# Two Experimental Variations
# On the Standard False Belief Task

Version: 1.1
Date Created: 15 July 2002
Date Last Modified: 18 September 2002

**TABLE OF CONTENTS**

**1. INTRODUCTION ..... 3**

**2. EXPERIMENTAL SUMMARY ..... 5**

**3. EXPERIMENT 1 ..... 5**

3.1 METHOD ..... 5
3.2 RESULTS ..... 6
3.3 DISCUSSION ..... 6

**4. EXPERIMENT 2 ..... 7**

4.1 METHOD ..... 7
4.2 RESULTS ..... 7
4.3 DISCUSSION ..... 8

**5. GENERAL DISCUSSION ..... 8**

**6. ACKNOWLEDGMENTS ..... 13**

**7. APPENDIX: DETAILED EXPERIMENTAL RESULTS ..... 14**

7.1 EXPERIMENT 1: INDIVIDUAL ANSWERS TO QUESTIONS POSED ..... 14
7.2 EXPERIMENT 2: INDIVIDUAL ANSWERS TO QUESTIONS POSED ..... 16
7.3 NOTES ON INDIVIDUAL ANSWERS TO EXPERIMENTS 1 AND 2 ..... 16

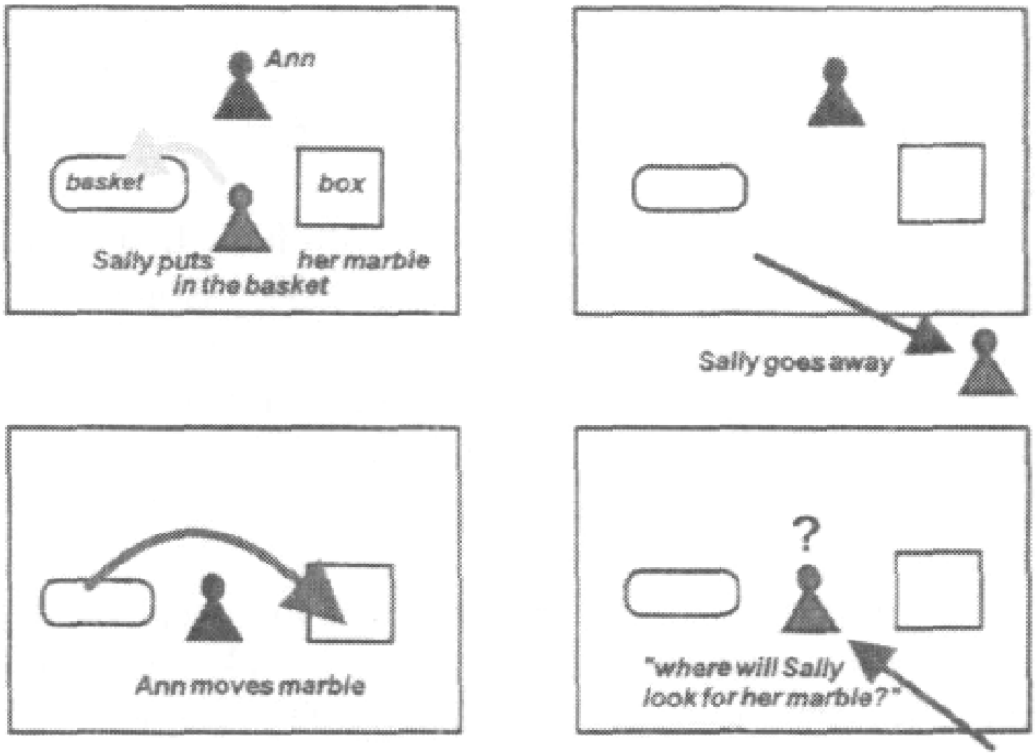
**8. REFERENCES ..... 17**

# 1. Introduction

What is Theory of Mind, what are false belief tasks, and why, almost 20 years after Wimmer and Perner first reported striking results when testing young children on false belief tasks, is there still a steady stream of experiments being performed and so much heated discussion of what leads to false belief task failure?

