

## Children's Difficulties Handling Dual Identity

I. A. Apperly and E. J. Robinson

*University of Birmingham, Birmingham, United Kingdom*

Thirty-nine 6-year-old children participated in a longitudinal study using tasks that required handling of dual identity. Pre- and posttest sessions employed tasks involving a protagonist who was partially informed about an object or person; for example, he knew an item as a ball but not as a present. Children who judged correctly that the protagonist did not know the ball was a present (thereby demonstrating some understanding of the consequences of limited information access), often judged incorrectly (1) that he knew that there was a present in the box, and (2) that he would search as if fully informed. Intervening sessions added contextual support and tried to clarify the experimenter's communicative intentions in a range of ways. Despite signs of general improvement, the distinctive pattern of errors persisted in every case. These findings go beyond previous studies of children's handling of limited information access, and are hard to accommodate within existing accounts of developing understanding of the mind. © 2001 Academic Press

*Key Words:* theory of mind; interpretation; partial knowledge; dual identity; intentionality; explaining action; context effects.

Children's understanding of the relationship between access to information and consequent knowledge has been investigated extensively. Children ages 3 years or even younger realize that a person who has no informative access to, say, the contents of a box, will remain ignorant of what is inside, even if they have touched, pushed, or stood upon the box (Pillow, 1989, 1993; Pratt & Bryant, 1990). Older children may remain confused about just what knowledge can be gained from what modality of access, as when 4-year-old children judge that they can identify color by feeling an object (O'Neill, Astington, & Flavell, 1992; Pillow, 1993; Robinson, Thomas, Parton, & Nye, 1997) and overestimate the knowledge to be gained from access to only an uninformative part of an object or picture or from an ambiguous utterance (e.g., Robinson & Whittaker, 1987; Robinson & Robinson, 1982; Ruffman & Olson, 1989; Taylor, 1988). By about 6 years, children have (a) grasped the relationship between a person's access to

The research reported was carried out in partial fulfilment of the degree of Ph.D. for the first author, who was supported by a studentship from the Biotechnology and Biological Sciences Research Council, UK. It was also supported financially by the Economic and Social Research Council, UK Grant R000237932.

Address correspondence and reprint requests to E. J. Robinson, Department of Psychology, Keele University, Keele, Staffordshire ST5 5BG, UK. E-mail: [e.j.robinson@keele.ac.uk](mailto:e.j.robinson@keele.ac.uk).

information and consequent knowledge and (b) recognize that people can interpret the same input in different ways.

However, this developmental picture is complicated by surprising inconsistencies in children's understanding of partial knowledge. Apperly and Robinson (1998) presented children with items with a dual identity, one visually apparent and one not, such as a dice that was also an eraser. A puppet protagonist (Heinz) was allowed to see the item (so found out that it was a dice) but was not allowed to feel (so did not know it was an eraser).

In one kind of task, the action prediction task, the dice/eraser was paired with an eraser of normal appearance. Having seen both items, Heinz observed as they were hidden in two locations. Children were then asked two questions: "Does Heinz know that the dice is a rubber?" ("rubber" is the standard British term for "eraser") and "Where will Heinz go and look to find a rubber?". Most 5- and 6-year-old children correctly denied that Heinz knew that the dice was a rubber, consistent with the previous finding that children this age understand the knowledge to be gained via a particular modality (O'Neill et al., 1992; Pillow, 1993; Robinson et al., 1997). Despite this, they were equally likely to predict that Heinz would go to either location to find a rubber: Despite acknowledging that Heinz did not know the dice was a rubber, many children predicted that he would go to the dice/rubber when he needed a rubber. Correctly answering the knowledge question first did not help children with their predictions of Heinz's search. Children's difficulty was not with the general form of the task: In a control condition in which Heinz was totally, rather than partially, ignorant of one of two pencil sharpeners, children predicted his search correctly. They were also successful on a standard deceptive box task (e.g., Gopnik & Astington, 1988).

Note that the protagonist's situation in this task differs crucially from that of a partially informed protagonist in appearance-reality tasks (e.g., Flavell, Green, & Flavell, 1986) or standard false belief tasks. Partial information access in such tasks leads to *misidentification* of an object (e.g., of a sponge as a rock, or of the pencils contained in a sweets tube as sweets) and thus a false belief. In contrast, Heinz did not have a false belief in our dual identity tasks since either label *correctly* identified the objects. Rather he had true but limited knowledge of the item's dual identity.

Apperly and Robinson's findings also differ importantly from those of Taylor (1988). Taylor found that despite correctly judging a small segment of a picture uninformative to a naïve viewer, children judged that he or she would know *other* information such as the name of the depicted character. In contrast, Apperly and Robinson found that children were inconsistent in their judgments regarding *the very same* information about the object in question. That is to say, children who judge correctly that Heinz does not know that the dice is a rubber appear to have already passed *both* of Taylor's tests by acknowledging not only that information access is restricted but that it will lead a specific lack of knowledge. Apperly and Robinson (1998) found that despite this apparent success, children predicted that Heinz would search as if fully informed. This suggests, contra Taylor, that chil-

dren's problems are not simply located in judging how much a protagonist knows following partial information access. It would appear that such judgments *can* be made with an understanding that still falls well short of adult-like competence.

A second task revealed that children who acknowledged that "Heinz does not know that the X is Y" (where X and Y co-refer) often failed to draw an obvious adult-like inference. In this task, a simple reformulation made the meaning of the knowledge question sensitive to the way reference to the object is secured. In the original question, "Does Heinz know that the *dice* is a rubber?" the referring expression ("dice") may be substituted for any other true description of the object, such as "green cube" or even a generic term like "thing," without altering the answer to the question (e.g., Does Heinz know that the green cube is a rubber? still has the correct answer "No"). However, in the apparently similar reformulation "Does Heinz know that there's a *rubber* in the box?"—to which adults consistently answer "No"—such a substitution yields "Does Heinz know that there's a green cube in the box?" [correct answer "Yes"]. Very few (14.5%) of the 5-year-olds tested answered the reformulated questions correctly (by saying "No"), although as before, many correctly said that Heinz did not know that the dice was a rubber: 69% across two tasks with different stimuli. These will be referred to as the "substitution-sensitive" and "substitution-insensitive" knowledge questions from now on.

From an adult perspective, success on the substitution-insensitive knowledge question ("Does Heinz know the X is a Y?") should be sufficient to answer the substitution-sensitive question ("Does Heinz know that there's a Y in the box?") or the action prediction correctly. Why was it not sufficient for the children in our samples? One possibility is that children fail to grasp the linguistic/pragmatic rules governing substitutability of co-referentials. This would affect their performance on the substitution-sensitive but not the substitution-insensitive question. This is the account offered by Russell (1992; see also Mitchell, 1996, and Perner, 1991) for results similar to ours from his (1987) study in which children answered questions about what a story character thought in a situation where he was partially informed.

However, such an account cannot explain why children's difficulty extends to Apperly and Robinson's action prediction task where there is no linguistic context that can be misinterpreted. But if children's errors are conceptual rather than interpretive, there are serious implications for accounts of children's developing mentalizing skills.

