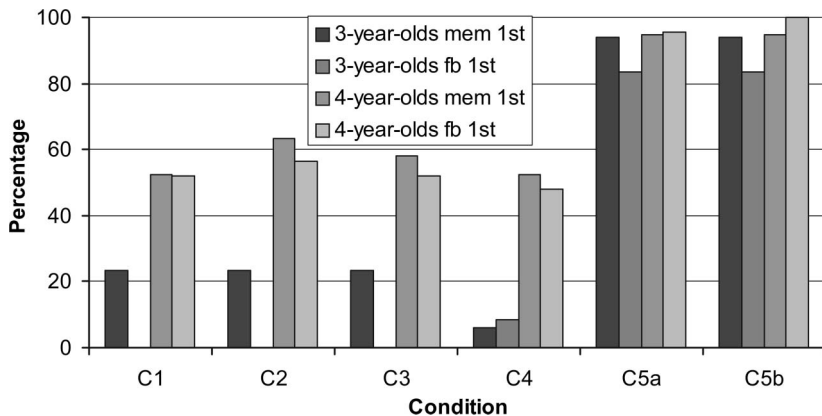


Figure 2. Percentage of children passing by age group and question order.

To investigate this observation further, Fisher Exact Probability analyses were conducted to determine whether there was an association between question order and performance on any of the individual tasks. For both 3- and 4-year-olds, there was no significant association between task performance and question order (all $ps > .05$). Chi-square analyses revealed that collapsed across age group the associations were non-significant (all $ps > .05$).

The effect of question order may have been masked by the scoring criterion (children were classed as failing the false belief question when they failed the memory question, regardless of their performance on the false belief question). To investigate the issue of question order more comprehensively, participants were assigned a combined score out of four according to success on the four false belief tasks (where 0 = failed all tasks and 4 = passed all tasks). A 2 (Question Order) \times 2 (Age Group) mixed-design analysis of variance (ANOVA) was conducted with the overall false belief score as the dependent variable. This analysis found no significant effect of question order upon performance, $F(1, 70) = 1.25$, $p > .05$, and no significant interaction between question order and age group, $F(1, 70) = 0.18$, $p > .05$. There was a significant effect of age group, with 4-year-olds performing significantly better than 3-year-olds, $F(1, 70) = 22.56$, $p < .001$. While there was no significant effect of question order upon performance, the overall mean false belief score was higher for both age groups when the memory question was asked first (3-year-olds $M = 0.88$, $SD = 0.11$; 4-year-olds $M = 2.42$, $SD = 1.87$), compared to when the false belief question was asked first (3-year-olds $M = 0.33$, $SD = 0.49$; 4-year-olds $M = 2.17$, $SD = 1.61$). For additional analyses performance was collapsed across question order.

Presence of the protagonist and performance

Phi coefficients were calculated to assess whether it was possible to collapse performance across the two protagonist present tasks and the two protagonist absent tasks. Results revealed that there was a significant association between performance on the two present tasks, $\phi(1, n = 71) = .67$, $p < .001$, and the two absent tasks, $\phi(1, n = 71) = .65$, $p < .001$, indicating consistent performance across the two versions. Therefore results were collapsed to give a present and absent "score".

A 2 (Age Group) \times 2 (Presence of Protagonist) mixed-design ANOVA revealed only a significant effect of age group, $F(1, 69) = 23.14$, $p < .001$, with children in the older age group achieving a higher overall score than those in the younger age group. There was no significant main effect of protagonist presence or a significant interaction between age and protagonist presence (both $F_s < 1$).

Discussion

In line with findings from numerous previous studies, there was clear evidence of an association between age and performance on all of the false belief tasks used in the current study, with 4-year-olds performing more accurately than 3-year-olds. More importantly, the study sought to determine whether these age-related patterns might vary depending on the question ordering and the presence of the protagonist at the time the false belief question was asked.

Association between question order and performance on dependent variables

There were no significant associations between question order and performance on the false belief questions. Reanalysis of the data with the less conservative scoring criterion (memory failers not classified as false belief failers) revealed a trend for both 3- and 4-year-old children to be more accurate on the false belief question when this was the second question asked (i.e., after rather than before the memory question) although this effect was not significant. Peter Mitchell (personal communication, 17th December 2003) suggests that this trend in performance may be due to a problem with inhibitory control on the child's part, in that they are not able to inhibit pointing to the current location of the object regardless of the question asked of them. On many occasions this would appear to be an acceptable explanation of the problem—the child often pointed to the object's current location when the experimenter began the question "Where . . ." without waiting to hear the question in full. When this occurred the child was asked to wait until the question was finished before responding. This highlights the potential importance of the inclusion of a reality control question, which is asked prior to the memory and false belief questions. With the inclusion of a reality question (e.g., "Where is the chocolate/book really?") the child is given the opportunity to make their automatic response of pointing to the object's current location. It could be the case that after the child has made this initial automatic response they may be more inclined to respond more accurately, or with more thought, to subsequent questions.