First, Theory of Mind. This (unfortunately unbelievably vaguely named!) concept generally refers to an agent's ability to understand mental states, in itself and others. Does an agent have a set of beliefs that guide its actions? Is the agent aware that some of those beliefs may be false? Is the agent aware that others may have beliefs as well, that some of those beliefs may be different from the agent's, and that some beliefs of others may also be false?

In order to test whether human children can answer questions like those posed above, Wimmer and Perner devised a simple task similar[1] to the scenario shown here (from Leslie, 2000):




[1] Wimmer and Perner's original task, known as the Maxi task, was slightly different; the setup shown here, first formulated by Baron-Cohen et al., is now considered the 'standard' false belief task.

Performance on such tasks fascinates in several ways, for example:

- Normally developing young children fail to perform the task correctly until about 4 years of age. Failures are visually quite astonishing to witness, as children generally give the appearance of having understood the task and questions posed, and quite decidedly give 'incorrect' answers.
- Autistic, blind, and deaf children perform significantly worse on the task than their normally developing peers, often displaying a delay of years.

Although the experimental evidence garnered from children performing false belief tasks has been both rich and reasonably consistent, the literature still abounds with theories attempting to explain the observed behaviour. Nevertheless, certain commonalities can be found among many of the theories; Surian and Leslie (1999) describe two of the competing camps as follows:

Despite almost 15 years of research on preschoolers' understanding of false belief, the reasons for the shift in performance between the ages of 3 and 4 years remains controversial. Currently there are two major theoretical positions on the question. According to one view, children construct a succession of 'theories' about behaviour that come to implicate a theoretical construct—mental states. This process is thought to culminate in a major discontinuity in development: namely, a 'theory-shift' at around 4 years of age in which the child makes the discovery that mental states are really representations. For example, Perner claims that 'what young children cannot do is to represent that a proposition can be given a different truth value than the one it has (as assigned by children themselves)' (Perner, 1991, p. 243). The theory-shift view is shared by a number of different models and theoretical proposals which may differ substantially on other respects such as the relative emphasis given to specific mechanisms of change and developmental precursors (e.g. Perner, 1991; Gopnik, 1993; Wellman, 1990). The key finding for the theory-shift view is that 3- (but not 4-) year-old children fail on tasks that test for the understanding of false belief.

Another major view assumes more conceptual continuity in preschool development and postulates an early emerging, domain specific, innate competence (e.g. Baron-Cohen, 1995; Fodor, 1992; Leslie, 1987; Leslie & Thaiss, 1992). For example, the early emerging abilities to pretend and to understand pretense-in-others have been explained in terms of the maturation of a specialized 'theory of mind mechanism' (ToMM). According to the theory of ToMM, an early competence is embodied in a pre-structured representational system that provides the child with a domain-specific and probably modular learning device. The existence of such a specialized learning device explains how a preschool child is able to attend to and thus learn about mental states and their causes (Leslie, 1994). For 'continuity theories' such as ToMM, the failure of 3-year-olds on false belief tasks is to be understood in terms of performance limitations of various kinds which prevent or attenuate the deployment of an inherent competence. This pinpoints one major area of disagreement between these two positions: whether the failure of 3-year-old children on false belief tasks is the result of conceptual deficit or the result of a performance squeeze. [end Surian and Leslie]

To these two camps, I would add a third, more general camp, that could be characterised as non-'theory' based and also non-'module' based. This camp sees the development of Theory of Mind occurring via generic, largely domain-independent mechanisms such as learning, frequent social interaction with other agents such as siblings, etc.

As described above, one of the possible explanations for failure of 3-year-old children is the inherent complexity of the task itself. Children are required to follow the actions of two different characters, note that Sally did not see the switch, remember both the original location of the marble and its new location, and process one or more fairly complex questions. The two experiments described in this paper constituted an attempt to improve performance by removing some of this complexity.

## 2. Experimental Summary

Two highly dissimilar experiments, both designed as simplifications of the standard false belief task, explored the robustness of young children's apparent inability to successfully perform false belief tasks.

In the first experiment, children were shown an animated video repeatedly in order to become thoroughly familiar with the scenario and its characters before attempting to pass a false belief task.