One possibility is that children who answer the substitution-insensitive knowledge question correctly have a genuine adult-like understanding of the protagonist's partial knowledge but fail to grasp its relevance for answering the action prediction or the substitution-sensitive knowledge questions. If so, the difference in difficulty could disappear if the surrounding context were more supportive and made the communicative aim of the experimenter clearer (Donaldson, 1978; Siegal, 1997, 1999). If this is correct, then Apperly and Robinson's (1998) results would inform developmental accounts focusing on how children come increas-

ingly to share an adult's perspective in an experimental procedure (or a similar adult-child interaction), but do not require us to rethink children's mentalistic understanding.

Another possibility, though, is that children's grasp of partial knowledge is not adult-like at the point when they only pass the substitution-insensitive knowledge question. If so, we need a new way of describing just how they do represent the protagonist's partial representation. Apperly and Robinson (1998) offer a tentative description that characterizes the test questions in terms of the representation of Heinz's knowledge on which the child must draw in order to answer correctly. Both the linguistic substitution-sensitive question and predicting action on the basis of partial knowledge require children to handle intensional contexts (e.g., Searle, 1983). Such contexts involve a meta-representation (i.e., a representation of Heinz's knowledge) whose conditions of satisfaction depend upon the features of the *primary representation* (e.g., Heinz's knowledge of the dice but not the eraser) rather than the features of the *object of that representation* (the actual dice/eraser). Handling intensional contexts requires children to represent Heinz's knowledge as held under some descriptions but not others. In contrast, success on the substitution-insensitive question might be achieved via a simpler processing route than adults employ, that would not support handling of intensional contexts. For this more radical option to be taken seriously, it is important to establish that the difference in difficulty between the substitution-insensitive question and the substitution-sensitive knowledge and action prediction questions is robust. That is the aim of the research presented here.

We conducted a longitudinal study involving a sample of 5- and 6-year-olds. Previous studies suggested that this is when children are just beginning to handle dual identity effectively. All children were tested in four sessions. In the first we administered a battery of pretests on the basis of which they were divided into matched experimental and control groups. One month later, the experimental group experienced the first set of supporting contexts involving *partial* knowledge of dual identity. Two weeks later they experienced the second set. During these sessions the control group experienced similar contexts involving *total* ignorance. Finally, 2 weeks later again we administered a battery of posttests to children in both experimental and control groups. The aim was to find out if the difference in difficulty remained between various questions that appear to tap the same understanding, when and after children experienced supporting contexts.<sup>1</sup>

<sup>1</sup>At first sight, the obvious way to test for change in the gap between substitution-sensitive and substitution-insensitive questions would be to have a matched control group without the supporting factors such as observation of action, and the thought bubble discussion of each protagonist's knowledge. However, Apperly and Robinson (1998) and Russell (1987) report that dual identity tasks can vary in difficulty. Thus we were concerned that the alterations to the stories required to omit the protagonist's incorrect action would result in a misleading "match," and thus unreliable absolute comparisons of difficulty. Instead we opted to focus on whether the difference between substitution-sensitive and substitution-insensitive questions remained significant despite our attempts to improve performance on substitution-sensitive questions.

## FIRST SESSION: PRETEST BATTERY

Children received a battery of four dual identity tasks taken from Apperly and Robinson (1998) and from Russell (1987). The main purpose of the pretesting was to enable us to create matched experimental and control groups for the subsequent sessions, but we had two subsidiary aims. First we were interested in examining the relationship between tasks within the pretest battery. Although the concrete tasks used by Apperly and Robinson (1998) were modeled on the story tasks used by Russell (1987), they were designed to be considerably simpler, with fewer embedded clauses in the intensional (substitution-sensitive) test questions: They asked, for example, "Does Heinz know there's a present in the box?" rather than "Can we say George was thinking 'I must find the man with curly red hair who stole my watch'?" Although, like Russell, the substitution-sensitive question form ("Does Heinz know there's a Y in the box?") was more difficult for children than the substitution-insensitive one ("Does Heinz know the X is a Y?") the tasks were not compared directly with Russell's

A final aim of the pretest was to check the surprising finding from Apperly and Robinson (1998) that children were not helped to answer the action prediction question "Where will Heinz go and look to find a rubber?" by being asked the substitution-insensitive question first: "Does Heinz know the dice is a rubber?" Since it shows *apparent* understanding, the substitution-insensitive question is an obvious way of trying to prompt children to answer not only the action prediction question but also the substitution-sensitive question correctly. Apperly and Robinson (1998) did not examine the latter relationship; thus it was important to check on the ineffectiveness of this prompt before it was rejected.

### *Method*

*Participants.* Forty-seven children were tested. Five failed to complete all four dual identity tasks because they refused to accept that the puppet could see or know anything. Results from 42 children (20 male, 22 female) ages between 5;7 and 6;6 ( $M = 6;0$ ,  $SD = 3.56$  months) entered the analyses. All children attended the same primary school with a lower middle class/upper working class catchment area in Birmingham, U.K., and spoke English as their first language.

*Materials.* In the two concrete dual identity tasks (Apperly & Robinson, 1998) we employed a dice/eraser which was paired with an eraser of normal appearance for the action prediction task, and a bouncy ball which was also described as a present. All items could be hidden in small tin boxes. The protagonist for these conditions was a puppet called Heinz. In the two story tasks we used vignettes in which a character came to have partial knowledge. One involved George and a thief with curly red hair who stole his watch while he was asleep, so George knew there was a thief but did not know what he looked like (from Russell, 1987). The second story was about a girl, Julie, who sees a man in a green coat approaching her house but does not know that he is the magician who is coming to entertain at her party.

*Procedure.* Children were all given a warm-up task to get them used to the experimenter and the use of vignettes and pictures. Children then received all four dual identity tasks in fully counterbalanced order. In the *action prediction task*, children were first allowed to look at the conventional eraser and the dice/eraser. They readily identified the objects by their visually apparent labels, a rubber (the label familiar to British children) and a dice. In order to agree on the second label for the dice, children were allowed to feel it, and it was demonstrated until the description was agreed. The descriptions were reiterated, emphasising the different modes of perceptual access “. . . so it looks like a dice but it feels like a rubber.” Each eraser was placed in a box, and children observed as Heinz looked inside the boxes, now placed apart on the table in front of the children. Their attention was drawn to the fact that Heinz had no other perceptual access. With the boxes closed, children were told that Heinz had a pencil drawing that he wanted to change, so he needed to find a rubber. Children were asked two questions in counterbalanced order between child: the substitution-insensitive form of knowledge question “Does Heinz know that the dice is a rubber?” and the action prediction question, “Where will Heinz go and look to find a rubber?” Note that for the action prediction question, children who make errors are equally likely to choose either location for Heinz to search. This means we cannot diagnose an individual response as right or wrong, but can only make comparisons on a group basis.

In the concrete *verbal task* involving a ball that was also a present, children first looked in the box at the ball and labeled it as such. To agree on the second label, children were conspiratorially informed “. . . this is going to be a present for Heinz, except we haven't told him and we don't want him to find out right now, so we'll have to whisper very quietly when he looks.” Children watched as Heinz looked inside the box and the fact that he had not been told about it also being a present was emphasized by conducting the task in whispers. Two knowledge questions followed: The substitution-insensitive question “Does Heinz know that the ball is a present?” and the substitution-sensitive question “Does Heinz know that there's a present in the box?”. The order of these questions was matched to the other task so that children got the substitution-insensitive question either first or second on both concrete tasks.