Presence of protagonist and performance

Performance on the two present and two absent tasks were significantly related, thus indicating that the presence of the protagonist had very little effect on performance, with the majority of children failing both the present and absent conditions, or passing both.

GENERAL DISCUSSION

In both studies 3-year-olds' performance on false belief tasks was poor. In Experiment 1 it was found that it made no difference to 3-year-olds' performance whether they were given a standard, or adapted false belief task. In contrast, 4-year-olds' performance was poorer when given the adapted version compared to the standard version of the false belief task. This leads to the tentative conclusion that 3-year-olds are unable to successfully pass a test of false belief, even when four performance-enhancing variables are included in one false belief task. Furthermore, the performance of 4-year-olds is not robust across false belief tasks—when changes are made to the task, as in the adapted version, the performance of 4-year-olds differs from their performance when the task is presented in standard format.

Experiment 2 demonstrated that the order in which test and control questions are asked may influence the performance of 3-year-olds on the false belief task. It is proposed that when 3-year-olds are asked the memory question first, they are more likely to pass a subsequent false belief question than when they are asked that false belief question first. Although this trend did not reach the standard alpha level of 5% significance the trend should be investigated further in future research as it may be the case that the relatively small sample size was responsible for the lack of significant effect in this study.

Furthermore, Experiment 2 investigated the role of the presence of a protagonist at the time that the false belief question is asked, and found no significant effect of presence of the protagonist. This finding suggests that the presence of the protagonist has not influenced performance on the false belief task in past studies, and will not do so in future studies. This is an important finding and contributes to the debate surrounding what does and does not influence children's performance on the false belief task (see Wellman et al., 2001, for a meta-analysis). However, this conclusion should be accepted with some caution as it may be that the variable of protagonist presence would have an effect if it were manipulated alongside other variables such as those manipulated in Experiment 1.

Several potential explanations have been proposed to explain why children fail tests of false belief yet pass other theory of mind tasks. Arguably, one of the most convincing explanations relates to the inhibitory demands posed by false belief tasks. As noted earlier, participants frequently failed the memory question when it was asked first in the second study. This could be due to the participant's automatic tendency to point to the true location of an object. Fodor (1992) and Bloom and German (2000) note that in order to pass a test of false belief participants must ignore the natural heuristic, which would state that the person who wants the object will go to

the location of the object in order to satisfy their desire, and instead state that the person who wants the object will look in the location where the object is not. This is not a natural occurrence—typically people act in a manner that satisfies their desires (Fodor, 1992), and young children may find it difficult to override this heuristic and consider what the person believes in addition to considering what they desire. This relates to behavioural inhibition—the child must override the prepotent response of pointing to the object's true location. However, a second type of inhibitory demand, knowledge inhibition, exists in the false belief task. When the false belief question is asked, the participant is aware of the object's true location, and therefore needs to inhibit this knowledge in order to assess what another person will know about the situation. Birch and Bloom (2003) found that young children find it very difficult to suppress their own knowledge about a situation in order to state what another, more ignorant, person will know. This also relates to the salience of the final location of the object mentioned earlier. It remains possible that if the object's final location is made less salient to the child participant, they may be more successful on the task. Although this was tested in Experiment 1 by having the child participant eat the object, it was not tested exclusively, and it is possible that the true effect of this variable was masked by effects of the other variables. An alternative way to reduce the child's knowledge of the final location, or reduce the salience of this location, would be to employ a design where the child is unaware of the object's final location, i.e., the participant knows that the object is no longer in the initial location, but they are unaware of its true location. It is possible that when such a manipulation occurs children might be more likely to pass the false belief task as they would have no knowledge of the final location.