In the second experiment, the children participated directly themselves, by hiding from a crèche leader and being interviewed while hiding.

Results of both experiments were highly consistent: neither variation succeeded in enabling children to successfully perform a false belief task.

## 3. Experiment 1

### 3.1 Method

The participants in this experiment were eleven children, recruited from an Amsterdam crèche serving primarily middle to upper socio-economic backgrounds. The children ranged in age from 2 years 11 months (2:11) to 4 years 0 months (4:0), with a mean age of 3 years 5 months (3:5). Children in this age range were chosen deliberately, as most children in this range are expected to fail the standard false belief task.

Each child was videotaped. Readers of this summary are strongly encouraged to view the videotapes as well.

The child was initially asked a series of questions (see questions 1-12 in section 5.1) intended to test her understanding of and ability to respond to sentences with

complement structures, future hypotheticals, counterfactuals, and true belief situations.

She was then shown an animated video three times, once in its entirety and twice focussing on the critical segment where a false belief scenario was generated. In that critical segment, Bob the Builder places his ladder against a shed and climbs up the ladder to start working on the roof. Meanwhile, Naughty Spud is looking for a way to get up in an apple tree to eat some apples. Naughty Spud sees Bob's ladder and sneakily takes it away, without Bob noticing. Bob then wants to take a lunch break and decides to come down off the roof…

During the final viewing, the video was paused at key moments to ask further questions (see questions 13-17 in section 5.1) involving true belief, false belief, and counterfactual situations.

Note: No attempt was made to ensure independence of this experiment from experiment 2, described below. Each child performed the same two experiments, in the same order (first experiment 1, then experiment 2).

### 3.2 Results

For a detailed summary of the children's answers to the questions posed during the experiment, see section 5.1.

Further, one filmed execution[2] of this experiment is available at:

http://staff.science.uva.nl/~dwood/fb2002/exp1_ge.wmv[3]

### 3.3 Discussion

It is clear from section 5.1, question 16, that the children were consistently unable to successfully perform this false belief task, thereby strongly confirming the results from more standard false belief tasks.

One of the primary differences between this false belief task and more standard ones was the attempt to thoroughly familiarise subjects with the scenario and its characters before attempting to pass a false belief task, in order to reduce the processing load of keeping track of various characters and events in real time. Although this seems to have succeeded admirably (even now, weeks later, tested children run up to me at the crèche and say "Naughty Spud took the ladder, didn't he?" ;-)), it also appears to have caused some related problems, notably that in subsequent viewings children were

[2] The subject is the son of the author, but was a highly typical subject. Please contact David Wood (dwood@inter.nl.net) if you are interested in viewing other video segments from these experiments. They have not been placed on the Internet because parents were not explicitly asked for permission to do so.

[3] This video is in Windows Media Format, and requires Windows Media Player 6.4 or later. It is also very large (24 MB). For optimum viewing, it is probably wisest to download the entire video to your client and then view it locally (volume should be turned up quite high as well). If you are unable to view it, please contact David Wood.

unable to keep track of the story's timeline and inhibit the use of information they had acquired from previous viewings but that was still in the future in the current viewing. See for example the answers of Fi. and Ge. to the **true** belief question "Where does Bob think that the ladder is?" (section 5.1, question 14). Both say "By the tree", even though at this point in the story the ladder has **never** been by the tree. As a further example of this, at the point where the child is asked the critical false belief question (section 5.1, question 16), it is not visually clear to the child that the ladder has in fact been placed by the apple tree. So at that point the two competing answers to the false belief question **should** be "By the shed." vs. "Taken away to some unknown destination by Spud.". This technique of reducing the salience or increasing the answering complexity of the 'incorrect' answer has been shown in other false belief experiments to improve performance (Wellman, Cross, & Watson, 2001), but in this experiment children were able to make use of previously acquired information that provided an equally salient alternative.

# 4. Experiment 2

## 4.1 Method

The participants in this experiment were the same as those in experiment 1, described above.

Each child was videotaped. Readers of this summary are strongly encouraged to view the videotapes as well.