On the two *story tasks* children were read vignettes and asked questions about characters who had come to have incomplete knowledge. For example in the vignette taken from Russell (1987) George has his watch stolen but does not see the thief. Children were asked about what George knew or was thinking with substitution-insensitive and substitution-sensitive questions based upon those used by Russell: (substitution-insensitive question) “Can we say that George knows that the thief is a man with curly red hair?”; (substitution-sensitive question) “Can we say that George is thinking ‘I must find the man with curly red hair who stole my watch?’” These questions were asked in the same order as for the concrete tasks. Finally, children were asked a control question (in which they must allow that George knows *something*) “Can we say that George is thinking ‘I must find the thief who stole my watch?’”, as in Russell (1987).

## Results

First, we compared children's performance on the two concrete tasks with performance on the two story tasks. Table 1 shows the incidence of correct responses to the substitution-insensitive questions in each of the four tasks. Each child was given two scores out of two according to the number of substitution-insensitive knowledge questions answered correctly in the concrete and the story tasks. They were categorized by age according to whether they fell in the first or second 6-month half of the sample. An analysis of variance (ANOVA) was computed on the substitution-insensitive question scores with task type (concrete vs story) as a repeated measure, and age category and question order (substitution-insensitive first or substitution-sensitive first) as between-subject variables. There was a significant main effect of task type ( $F(1, 38) = 6.41, p = .016$ ). While the means for concrete and story tasks clearly were different ( $M = 1.81, 1.55; SD = .67, .45$ , respectively), Table 1 shows that this was mainly due to one of the story tasks being particularly hard and one concrete task being easier than the average. There were no significant effects of age or question order ( $ps > .12$ ) and all interactions were nonsignificant ( $ps > .19$ ). This suggests that the simplifications of Russell's original tasks did reduce children's problems with these questions somewhat.

Performance on the substitution-sensitive knowledge questions was complicated by the fact that the action task had a theoretical baseline of 50%: If children mistakenly think that Heinz is fully informed he could search in *either* location. Performance on the three tasks with comparable baselines was therefore considered on an individual basis (see Table 1 for scores). There was no difference in difficulty between the substitution-sensitive questions on the story task about George and the thief and the concrete ball/present task. However, the substitution-sensitive knowledge question in the story about Julie and the magician was significantly more difficult both than that in the story about George and the thief and

TABLE 1  
Number (Percentage) of Correct Answers to the Questions in the Pretest Battery of 4 Tasks

| Question type   | Concrete tasks |                   | Story tasks          |                        |
|---|----------------|-------------------|----------------------|------------------------|
|   | Verbal         | Action prediction | George and the thief | Julie and the magician |
| Substitution-insensitive; e.g., "Know X is Y?"                    | 36/42 (86%)    | 40/42 (95%)       | 37/42 (88%)          | 28/42 (67%)            |
| Substitution-sensitive; e.g., "Know Y in box?"/ "Where go for Y?" | 20/42 (48%)    | 24/42 (57%)       | 15/42 (36%)          | 6/42 (14%)             |
| Control   | n/a            | n/a               | 33/42 (79%)          | 36/42 (86%)            |

that in the ball/present task (both  $ps < .006$  by sign test). Russell (1987) also found differences in difficulty between his stories, so our results are consistent with his in this respect. Clearly the particular characteristics of the vignettes influence the difficulty of the substitution-sensitive knowledge question, and it seems not *necessarily* to be the case that the complex question form used by Russell, "Can we say . . . ?" makes the substitution-sensitive questions on his tasks substantially harder than our concrete tasks. What is more important here is the relative difficulty within each task or story of the substitution-insensitive and substitution-sensitive knowledge questions. For all four tasks (including action prediction), children were significantly more likely to answer the substitution-insensitive knowledge question correctly than the substitution-sensitive knowledge/action prediction question (all  $ps < .0002$  by sign test).

Next we examined the effect on the substitution-sensitive knowledge questions and the action prediction question of being asked the substitution-insensitive knowledge question first. Apperly and Robinson (1998) found no effect of order in their action prediction task, but did not examine order effects in the purely verbal task. As above, each child was given one combined score for the concrete tasks and one for the story tasks. An ANOVA was computed on these scores with task type (concrete vs story) as a repeated measure, and age category and question order (substitution-insensitive first or substitution-sensitive first) as between-subject variables. There were significant main effects of task type ( $F(1, 38) = 17.4, p < .0001$ ) with concrete tasks being easier than story tasks ( $M = 1.05, .50; SD = .71, .70$ , respectively). (Recall though that the baseline for the action prediction question was 50%, so this comparison is misleading.) There was also a main effect of age ( $F(1, 38) = 6.53, p = .015$ ) with the younger children performing less well than the older children ( $M = 1.14, 2.00; SD = 1.17, .92$ , respectively). Importantly, order of presentation of the questions (i.e., substitution-insensitive first or substitution-sensitive first) had no significant effect ( $F(1, 38) = 1.04, p = .313$ ). There were no significant interactions between question order and age ( $F(1, 38) = 1.04, p = .313$ ), or between task type and either age or question order ( $F(1, 38) = .20; .00$ , both  $ps > .66$ ), respectively. These analyses show that children were not helped to answer the substitution-sensitive knowledge or action prediction questions correctly by being asked the substitution-insensitive knowledge question first. The implications of these surprising findings are considered in the final discussion.

Finally, children made some errors on the *control questions* of the story tasks (in contrast to perfect performance in Russell's 1987 study): 3 made two errors, 9 made one error, and 30 made no errors. Observation of the data suggested that there was no systematic relationship between these errors and performance on the foregoing substitution-insensitive and substitution-sensitive questions. A possible explanation is that, since in the current study the control questions always came last in the set of three, lapses in concentration would result in proportionately more errors than Russell found.

On the basis of their performance on the substitution-sensitive knowledge (and action prediction) questions in the four pretest tasks, children were divided into



two groups: An experimental "partial" group,  $N = 22$ , and a control "total" group,  $N = 20$  (see Table 4). As far as possible, equal numbers of children gaining each score from 0 to 4 were assigned to the two groups. Selection for each group was otherwise random. A one-way ANOVA based upon performance on the substitution-sensitive question revealed there to be no significant difference between groups ( $F(1, 41) = .0002, p > .99$ ). These groups had different experiences in the second and third sessions.

## SECOND SESSION: EXPLAINING ACTIONS

Bartsch (1998) found that 3- to 4-year-old children are better able to explain action on the basis of a false belief than to predict such action before it has occurred (see also Bartsch & Wellman, 1989; Robinson & Mitchell, 1995; but see Moses & Flavell, 1990; Wimmer & Mayringer, 1998). Two potential reasons for such facilitation are that (a) the incorrect action provides the child with additional clues to the protagonist's ignorance, over and above his or her lack of information access, and (b) following the incorrect action the child need not deal with a protagonist's conflicting desires and beliefs. Since these reasons apply equally to settings involving partial, rather than complete, ignorance, children might also find it easier to handle the resulting intensional contexts when they have seen incorrect action based on partial knowledge. If they do, they might or might not subsequently perform better on the standard substitution-sensitive knowledge question. If there were any such effects, we would also need to decide whether to interpret this as genuine new understanding, or simply a consequence of clarifying the task (e.g., Siegal, 1997).