Overall, these studies have found that participants aged both 3 and 4 years find it difficult to pass a test of false belief even when variables thought to enhance their performance are incorporated into the design of the task. This argues against the early competence view of failure on false belief tasks due to task constraints. However, it does support the notion of conceptual change. Supporters of the conceptual change account of false belief development argue for the presence of a real change in the way that children perceive other people during the period between 3 and 4 years. However various theories have emerged under the umbrella heading of “conceptual change accounts”, which claim that different conceptual functions are responsible for the developmental trend demonstrated in false belief task performance. Some theories claim that an inability to understand representations is the cause of failure on the false belief task (e.g., Perner, 1991), whereas others claim that it is underdeveloped executive functioning skills that cause failure on the false belief task (e.g., Moses, 2001, emergence account). This study does not uniquely support or refute either of these

theories of failure on the false belief task, although it does support the view that the child undergoes a change in the way that they conceptualize belief between the ages of 3 and 4 years.

Manuscript received 1 March 2005

Revised manuscript accepted 2 May 2006

First published online 18 May 2007

REFERENCES

- Bartsch, K., & Wellman, H. (1989). Young children's attribution of action to beliefs and desires. *Child Development, 60*, 946–964.
- Birch, S. A. J., & Bloom, P. (2003). Children are cursed: An asymmetric bias in mental-state attribution. *Psychological Science, 14*, 283–286.
- 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.
- Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development, 72*, 1032–1053.
- Carlson, S. M., Moses, L. J., & Breton, C. (2002). How specific is the relation between executive function and theory of mind? Contributions of inhibitory conflict and working memory. *Infant and Child Development, 11*, 73–92.
- Chandler, M., Fritz, A. S., & Hala, S. (1989). Small-scale deceit: Deception as a marker of two-, three-, and four-year-olds' theories of mind. *Child Development, 60*, 1263–1277.
- Chandler, M., & Hala, S. (1994). The role of personal involvement in the assessment of early false belief skills. In C. Lewis & P. Mitchell (Eds.), *Children's early understanding of mind* (pp. 403–425). Hove, UK: Lawrence Erlbaum Associates, Ltd.
- Fodor, J. A. (1992). A theory of the child's theory of mind. *Cognition, 44*, 283–296.
- Hala, S., & Chandler, M. (1996). The role of strategic planning in accessing false-belief understanding. *Child Development, 67*, 2948–2966.
- Hala, S., Chandler, M., & Fritz, A. S. (1991). Fledgling theories of mind: Deception as a marker of 3-year-olds' understanding of false belief. *Child Development, 62*, 83–97.
- Hogrefe, G. J., Wimmer, H., & Perner, J. (1986). Ignorance versus false belief: A developmental lag in attribution of epistemic states. *Child Development, 57*, 567–582.
- Hughes, C., Adlam, A., Happé, F., Jackson, J., Taylor, A., & Caspi, A. (2000). Good test–retest reliability for standard and advanced false-belief tasks across a wide range of abilities. *Journal of Child Psychology and Psychiatry, 41*, 483–490.
- Koos, O., Gergely, G., Csibra, G., & Biro, S. (1997, April). *Why eating Smarties makes you smart? Understanding false belief at the age of 3*. Poster presented at the Society for Research in Child Development Biennial Meeting, Washington, DC.
- Lewis, C., & Osborne, A. (1990). Three-year-olds' problems with false belief: Conceptual deficit or linguistic artefact? *Child Development, 61*, 1515–1519.
- Mitchell, P., & Lacoche, H. (1991). Children's early understanding of false belief. *Cognition, 39*, 107–127.
- Moses, L. J. (2001). Executive accounts of theory-of-mind development. *Child Development, 72*, 688–690.
- Nadel, J., & Melot, A. (2001). How clear is the 4-year-olds' "clear-cut change" in understanding mind? *Cognitive Development, 15*, 153–168.
- Perner, J. (1991). *Understanding the representational mind*. London: MIT Press.
- Perner, J., Leekam, S. R., & Wimmer, H. (1987). Three-year-olds' difficulty with false belief: The case for a conceptual deficit. *British Journal of Developmental Psychology, 5*, 125–137.

- Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *The Behaviour and Brain Sciences*, 1, 515–526.
- Russell, J. (1996). *Agency. Its role in mental development*. Hove, UK: Lawrence Erlbaum Associates, Ltd.
- Siegal, M., & Beattie, K. (1991). Where to look first for children's knowledge of false beliefs. *Cognition*, 38, 1–12.
- Sullivan, K., & Winner, E. (1991). When 3-year-olds understand ignorance, false belief and representational change. *British Journal of Developmental Psychology*, 9, 159–171.
- Sullivan, K., & Winner, E. (1993). 3-year-olds' understanding of mental states: The influence of trickery. *Journal of Experimental Child Psychology*, 56, 135–148.
- Wechsler, D. H. (1990). Wechsler Preschool and Primary Scale of Intelligence (revised). Kent, UK: the Psychological Corporation.
- 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.
- Woolley, J. D. (1995). Young children's understanding of fictional versus epistemic mental representations: Imagination and belief. *Child Development*, 66, 1011–1021.