The child was initially seated on a sofa, together with a crèche leader and an experimenter. The crèche leader then left, saying that she would be right back. Experimenter proposed to child to play hide-and-seek, hiding in a nearby tent. Each child enthusiastically agreed. Once in the tent, the child was asked the questions in section 5.2.

Note: No attempt was made to ensure independence of this experiment from experiment 1, described above. Each child performed the same two experiments, in the same order (first experiment 1, then experiment 2).

## 4.2 Results

For a detailed summary of the children's answers to the questions posed during the experiment, see section 5.2.

Further, one filmed execution of this experiment is available at:

http://staff.science.uva.nl/~dwood/fb2002/exp2_ge.wmv

### 4.3 Discussion

It is clear from section 5.2, questions 2 and 3, that the children were consistently unable to successfully perform this false belief task, thereby strongly confirming the results from more standard false belief tasks.

This false belief task differed from more standard tasks in some dramatic ways. First, the scenario was extremely short and simple. Second, the child herself performed one of the roles. Third, all actors were real people, not dolls or video characters. Nevertheless, the results show an astonishing consistency.

One experimental shortcoming in this experiment is that we only asked false belief questions in a complementation-based form, i.e.:

Where does Debbie/Tamara **think that** we are?

and

Does Debbie/Tamara **think that** we're here (in the tent), or there (on the sofa)?

Asking questions of the form:

Where is Debbie/Tamara going to look for us?

or

Where is Debbie/Tamara going to look for us first?

might have given different results. (See Siegal and Beattie (1991) for an example of how minor modifications to question wording may produce dramatic perfomance changes.)

Another possibly interesting minor variation on this experiment would be to replace the adult seeker with a child seeker, to test the prima facie unlikely hypothesis that the child subject believes that adults are omniscient whereas peers are not, and therefore the child believes that she is giving a trivially correct answer to a true belief task instead of an incorrect answer to a false belief task.

## 5. General Discussion

Communicating by means of natural language with children of mean age 3:5 is an experimental minefield. Given that comprehension still significantly outstrips production capability at that age, it's difficult to verify whether a question is being understood properly, or even heard properly. Standard ways of ensuring this with older children and adults, such as using written language, asking subjects to repeat what they were just asked, or asking subjects whether they fully understood the question, simply aren't available or reliable. This makes approaches to Theory of Mind testing such as that taken by O'Neill (1996) particularly interesting, in that they avoid a heavy dependency on natural language skills. O'Neill tested children's

awareness of the knowledge states of their parents by placing a toy in a for the child inaccessible position, with the parent either present or absent, and was able to show that the child's subsequent form of communication with the parent regarding the toy was indeed dependent on whether the parent had been present or absent when the toy was made inaccessible. (But even in that study O'Neill bumped up against natural language limitations, for example when asking control questions to ascertain whether the child remembered where the toy actually was. Also, when testing 2-year-old subjects, much of their verbal production was unintelligible to both experimenters and parents.)

To take a tangible example of the difficulties surrounding natural language communication, were our test subjects able to understand the complementation-based questions they were asked? Unfortunately it remains difficult to say. At one extreme, Fl. (3:11) even produced one sentence using complementation (section 5.1, question 16), but this was in fact the only such sentence produced during about 100 minutes of video footage! Fl. was incidentally also the only child in both experiments who showed initial signs of successfully performing the task. At the other end of the scale, in response to the question "Where does Debbie (not present) think that you are?", both Di. (2:11) and Ge. (2:11) answered "I'm upstairs.", which would seem to indicate they were answering the stripped down question "Where are you?". Most other children gave only the one word answer "Upstairs.", which leaves open exactly which question they were answering.

Another ambiguous linguistic/pragmatic area arose when the children were confronted with an unknown word or didn't know an answer. See for example the sometimes hilarious exchanges around the 'teriyaki' questions, where several children cheerily and emphatically affirmed that they liked teriyaki, only to immediately concede that they didn't know what it was.

One question that clearly seemed to overwhelm the children was the counterfactual question "If Spud hadn't taken the ladder, where would the ladder be?". All 11 children answered incorrectly, and many looked visibly baffled while the question was being posed. Some children even attempted to answer at the comma pause, as if they'd processed the if clause portion as a full-fledged question.