Children in the two groups were read different sets of two vignettes with accompanying pictures, and were invited to explain a character's incorrect action. In each group, the action was performed in ignorance but the type of ignorance varied between group. In the experimental "partial" group it was partial since the protagonist only knew one of an item's two identities; in the control "total" group the ignorance was total (and the protagonist had a false belief). In all other respects the two groups had the same general experience with vignettes and the experimenter. For each, there was a subsequent discussion of the discrepancy between the character's knowledge using pictures of their thought bubbles followed by a re-posing of the explanation question. Wellman, Hollander, and Schult (1996) report that children considerably younger than those in our sample readily understand thought bubbles as a means of expressing someone's thoughts.

### *Method*

*Procedure.* Each child was read two illustrated vignettes with accompanying questions and discussion. In the experimental, *partial group* the stories involved a protagonist acting on the basis of partial knowledge. One story involved Todd:

"It is Todd's birthday and his Mummy and Daddy have arranged a party for him. Lots of his friends and people from the street where he lives have come to his house. Todd doesn't know, but his Mummy has asked Mr Jones from next

door to dress as a clown and do a show for the children. Look, there he is, getting ready to do the show for the children, and all the parents are in the other room talking. Just then Todd's auntie gives him a plate with some food and asks him to take it to Mr Jones. Todd walks all around, looking for Mr. Jones. He goes into all the different rooms and sees the clown getting ready for the show. But after a couple of minutes, he comes back saying that he can't find Mr Jones."

At the end of the story children were asked three questions in the following order:

(1) (action explanation) Why can't Todd find Mr. Jones?

(2) (substitution-sensitive knowledge question) Does Todd know that Mr. Jones is right over there in the corner?

(3) (substitution-insensitive knowledge question) Does Todd know that Mr. Jones is the clown?

Children were then asked "Shall we have a think about what Todd knows?" and were shown a picture of Todd with a thought bubble picturing Mr. Jones dressed normally. They were told, "look, he's thinking of what Mr. Jones looks like normally!" and were encouraged to describe his appearance. They were then asked, "Shall we look at what his Auntie is thinking?" and were shown a picture of Auntie with a thought bubble showing Mr. Jones half-way into his clown costume. They were told, "Look, she knows that Mr. Jones is dressed as a clown" and were encouraged to describe his appearance. Finally they were again asked

(4) Why can't Todd find Mr. Jones?

The questions were always asked in the same order, ensuring that the protagonist's behavior was explained in the same way for each child. The second story was about a girl whose bag was stolen while she was in a burger bar. She saw the thief disappearing out the door but did not see that he had a beard (see Appendix 1). The two vignettes were presented in counterbalanced order between children.

In the control *total group* the stories involved a protagonist acting on the basis of total ignorance of a change in a state of affairs (and hence a false belief). In one case the outdated belief was about the color of the walls of someone's room and in the other, the changed location of a chocolate cake (see Appendix 2). In all other respects, the stories, questions, and discussion of the protagonist's knowledge were similar to those experienced by the partial group.

### *Treatment of Results*

While a model answer to the action question about the clown might have been "He can't find Mr. Jones because he doesn't know that Mr. Jones is a clown" we felt that a strict criterion of this kind was unsatisfactory since it was in danger of missing cases where the mental state of the protagonist was pragmatically implied rather than explicitly mentioned. Instead, answers were deemed correct if they made appropriate mention of the protagonist's partial ignorance. In the above story, children asked "Why can't Todd find Mr. Jones" passed if they said "cos he's dressed up like a clown" and failed if they said, for example, "cos he's hiding" "cos he's lost" or "I don't know." For the second vignette, the passers

mentioned the protagonist's lack of perceptual access to the relevant feature: "Cos she didn't see his black beard." Allusion to nonspecific ignorance "she didn't know that he stole the bag" counted as a fail, along with failure to supply an answer. Eleven randomly selected sets of responses were blind-coded from the tape recordings by a second person (E.J.R.). There was only one discrepant decision and this was easily settled by discussion.

Similar criteria to those described above were used in when scoring the control total group's responses.

### Results

*Partial group.* From Tables 1 and 2 it appears that children were better at explaining action and answering a substitution-sensitive question in the current situations than they had been on the pretest. Since there was no group who received comparable substitution-sensitive and substitution-insensitive questions without supporting context, this apparent improvement could not be quantified. However, our real interest was in whether a supporting context would close the gap in difficulty between substitution-sensitive and substitution-insensitive questions. If performance improved but the gap remained, we would have succeeded only in showing that the variation in overall task difficulty observed by Apperly and Robinson (1998) and Russell (1987) could be manipulated deliberately. Thus, we were principally interested in whether, having witnessed and explained Todd's action (Q1: Why can't Todd find Mr. Jones?), children still found the substitution-sensitive knowledge question (Q2: Does Todd know that Mr. Jones is right over there in the corner?) harder than the substitution-insensitive one (Q3: Does Todd know that Mr. Jones is a clown?). It is clear that they did: There were 12 children who made errors on Q1 (action explanation) and not Q3 (substitution-insensitive question), and 8 children who made errors on Q2 (substitution-sensitive question) and not Q3 (substitution-insensitive question), with none showing the opposite pattern, both  $ps < .008$  by sign test (see Table 2). There was no significant difference between action prediction and the substitution-sensitive question: Seven children made errors on Q1 and not Q2 with two showing the opposite pattern ( $p = .09$  by sign test). There were no effects of the order of task presentation on children's ability to explain action or answer a substitution-sensitive question for either task (all  $\chi^2(1, 22) < .26$ , all  $ps > .5$ ). Thus, while increasing the amount of contextual support appeared to improve overall performance, there was no evidence that this had any effect upon children's comprehension of the similar task that followed.

TABLE 2

Numbers of Children Scoring 0, 1, or 2 on the Four Questions of the Partial Tasks of Session 2

| Total score | Q1 Explaining action | Q2 (substitution-sensitive) | Q3 (substitution-insensitive) | Q4 Explaining action |
|-------------|----------------------|-----------------------------|-------------------------------|----------------------|
| 0           | 2                    | 0                           | 0                             | 2                    |
| 1           | 10                   | 8                           | 0                             | 3                    |
| 2           | 10                   | 14                          | 22                            | 17                   |

On Q4 (repetition of “Why can’t Todd find Mr. Jones?”) children found it easier to explain action after the other questions and the discussion of the knowledge states of the story protagonists. In an ANOVA with the first and second action explanation questions as repeated measures and age as a between subject variable, there was a main effect of question ( $F(1, 20) = 10.2, p = .005$ ) with performance on Q1 better than on Q4 ( $M = 1.36, 1.1.68; SD = .66, .65$ , respectively). Inspection of the data revealed this change to be due to seven children improving from a score of 1 on the first explanation question to a score of 2 when it was repeated. However, this could have been an artifact of the scoring criteria, by which children who had been encouraged simply to mention Mr. Jones’ appearance (regardless of who knew about it) might have seemed to improve. All age effects were nonsignificant.