APPENDIX

False belief task 1—The chocolate task

In each of the following conditions, the story was read to the child while two live actors acted out the scenario.

This is Lauren and she's waiting for Sarah to come home. Sarah returns from her shopping trip. [Sarah enters with bags] She bought chocolate for a cake. [Sarah passes Lauren the chocolate] Lauren is going to help her put away the things. Lauren asks Sarah, "Where will I put the chocolate?" "In the red cupboard", says Sarah. [spoken by Sarah]

Lauren puts the chocolate into the red cupboard. Lauren remembers exactly where she put the chocolate so that she could come back and get some later. She loves chocolate. Then she leaves for the shops. [Lauren leaves] Sarah starts to prepare the cake. [Sarah uses a spoon to stir an imaginary cake mix in a bowl] She needs the chocolate next, so she takes it out of the red cupboard. She's just about to put a piece into the bowl when she remembers that she forgot to buy the eggs for the cake. Sarah puts the chocolate away again, but she does not put it back into the red cupboard, she puts it the green cupboard. Then she goes to her neighbour to borrow some eggs.

In the protagonist present condition, the story continued as follows: Here comes Lauren back from the shops [Lauren enters room and stands in front of two cupboards], hungry and she wants to get some chocolate. She still remembers where she had put the chocolate. Where will Lauren look for the chocolate? Do you remember where Lauren put the chocolate in the

beginning? [with questions in reverse order for memory question first groups of participants]

In the protagonist absent condition, the story continued as follows: Lauren is coming back from the shops soon, she'll be hungry and wanting some chocolate. She still remembers where she had put the chocolate. Where will Lauren look for the chocolate? Do you remember where Lauren put the chocolate in the beginning? [with questions in reverse order for memory question first groups of participants]

False belief task 2—The book task

This is Lauren and she's waiting for Sarah to come home from the library. Sarah is late, but here she comes now. [Sarah enter with book] She got a book which Lauren wanted. [Sarah hands Lauren the book] Lauren asked Sarah where it would be safe to keep the book. "Where will I put the book?" [spoken by Lauren] Sarah says, "In the green drawer."

Lauren puts the book into the green drawer. Lauren remembers exactly where she put the book so that she could come back and get it later. She loves reading books. Then she leaves for the shops. [Lauren leaves] Sarah starts to do some cleaning. She dusts the surfaces and then the green drawer. She opens it up and takes the book out and then dusts inside. Then she closes the drawer and dusts inside the red drawer. Then she remembers to put the book away, but she doesn't put it in the green drawer, Sarah puts the book into the red drawer. [actions while story is read] Now she realises that she needs some more polish, so she goes to her neighbour to borrow some.

In the protagonist present condition the story proceeded as follows: Here comes Lauren back from the shops, [Lauren enters room and stands in front of two cupboards] bored and she wants to get the book. She still remembers where she had put the book. Where will Lauren look for the book? Do you remember where Lauren put the book in the beginning? [with questions in reverse order for memory question first groups of participants]

In the protagonist absent condition the story proceeded as follows: Lauren is coming back from the shops soon, she'll be bored and wanting her book. She still remembers where she had put the book. Where will Lauren look for the book? Do you remember where Lauren put the book in the beginning? [with questions in reverse order for memory question first groups of participants]

True belief story

This is Sarah and she's waiting for Lauren to come home. Lauren returns from her friend's house with a picture of her friend's new dog. Sarah asks

Lauren, "Can I see the picture?" [spoken by Sarah] "No, I'm putting it in this red box," says Lauren.

Lauren puts the picture into the red box. Lauren remembers exactly where she put the picture so that she could come back and look at it later. The dog was so cute! Then she leaves for the shops. [Lauren leaves] Sarah was angry with Lauren for not showing her the picture so she goes to the red box and takes the picture out. She says, "What a cute dog!" [spoken by Sarah] and then puts the picture back into the red box before she goes to make some coffee.

Here comes Lauren back from the shops. [Lauren enters room and stands in front of two boxes] She wants to look at the picture now. She still remembers where she had put the picture. Where will Lauren look for the picture? Do you remember where Lauren put the picture in the beginning?