One other interesting observation (from Michiel van Lambalgen): Deixis (pointing) seemed to be an important means of communication. Whenever a question could be answered by pointing, e.g., "Where is the ladder now?", children tended to do so, and would only answer verbally if further prompted to do so.

Returning briefly to the key differentiators distinguishing the various 'camps' discussed at the beginning of this paper, based on our and others' experimental evidence, what do I believe underlies this fascinating failure of children to successfully perform false belief tasks:

- Rapid Acquisition of a 'Theory of Mind' Around the 4th Birthday? I find this "theory" theory so vague as to be almost impossible to evaluate. O'Neill's findings would seem to contradict it. As further counterevidence, Gergely,

Nádasdy, Csibra, and Biró (1995) show in a fascinating experiment that even 1-year-old children attribute rational intentions to agents and are surprised if such agents stop acting in accordance with previously displayed intentions.

- Immaturity of a Theory of Mind 'module' in the brain? Unlikely. The simple fact that all of autism, blindness, and deafness significantly affect false belief task processing, while affecting highly diverse regions of the brain, would seem to argue against a highly localised and specialised Theory of Mind module.
- General Task Complexity and Processing Demands? Yes, at least partially. Both the linguistic complexity of the questions asked, as well as (at least in experiment 1) simply keeping track of the story are placing heavy processing demands on the children and leading to confusion and errors.
- Insufficient Linguistic Skill? Yes, although the critical question here remains whether increased Theory of Mind drives increased linguistic competence, whether the causality is the other way around, or whether the two capabilities simply co-develop.
- Insufficient Social Interaction? Yes. A number of interesting studies, such as Perner, Ruffman, & Leekham (1994), have shown that the presence of siblings is a substantial predictor of false belief task performance. This seems to be a very strong argument that Theory of Mind is something we acquire through learning, and that that learning occurs through interaction with others.

Nevertheless, all of this remains nothing more than unproven hunches and speculation; how can things be further clarified? Our experiments shed no new light, but I believe they were a step in the right direction, an attempt to, in the words of a certain famous physicist, "make everything as simple as possible, but no simpler". We just didn't go far enough.

I argue that the fundamental shortcoming is with the false belief task itself, and that the task needs to be fundamentally simplified in order to become of any scientific use. The 'standard' false belief task requires the participation of far too many cognitive abilities to ever permit concrete, detailed conclusions to be drawn about anything. It's a fancy parlour trick, with results that make the guests shake their heads in amazement, but that's about all it is. (To be fair, it's apparently clinically effective in determining whether a 10-year-old child is autistic, but this sort of 'broad brush' diagnosis is a far cry from being able to make detailed predictions about Theory of Mind capabilities.)

So what can be simplified? I see at least two possibilities: the task itself, and the subjects. Has anyone tried to take steps in these directions? Yes, to a certain degree. Some scholars have been studying false belief task performance in non-human animals. Here, Susan Hurley summarises two such efforts:

One version of a nonverbal mind-reading test uses a hider/communicator paradigm, and has been applied to children and to chimps (Call et al, 1999, 2000; work is in

progress with bottlenose dolphins). In this paradigm, the subject perceives two opaque boxes, which are then hidden from her view. The hider then proceeds to hide a reward in one of the two boxes, while the communicator watches. The subject cannot see which box the hider puts the reward in. But the subject can see that the communicator can see which box the hider puts the reward in. The barrier is removed, and the subject is then allowed to choose between the two opaque boxes, still unable to see which contains the reward. The communicator truthfully indicates to the subject which box the reward is in. The subject learns to choose the box the communicator has indicated in order to obtain the reward.

In the critical trials, the procedure is altered as follows. After the hider has put the reward in one of the boxes, the communicator leaves the scene. The barrier is removed so that the subject can see the boxes, though not which box contains the reward. While the subject watches, the hider switches the positions of the boxes. The communicator then returns, and indicates the box not containing the reward to the subject, since this box is now in the position of the box into which the communicator saw the reward placed. The subject can choose either the box indicated by the communicator, or the other box. The correct response is to choose the other box, since the communicator did not see that the boxes were switched and thus has a false belief about which contains the reward.