*Total group.* For each question type children received a composite score of 0, 1, or 2. One child got both of the first action explanation questions (Q1) incorrect, 7 got one correct, and 11 got both correct. Errorless performance on Q2 suggested that successfully denying the protagonist’s knowledge was easier. However, inspection of the data revealed all but one of these errors to be on the decorating story, with children tending to say that Lisa told her friend that Arthur’s walls were orange “because she didn’t like them.” Such responses were scored as errors since they did not refer to her ignorance of the new color. But these answers may not have been inappropriate since there may be an adult reading of the situation that sees Lisa’s dislike of the color rather than her ignorance of the change as the relevant reason for her telling her friend that the walls are orange. Discussing Lisa’s knowledge with the experimenter did little to change children’s explanations of her action when they were again asked Q3: 7 children got one question correct and 12 got both correct.

In sum, as expected, children had little difficulty answering the knowledge question in the total group or the substitution-insensitive knowledge questions in the partial group. Seeing an inappropriate action caused some overall improvement in children’s ability to recognize the implications of the story protagonist’s partial access. However, seeing an inappropriate action based on partial knowledge did not remove the gap in difficulty between the substitution-insensitive and substitution-sensitive knowledge questions. Nor was there any evidence of transferred understanding from one scenario to another.

### THIRD SESSION: RHETORICAL QUESTIONS

Two weeks later each group was read further sets of two illustrated vignettes, and again the story protagonist engaged in incorrect behavior based on partial or complete ignorance. This time we tried to make the protagonist’s problem even more salient by having other story characters discuss his/her knowledge explicitly.

#### *Method*

*Procedure.* In the *partial group* there were three means by which children might be assisted in answering the substitution-sensitive knowledge questions.

First, the stories involved a protagonist acting incorrectly, given what the child knows, on the basis of partial knowledge. Second, the substitution-sensitive knowledge question was first posed rhetorically by one story protagonists before being addressed to the child; for example “Doesn’t he know that his present is in the bag?” Following the child’s answer, another story protagonist also responds to the question by contrasting the subject’s knowledge and ignorance of the object (e.g., “No, because he doesn’t know that the dinosaur is his present”), the aim being to clarify the meaning of the substitution-sensitive question without confronting a child’s (possibly incorrect) answer directly. Third, subsequent questions led the child through the protagonist’s information access, including a substitution-insensitive question about their knowledge, before re-posing the substitution-sensitive question.

One of the stories follows: Mary and her Mum are going to visit her Gran at her house. Her Gran has just bought a new china ornament in the shape of a rabbit. But you can also take the top off and use it as a pot and her Gran is keeping sweets in it. When she arrives, her Gran sends Mary to get some sweets, “They’re in the pot in the lounge” she says. Mary goes into the lounge, looking for a pot. She looks on the shelf and sees a clock and the china rabbit, and she looks on the table and sees the fruit bowl. Then she calls back to her Mum and her Gran saying, “I can’t find the sweets anywhere.” Her mum says to her Gran “Doesn’t she know that the pot is on the shelf?”

(Q1) (substitution-sensitive knowledge question) Does Mary know that the pot is on the shelf?

Her Gran says to her mum, “Well maybe she doesn’t know that the rabbit is a pot.”

(Q2) Did Mary see the rabbit?

(Q3) Can she tell just by looking that it’s a pot?

(Q4) (substitution-insensitive knowledge question) Does Mary know that the rabbit is a pot?

(Q5) (repeated substitution-sensitive knowledge question) Does Mary know that the pot is on the shelf?

The other story concerned a family Christmas shopping trip in which a protagonist sees his parents buying a toy dinosaur among a variety of other toys, but doesn’t realize that it’s a present for him (see Appendix 3).

Stories for the control *total group* involved a protagonist acting on the basis of a false belief (total ignorance). One involved the changed location of a handbag in a restaurant; the other, the location of some pens in a tin inside a cupboard. In all other respects they were similar to the partial group’s although there were necessarily fewer questions about information access (See Appendix 4).

Story pairs were presented in counterbalanced order between child. Questions were always asked in the same order.

### *Results*

Unlike the action explanation questions of session 2, all of the questions in this session had yes/no answers and could be scored simply.

*Partial group.* As before, our main interest was less in the level of overall performance than in whether the difference in difficulty between substitution-insensitive and substitution-sensitive knowledge questions remained in this context, which appears to make the protagonist's partial knowledge particularly salient. The other questions were designed only to lead children through the protagonists' partial information access (see Table 3 for summary of performance). ANOVAs were computed with substitution-sensitive questions (Q1 or Q5; Does Mary know that the pot is on the shelf?) and substitution-insensitive Q4 (Does Mary know that the rabbit is a pot?) as repeated measures and age category as a between subject variable. Substitution-sensitive Q1 (and Q5),  $M = 1.36$ ,  $SD = .85$  ( $M = 1.18$ ,  $SD = .73$ ) remained more difficult than substitution-insensitive Q4,  $M = 1.77$ ,  $SD = .53$  (Q1:  $F(1, 20) = 5.55$ ,  $p = .029$ ; Q5:  $F(1, 20) = 11.58$ ,  $p = .003$ ). Age was not significant in either comparison (all  $ps > .60$ ). In a similar analysis of variance comparing the two substitution-sensitive questions (Q1 and Q5) there were no significant differences (all  $ps > .47$ ). As before, order of task presentation did not affect children's ability to answer the first substitution-sensitive question (both  $\chi^2(1, 22) < 2.7$ ), again suggesting that there was no transfer of understanding from one task to another.

*Total group.* Two children had left the school by the time of testing, reducing the sample size to 18. Composite scores across the two tasks were calculated. Children made no errors judging whether the protagonist knew the critical piece of information (Q1). On the other key question (Q4: What does s/he think. . . .) one child got both incorrect, 3 got one correct, and 14 got both correct.

In sum, the difference in difficulty between the Partial group's substitution-sensitive and substitution-insensitive knowledge questions was not removed. This time, the explicit, structured discussion of the protagonists' information access and knowledge did not improve children's performance on the repeated question. This result is possibly more reliable than the opposite finding in the second session, where children were better at explaining a protagonist's incorrect action the second time, since in this case a substitution-sensitive knowledge question was repeated. Given the consistent evidence from the experiments of Apperly and Robinson (1998; Apperly, 1999) about the nature of children's errors on these questions, their answers could be scored with more certainty than the explanations in the second session. Despite improvements, these data also show chil-

TABLE 3  
Numbers of Children Scoring 0, 1, or 2 on the Five Questions in the Partial Tasks of Session 3

| Total score | Q1<br>(substitution-<br>sensitive) | Q2 (e.g., Did<br>Mary see<br>the rabbit?) | Q3 (Can she<br>tell just by<br>looking. . . ?) | Q4<br>(substitution-<br>insensitive) | Q5<br>(repeated<br>substitution-<br>sensitive) |
|-------------|------------------------------------|---|--|--------------------------------------|--|
| 0           | 5                                  | 1   | 1  | 1                                    | 4  |
| 1           | 4                                  | 3   | 6  | 3                                    | 10   |
| 2           | 13                                 | 18  | 15   | 18                                   | 8  |

dren's problems with intensional contexts to be remarkably robust, persisting even when they have just heard a very similar substitution-sensitive question answered by one of the story protagonists and insensitive to a previous, similar task.