When this nonverbal test of mind-reading ability is applied to children of varying ages, the results are strongly correlated with the results of verbal false belief tests. In general, children under 4 make in wrong response, and select the box the communicator indicates in the false belief trials as well as the control trials. Children over 5 make the correct response-- they select the box not indicated by the communicator when the boxes are switched, though they select the box indicated by the communicator in the control trials. Chimps fail profoundly in the false belief trials.

If this test is accepted as an indication of the ability to reason about the mental states of others, chimps appear to lack such ability. However, these results contrast with results for chimps of a different non-verbal mind-reading test, suggesting that the ability to reason about the mental states of others may be context-dependent.

This second version of a nonverbal mind-reading test uses a dominant/subordinate paradigm (Hare et al 2000, in press). The dominant and subordinate chimps compete for food. In some conditions the dominants had not seen the food hidden, or food they had seen hidden was moved to a different location when they were not watching (whereas in control conditions they saw the food being hidden or moved). At the same time, subordinates always saw the entire baiting procedure and could monitor the visual access of their dominant competitor as well. The results are that subordinates more often approach and obtain the food that dominants have not seen hidden or moved. Similarly, the subordinate gets more food when a new dominant chimp is substituted for the one who saw the baiting. This suggests a kind of mind-reading ability on the part of chimps: that subordinates are rationally sensitive to what dominants did or did not see during baiting.

How should the apparent ability of chimps to reason about the mental states of others in the dominant/subordinate paradigm be reconciled with the apparent lack of this

ability in the hider/communicator paradigm? It is too soon to say with any confidence. An empirical speculation of interest here, however, is that the hider/communicator paradigm provides a context in which there is cooperation over finding food, while the dominant/subordinate paradigm provides a context in which there is competition over finding food. It may be natural for chimps to compete over food in such a way that their ability to reason about the mental states of others is tuned to competitive practical contexts rather than cooperative ones. This provides another possible illustration of how practical reasons might be context-bound, and fail to have full conceptual generality. [end Hurley]

Nevertheless, although these experimental designs do eliminate one or more of the major shortcomings of the standard false belief task, such as a heavy dependence on natural language, these designs still seem **far** too complex and vague to give useful results. So what might a simpler but more powerful false belief task look like? As a starting point, these would seem to be some desirable characteristics:

- Avoid reliance on natural language.
- Require as few participants and steps as possible.
- Use conspecifics or peers if possible. In other words, don't have chimps attempting to predict the actions of another species, or human babies attempting to predict the actions of adult humans.
- Avoid rote learning (such as that required in the Call et al. experiment described above).

Based on these criteria, here's a basic design for a false belief task that could be used to evaluate **any** intentional agent, human or otherwise, including robots:

- Like all the other experiments described so far, set up two stations, e.g., boxes in which something could be concealed. Ensure that the boxes are physically distant from one other.
- Take two agents, and generate a strong but differing intention in each. For example, take two chimps and allow one, agent 1, to become extremely hungry and the other, agent 2, extremely thirsty. Or two robots, one programmed to seek a ball, while the other seeks physical contact with other robots.
- While agents 1 and 2 are present, place resolution of agent 1's intention, e.g., food, in one of the boxes.
- While agents 1 and 2 are present, attach resolution of agent 2's intention, e.g., drink, to agent 1.
- Remove agent 1 from the location of the two boxes so that it can't witness a box swap.
- Swap boxes.
- Begin monitoring the activity of agent 2, in particular where it waits for agent 1.

- Release agent 1 so it can return to the location of the boxes.

The point of the whole exercise is to generate in agent 2 a compelling desire to correctly predict the movements of agent 1 and move to meet agent 1 as quickly as possible, and then monitor how successfully agent 2 accomplishes this task.

This proposed false belief task is not more than a sketchy outline, and could probably be simplified even further. Again, the larger argument I'm trying to make is that in the whole experimental Theory of Mind-domain, we need more Gergely-style "close to the hardware" studies, and fewer "cute kids making cute mistakes" studies. Otherwise we'll never break out of this cycle of vague studies permitting 1001 conclusions, everybody picking their favorite, and then tenaciously defending it for another 20 years…

## 6. Acknowledgments

The experiments reported here were conducted at the Villa Kakelbont crèche in Amsterdam. I would like to thank the parents of all children who participated, and in particular thank Jessica, Debbie, and Tamara of the Villa for their enthusiastic cooperation and participation!