#### FOURTH SESSION: POST-TEST BATTERY

##### *Method*

*Procedure.* As for the pretest, children were given four partial knowledge tasks in fully counterbalanced order. There were two new concrete tasks, which were acted out with Heinz the puppet. The action prediction task used a baby's ball containing a hidden bell that could therefore be labeled "ball" or "bell." This was paired with a small bell of normal appearance. The verbal task employed a ruler that was also an eraser. Children were introduced to the stimuli at the beginning of each task and the various labels agreed. They watched as Heinz looked but did not hear or feel the items. On the action prediction task children were asked an substitution-insensitive dual identity question ("Does Heinz know that the ball is a bell?") and an action prediction question ("Where will Heinz go and look to make some noise?") in counterbalanced order. On the verbal task children were asked substitution-insensitive ("Does Heinz know that the ruler is a rubber?") and substitution-sensitive ("Does Heinz know that there's a rubber in the pencil case?") questions about Heinz's knowledge. There were two new story tasks in which children were asked substitution-insensitive, substitution-sensitive, and control questions about vignettes. One involved a character who met a lady in a blue jumper at the paper shop, but did not know that she was a teacher and the other, a person seeking a new home who likes a white house but does not know it is the Mayor's house.<sup>2</sup> Substitution-insensitive and substitution-sensitive questions were asked in the same order as for the concrete tasks, and the control question (in which they must allow substitution) was always asked last.

##### *Results*

One child left the school before posttesting, so for partial group  $N = 21$ , for the total group,  $N = 18$ . Children received combined scores for both the substitution-insensitive and the substitution-sensitive questions, ranging from 0 (all incorrect) to 4 (all correct). Table 4 summarizes these data for the partial and total groups on the posttest and pretest. It is clear that there remains a large overall difference in difficulty between the substitution-insensitive and substitution-sensitive questions.

*Within posttest analyses.* As for the pretest, performance on the substitution-sensitive knowledge questions of the two story and one concrete task with comparable baselines was compared on an individual basis (see Table 5 for scores). There was no difference in difficulty between the two story tasks but both were

<sup>2</sup>This story was adapted from Russell (1987).

TABLE 4  
 Frequency (Percentage) of the Two Groups' Performance on Substitution-Insensitive and Substitution-Sensitive Questions of the Pre- and Posttest Battery

| Total score on substitution-insensitive question | Partial  |          | Total    |          |
|--|----------|----------|----------|----------|
|  | Pretest  | Posttest | Pretest  | Posttest |
| 0  | 0 (0%)   | 0 (0%)   | 0 (0%)   | 0 (0%)   |
| 1  | 2 (9%)   | 0 (0%)   | 1 (5%)   | 0 (0%)   |
| 2  | 3 (14%)  | 1 (5%)   | 1 (5%)   | 1 (6%)   |
| 3  | 3 (14%)  | 4 (19%)  | 7 (35%)  | 4 (22%)  |
| 4  | 14 (63%) | 16 (76%) | 11 (55%) | 13 (72%) |
| Total  | 22       | 21       | 20       | 18       |

  

| Total score on substitution-sensitive/action question | Partial |          | Total   |          |
|---|---------|----------|---------|----------|
|   | Pretest | Posttest | Pretest | Posttest |
| 0   | 4 (18%) | 1 (5%)   | 3 (15%) | 1 (6%)   |
| 1   | 8 (36%) | 5 (24%)  | 8 (40%) | 6 (33%)  |
| 2   | 5 (23%) | 8 (38%)  | 6 (30%) | 7 (39%)  |
| 3   | 4 (18%) | 3 (14%)  | 1 (5%)  | 2 (11%)  |
| 4   | 1 (5%)  | 4 (19%)  | 2 (10%) | 2 (11%)  |
| Total   | 22      | 21       | 20      | 18       |

harder than the verbal concrete task (both  $ps < .0001$  by sign test). There is no obvious reason for the good performance on the concrete verbal task, other than children's prior exposure to the dice/eraser, which may have led them to think that Heinz would not fall for the same trick again. However, even with this unusual performance, children's 100% performance on the substitution-insensitive question was significantly better ( $p = .016$  by sign test). There was no evidence of a difference between groups: Four of these errors were in the total group, three in the partial group. Performance on the substitution-insensitive question was also significantly better for the other three tasks.

Next we examined the effect on the substitution-sensitive knowledge questions and the action prediction question of being asked the substitution-insensitive knowledge question first. An ANOVA was computed with task type (concrete vs

TABLE 5  
 Number (Percentage) of Correct Answers to the Questions in the Posttest Battery of 4 Tasks.

| Question type            | Concrete tasks |                   | Story tasks          |                        |
|--------------------------|----------------|-------------------|----------------------|------------------------|
|                          | Verbal         | Action prediction | George and the thief | Julie and the magician |
| Substitution-insensitive | 39/39 (100%)   | 39/39 (100%)      | 34/39 (87%)          | 32/39 (82%)            |
| Substitution-sensitive   | 32/39 (82%)    | 21/39 (54%)       | 14/39 (36%)          | 13/39 (33%)            |
| Control                  | n/a            | n/a               | 34/39 (87%)          | 30/39 (77%)            |



story) as a repeated measure and age category, group (partial or total), and question order (substitution-insensitive first or substitution-sensitive first) as between-subject variables. There were no significant between-subject main effects (all  $ps > .14$ ). Within subjects, there was a significant effect of task type ( $F(1, 31) = 27.4, p < .0001$ ) with the concrete task easier than the story task ( $M = 1.36, .69; SD = .63, .80$ , respectively). There was also a significant interaction between task type and question order ( $F(1, 31) = 6.36, p = .017$ ). Examination of the data reveals this interaction to be primarily due to an effect in the story task performance with children who received the substitution-sensitive question first making more errors than those receiving the substitution-insensitive question first ( $M = .42, .95; SD = .77, .76$ , respectively). For the concrete tasks, there were no significant effects between the groups ( $M = 1.35, 1.37$ ), and all other effects, including the interactions with group, were also nonsignificant (all  $ps > .13$ ).

Performance on the *control* questions of the story tasks was similar to that in the pretest: 27 children got both correct, 10 made one error, and 2 made two. Examination of the data suggested that there was no systematic relationship between these errors and performance on the foregoing substitution-insensitive and substitution-sensitive questions.

*Comparison of pre- and posttest performance.* An ANOVA was computed with pre- and posttest scores on the substitution-insensitive and substitution-sensitive questions as repeated measures, and group (total vs partial) and age category as between-subject variables. There was a significant main effect of testing time—pre- vs posttest—( $F(1, 35) = 6.19, p = .018$ ) and it is clear from the data that this is due to superior performance on the posttest on both substitution-insensitive and substitution-sensitive questions ( $M = 3.69, 2.05; SD = .56, 1.12$ , respectively, versus  $M = 3.31, 1.64; SD = .95, 1.11$ , respectively, on the pretest). Importantly though, this occurred regardless of children's exposure on the foregoing studies: The interaction between testing time and group was nonsignificant ( $F(1, 35) = .61, p = .44$ ): The extent of change was no different between experimental and control groups. As might be expected, there was a large main effect of question type (substitution-insensitive vs substitution-sensitive):  $F(1, 35) = 138.2, p < .0001$  (see above for means and standard deviations). However, there were no significant interactions between question type and testing time or between question type, testing time, and group ( $F(1, 35) < .001, p = .988; F(1, 35) = .36, p = .55$ , respectively). Thus, it appears that the earlier sessions did not significantly decrease the difference between the substitution-insensitive and substitution-sensitive questions and in particular, sessions involving partial knowledge had no more long-term effect than those involving total knowledge.

In sum, no clear evidence was found for any *persisting* effect of the two earlier sessions on children's ability to handle intensional contexts as measured by the substitution-sensitive linguistic and partial knowledge action prediction questions.

## FINAL DISCUSSION AND CONCLUSIONS

The current study investigated the generality of children's problems with dual identity across a range of situations aimed to support good performance. In the pre- and posttest batteries, children completed a total of eight dual identity tasks and this substantial body of data was broadly consistent with the findings of Apperly and Robinson (1998; Apperly, 1999; Russell, 1987): Intensional contexts were in every case more difficult for children than questions about the same partial knowledge which had the substitution-insensitive form. Furthermore, it was generally the case that receiving an substitution-insensitive question first had no effect on children's success with substitution-sensitive, intensional questions.

Within both the second and third sessions in which we assessed the effects of supporting context, children often appeared better at answering substitution-sensitive questions about dual identity than they were on the pretest versions but importantly, were always better still on substitution-insensitive questions. In the third session one of the substitution-sensitive questions not only followed observation of a protagonist's (partially) ignorant behavior but also the answering of a similarly phrased substitution-sensitive question by one of the story characters, suggesting that the errors were indeed robust. Finally, there were no effects of the order of task presentation within sessions 2 and 3, suggesting that children were not learning anything that would generalize to another situation, even in the short term.

Comparing pre- and posttest performance, there seemed to be little or no effect of the intervening sessions, suggesting once more that children learned nothing persistent or generalizable from the supportive contexts.

Our results provide little support for the possibility arising from the views of Donaldson (1978) and Siegal (1997, 1999) that the difference in difficulty between the substitution-insensitive and substitution-sensitive knowledge questions is due to children's failure to grasp the experimenter's communicative intent with the substitution-sensitive (intensional) question. The difference remained even after a structured discussion of what each story protagonist knew of the information access that led to the partial knowledge of one protagonist. Furthermore, the absence of any improvement from the first task to the second suggests that children were not merely getting a better grasp of the experimenter's communicative intentions.

However, children did find it easier to explain action or answer a substitution-sensitive question after observing the incorrect behavior, than to answer the unsupported substitution-sensitive knowledge questions in the pre- and posttests. This seems consistent with performance or information processing difficulties similar to those invoked to account for errors on standard false belief tasks (e.g., Fodor, 1992; Leslie, 1987, 1994; Mitchell, 1996). Furthermore, it is plausible, as Moses and Flavell (1990) argue regarding false belief experiments (see also Fodor, 1992), that explaining action is easier than predicting action because the former case provides more evidence of the mental state which, in effect, can be read from the behavior.

However, as a sole explanation for children's errors in the dual identity tasks we believe that such accounts would be inadequate. First, they fail to explain why children remained significantly better on the substitution-insensitive question than on the substitution-sensitive questions, or the explanation of the protagonist's action. Second, they fail to explain how observing behavior might help children derive an answer to the substitution-sensitive questions, while apparently successful recollection of the character's knowledge in the substitution-insensitive question does not. An assumption of intact understanding that is masked by poor general performance is surely challenged by these differences between children's handling of situations that—to an adult at least—appear very similar in both content and complexity.

Russell's (1996) notion of executive competence—rather than performance—errors may resolve this problem. On this account there is no latent understanding to be frustrated by simple errors of performance because the understanding actually *consists in* children's ability to “. . . regulate their attention in such a way that they are able to think explicitly and at will about certain mental or abstract properties and processes such as beliefs . . . and logical relations” (ibid, p.210). It is perhaps tempting to view competence errors as the far end of a continuum of performance errors at which there are no further task-specific demands that could be removed to reveal children's understanding. But then we must ask what basis we have for positing an understanding that children cannot, even in principle, manifest. Russell denies that there is any such basis because on his view, understanding is grounded in the child's own agency. Thus, certain kinds of thought are simply absent from children who lack the requisite executive abilities.

The suggestion of qualitative change in children's thinking due to executive development gives Russell's account predictive potential. While purely performance-based accounts seem condemned to explain the form of children's errors with only *ad hoc* appeals to salience, a competence account—coupled with an adequate description of the representations controlled by the executive processes—could predict what tasks should be within reach of a child at a particular point in development. However, Russell's focus on the necessary role for agency in cognitive development leaves his account of the representational system that the agent develops and controls underspecified. What we need is a description of the differences between dual identity and other, apparently similar, mentalistic tasks that children find easier. The notion of changes in executive competence might then contribute importantly to an account of our findings as the reason for children's remarkable, and rationally inconsistent, pattern of errors.

We can outline some properties that such an account should have. First, to be compatible with an executive competence account, it should provide suitable content for explicit, willful manipulation. This requirement bears upon the distinction made by Dennett (1969) between personal level and subpersonal level explanations. A subpersonal level explanation concerns the nonconscious “behind the scenes machinery” underlying a particular aspect of behavior. Fodor's language of thought (Fodor, 1975) with its reliance upon syntactic symbol manipulation

exemplifies a representational account at the subpersonal level. In contrast, a personal level explanation involves entities one would normally attribute to a person: Our folk psychological notions of intention, reasons, and mental states, for example. It is clear that representations to be manipulated explicitly and willfully should be described at this personal level.

Second, it must provide some basis for distinguishing the demands of handling false beliefs and our substitution-insensitive question from mentalistic tasks like our substitution-sensitive and action prediction questions that children find much harder. Without this there is nothing on which an executive competence (or other general cognitive account) can achieve any purchase.

Finally, any representational description should take account of the fact that the referential and logical phenomena associated with mentalistic contexts are far from unique. Important similarities are found with modal contexts (concerning necessity and possibility), and when considering outdated, hypothetical and spatially displaced situations (e.g., Carnap, 1947; Dowty, Wall, & Peters, 1982; Fauconnier, 1985; Quine, 1953). But while this similarity has occasionally been acknowledged in the literature on children's developing mentalizing skills (e.g., Apperly & Robinson, 1998; Perner, 1991), there has been little exploration of how these domains might co-emerge or interact. The case for such an exploration has been made more compelling by recent empirical work suggesting links between children's performance on false belief and more general counterfactual and conditional reasoning tasks (e.g., Frye, Zelazo, & Palfai, 1995; Riggs, Peterson, Robinson, & Mitchell, 1998; Robinson & Beck, 2000). An adequate account of how these varied contexts are represented mentally could both clarify the nature of such links and pave the way for further research into how children's folk psychology continues to develop well after they have mastered false belief.

## APPENDIX 1

### Second Vignette from the "Partial" Group of Study 1

One day Claire was in MacDonald's. There she is at her table eating her burger, and look, that's her bag on the chair next to her. When she finished her food, she needed to go to the toilet, so she left her bag on the chair because she would only be a minute. But while she was away, a thief came and stole her bag. Look, the thief was a man with a big black beard! When Claire came back from the toilet, she saw a man disappearing out of the door with her bag. Stop! she shouted, but the man had already gone around the corner. Luckily, there was a policeman nearby and she ran over and said, "Quick, help me, a man has just stolen my bag!"