# 7. Appendix: Detailed Experimental Results

## 7.1 Experiment 1: Individual Answers to Questions Posed

| Question | Mi. (3:2) | Ba. (4:0) | Li. (3:8) | Di. (2:11) | Jo. (3:5) | Fi. (3:11) | Fl. (3:11) | Gu. (3:9) | Ch. (3:1) | Ge. (2:11) | Ky. (3:3) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | | | |
| (Prior to viewing video.) | | | | | | | | | | | |
| | | | | | | | | | | | |
| 1) Do you like cookies? | Y | Y | N | Y | Y | Y | Y | Y | Y | Y | Y |
| 2) Do you think that D/T (present) likes cookies? | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y | Y |
| 3) Do you like coffee? | N | N | N | Y | Y | Y | Y | N | Y | When I'm big. | N |
| 4) Do you think that D/T (present) likes coffee? | Y | N | Y | Y | Y | Y | Y | Y | Y | (Asks:) Do you like coffee? | Y |
| 5) If you were really big, would you like coffee? | N | Y | N | Y | Y | Y | Y | N | Y | Yes, very much. | Y |
| 6) If D/T were really small, would D/T like coffee? | N | N | N | Y | N | N | Y | Y | Y | A little bit. | Y |
| 7) Do you like teriyaki? | | | | Y | Y | Y | Y | N | Y | N | (No ans.) |
| 8) Do you know what teriyaki is? | | | | Y | N | Y | N | N | N | Y | (No ans.) |
| 9) What is teriyaki? | | | | (No ans.) | | A kind of juice. | | | | I don't know. That's a difficult word for me. | |
| 10) Where do you think that Debbie (not present) is? | Outside. | Dwnstrs. | Upstairs | Kitchen. | Dwnstrs. | Dwnstrs. | Dwnstrs. | Dwnstrs. | Outside. | Dwnstrs. | Dwnstrs. |
| 11) Where does Debbie (not present) think that you are? | Don't Know. | Upstairs. | Upstairs | I'm upstairs. | Debbie doesn't know that very well. | Upstairs. | Upstairs. | Upstairs. | (No ans.) | I'm upstairs. | (No ans.) |
| 12) Do you think that you'll come back to play tomorrow? | Y | Y | N | Y | Y | Y | Y | Y | Y | Y | Y |
| | | | | | | | | | | | |
| (Ladder still leaning against house.) | | | | | | | | | | | |
| | | | | | | | | | | | |

| Question | Mi. (3:2) | Ba. (4:0) | Li. (3:8) | Di. (2:11) | Jo. (3:5) | Fi. (3:11) | Fl. (3:11) | Gu. (3:9) | Ch. (3:1) | Ge. (2:11) | Ky. (3:3) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 13) Where is the ladder now? | In the tree. | Edge of the roof. | By the house. | By Bob the Builder. | Against the roof. | Against the roof. | Against the shed. | Against the roof. | I don't know. | By the house. | (Indicates only by pointing and touching TV.) |
| 14) Where does Bob think that the ladder is? | In the house. | There. By the house. | By the house. | (No ans.). | Bob stops. | By the tree. | Against the wall. | Against the wall. | I don't know. | By the tree. | (No ans.) |
| | | | | | | | | | | | |
| (Spud takes ladder away.) | | | | | | | | | | | |
| | | | | | | | | | | | |
| 15) Did Bob see that Spud took the ladder? | | N | Y (Corr.) | N | Y (Corr.) | Y (Corr.) | N | N | Y (Corr.) | N | (No ans.) |
| | | | | | | | | | | | |
| (Ladder is no longer visible.) | | | | | | | | | | | |
| | | | | | | | | | | | |
| 16) Where does Bob think that the ladder is? | Tree. | By the tree. | Spud has it. | By the tree. | The ladder is gone. By the tree. | By the tree. | Spud took it. (Question repeated.) He thinks that it's fallen. | Gone. | I don't know. | In the mud. | (No ans.) |
| 17) If Spud hadn't taken the ladder, where would the ladder be? | In the tree. | By the tree. | By the tree. | (No ans.) | By the tree. | By the tree. | Fallen. | By Spud. | I don't know. | I don't know. | By Spud. |

## 7.2 Experiment 2: Individual Answers to Questions Posed

| Question | Mi. (3:2) | Ba. (4:0) | Li. (3:8) | Di. (2:11) | Jo. (3:5) | Fi. (3:11) | Fl. (3:11) | Gu. (3:9) | Ch. (3:1) | Ge. (2:11) | Ky. (3:3) |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1) Did D/T see us go and hide? | | N | N | N | N | N | N | N | N | N | Y (Corr.) |
| 2) Where does D/T think that we are? | Hiding. | In the tent. | In the tent. | In the tent. | I'm in the tent. | In the tent. | Here. | Here. | Here. | In the tent. | Upstairs. |
| 3) Does D/T think that we're here (in the tent), or there (on the sofa)? | Here. | Here. | Here. | Here in the tent. | (No ans.) | Here. | There. | Here. | Here. | Here. | In the tent. |

## 7.3 Notes on Individual Answers to Experiments 1 and 2

- Pragmatic information, not spoken, is contained in parentheses.
- 'Y' indicates either a verbal affirmative or decisive nodding. 'N' indicates either a verbal negative or decisive shaking of the head.
- D/T represents one of Debbie or Tamara. These two crèche leaders participated in the experiment, alternating their roles. One was always present, one was always absent.
- A blank table cell means no question was asked. A question that was posed but not answered is indicated by (No ans.).
- Testing took place in a playroom above the main crèche area, hence the references to upstairs and downstairs.
- (Corr.) indicates that the child gave an incorrect answer and was explicitly corrected by the experimenter.

## 8. References

Bloom, P. (2000). How Children Learn the Meanings of Words. MIT Press, Cambridge, Mass.

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

Bowerman, M. (1986). First Steps in Acquiring Conditionals. Traugott / ter Meulen / Reilly / Ferguson (Eds.), 285-307.

Breheny, R. (2001). Meaning, Communication and Theory of Mind. *Proceedings of 23rd Annual Conference of the Cognitive Science Society, 2001*.

Garfield, J., Peterson, C., & Perry, T. (2001). Social Cognition, Language Acquisition and the Development of Theory of Mind. *Mind and Language, Volume 16, Number 5,* 494-541.

Gergely, G., Nádasdy, Z., Csibra, G., & Biró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition, 56*, 165-193.

German, T., & Leslie, A. (2001). Children's inferences from 'knowing' to 'pretending' and 'believing'. *British Journal of Developmental Psychology (2001), 19*, 59–83.

German, T., & Leslie, A. (2000). Attending to and learning about mental states. In P. Mitchell & K. Riggs (Eds.), *Reasoning and the mind*, 227–252.

Hurley, S. (2001). The Space of Reasons vs. the Space of Inference: Reply to Noë.

Karmiloff, K., & Karmiloff-Smith, A. (2001). *Pathways to Language*.

Leslie, A. (2000). "Theory of Mind" as a Mechanism of Selective Attention. *THE NEW COGNITIVE NEUROSCIENCES*, 1235-1247.

O'Neill, D.K. (1996). Two year-old children's sensitivity to parent's knowledge state when making requests. *Child Development, Volume 67*, 659-677.

Perner, J., Leekam, S., Myers, D., Davis, S., & Odgers, N. (In press.) Misrepresentation and referential confusion: Children's difficulty with false beliefs and outdated photographs. *British Journal of Developmental Psychology*.

Scholl, B., & Leslie, A. (2001). Minds, Modules, and Meta-Analysis. *Child Development, Volume 72, Number 3*, 696-701.

Surian, L., & Leslie, A. (1999). Competence and performance in false belief understanding: A comparison of autistic and three-year-old children. *British Journal of Developmental Psychology, 17*, 141–155.