(Q1) Why didn't Claire tell the policeman that a man with a black beard had stolen her bag?

(Q2) Does Claire know that a man with a black beard stole her bag?

(Q3) Does Claire know that the thief was a man with a black beard?

(Q4) Why didn't Claire tell the policeman that a man with a black beard had stolen her bag?

## APPENDIX 2

## Vignettes from the "Total" Group of Intervention 1

(Vignette A) Arthur has just moved into his new house and he is showing his friend Lisa around. When she sees the walls of his sitting room painted orange, she says, "Oh, wouldn't they look a lot nicer if they were blue?" Arthur thinks she is probably right, so after she has left he rushes out to buy some blue paint and starts work painting the walls. He works very hard all afternoon, and finishes painting all the walls blue.

When Lisa got home after visiting Arthur, she was talking to a friend about Arthur's new house. "It's mostly very nice, but I don't like his sitting room. The walls are orange!"

Children were asked:

(Q1) (action explanation) Why did Lisa say that Arthur's sitting room walls were orange?

(Q2) (knowledge) Does Lisa know that Arthur's walls are blue now?

Children were then asked, "Shall we think about what Lisa knows?" and were shown a picture of Lisa with a thought bubble showing Arthur's walls painted orange. This was contrasted with a picture of the walls "as they are really." Children were then asked:

(Q3) Why did Lisa say that Arthur's walls were orange?

(Vignette B) Martin and Nina are having some cake to eat in their kitchen. Martin puts the cake away in the blue cupboard and then goes out to play. While he is outside, Nina is doing some cleaning and she takes the cake out of the blue cupboard. When she has finished cleaning she puts the cake away again, this time in the red cupboard. Here comes Martin, and he's hungry again after playing outside and he wants to eat some more cake. He goes to look in the blue cupboard.

(Q1) Why is Martin looking in the blue cupboard?

(Q2) Does Martin know that the cake is in the red cupboard?

(Q3) Why is Martin looking in the blue cupboard?

## APPENDIX 3

## Second Vignette from the "Partial" Group in Study 2

One day, not long before Christmas, Daniel and his mummy and daddy went shopping in a big toy shop. They had lots of things to buy because this year, all of Daniel's cousins were coming to stay, and so they had to buy presents for all of them. Look, their trolley is already quite full of toys isn't it? While he was away looking around, his mummy saw a toy dinosaur on the shelf and she said to daddy, "Daniel would really like this for his present." Daniel didn't see them or hear what they were saying. When they paid for the toys, Daniel saw the dinosaur amongst all the other toys, but he didn't ask whether any were for him.

In the car, Daniel jumped into the back seat, knocking over one of the bags of toys. His dad said to his mum, "He should be careful, does he know that his present is in the bag?"

(Q1) Does Daniel know that his present is in the bag?  
His mum said “No, he just *saw* the dinosaur, he can’t have guessed that it was his present.”

(Q2) Did Daniel see the dinosaur?

(Q3) Did Daniel hear that it was his present?

(Q4) Can he tell just by looking that it’s his present?

(Q5) Does Daniel know that the dinosaur is his present?

(Q6) Does he know that *his present* is in the bag?

## APPENDIX 4

### Vignettes from the “Total” Group of Study 2

(Vignette A) One day Claire was in MacDonaldis. There she is at her table eating her burger, and look, that’s her bag on the chair next to her. When she finished her food, she needed to go to the toilet, so she left her bag on the chair because she would only be a minute. But while she was away, the waiter noticed the bag left on the seat and took it away to keep it safe. When Claire came back from the toilet, she saw that her bag had disappeared and shouted, “Quick, help me, a thief has just stolen my bag!” Behind the counter, the cook said to waiter, “Doesn’t she know that we have it here?”

(Q1) Does Claire know that they have her bag behind the counter?

The waiter said “Maybe she doesn’t know we just took it away to keep it safe”

(Q2) Did Claire see the waiter take her bag?

(Q3) Can she tell who took her bag?

(Q4) Who does Claire think took her bag?

(Vignette B) Sam and his mum are staying at his gran’s house. One afternoon, Sam wants to do some drawing so he asks his gran if he can borrow some pens. He doesn’t know that his gran keeps all her pens in an old biscuit tin. His gran says “of course you can borrow some pens, you’ll find them in the cupboard in the lounge.”

Sam goes to look in the cupboard and he sees lots of books some boxes with games in them and an old biscuit tin. He calls back to his Gran saying “I don’t think the pens are in the cupboard.”

His gran says to his mum, “Doesn’t he know that the pens are in the biscuit tin?”

(Q1) Does Sam know that the pens are in the biscuit tin?

His mum says, “Maybe he doesn’t know.”

(Q2) Did Sam see the biscuit tin?

(Q3) Can he tell just by looking that it’s got pens inside?

(Q4) What does Sam think is in the biscuit tin?

## REFERENCES

- Apperly, I. A. (1999). *Children’s mental representation of referential relations: Representational partitioning and “theory of mind.”* Unpublished Ph.D. thesis, University of Birmingham: Birmingham, UK.
- Apperly, I. A., & Robinson, E. J. (1998). Children’s mental representation of referential relations. *Cognition*, **63**(3), 287–309.

- Bartsch, K. (1998). False belief prediction and explanation: Which develops first and why it matters. *International Journal of Behavioural Development*, **22**(2), 423–428.
- Bartsch, K., & Wellman, H. (1989). Young children's attribution of action to beliefs and desires. *Child Development*, **60**, 946–964.
- Carnap, R. (1947). *Meaning and necessity*. Chicago: Univ. of Chicago Press.
- Carpendale, J., & Chandler, M. (1996). On the distinction between false belief understanding and subscribing to an interpretive theory of the mind. *Child Development*, **67**, 1686–1706.
- Chandler, M. (1988). Doubt and developing theories of mind. In J. W. Astington, P. L. Harris, & D. R. Olson (Eds.), *Developing theories of mind* (1st ed., Vol. 1, pp. 387–414). Cambridge: Cambridge Univ. Press.
- Dennett, D. C. (1969). *Content and consciousness*. London: Routledge & Kegan Paul.
- Donaldson, M. (1978). *Children's minds*. Glasgow: Fontana.
- Dowty, D. R., Wall, R. E., & Peters, S. (1981). *Introduction to Montague semantics*. Dordrecht: D. Reidel.
- Fauconnier, G. (1985). *Mental spaces: Aspects of meaning construction in natural language*. Cambridge, MA: MIT Press.
- Flavell, J. H., Green, F. L., & Flavell, E. R. (1986). Development of knowledge about the appearance–reality distinction. *Monographs of the Society for Research in Child Development*, **51**(1, Serial No. 212).
- Fodor, J. A. (1975). *The language of thought*. New York: Crowell, TY.
- Fodor, J. (1992). A theory of the child's theory of mind. *Cognition*, **44**, 283–296.
- Frye, D., Zelazo, P. D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cognitive Development*, **10**, 483–527.
- Gopnik, A., & Astington, J. W. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearance–reality distinction. *Child Development*, **59**, 26–37.
- Leslie, A. M. (1987). Pretence and representation: The origins of "Theory of mind." *Psychological Review*, **94**, 412–426.
- Leslie, A. (1994). Pretending and believing: Issues in theory of ToM. *Cognition*, **50**, 211–238.
- Mitchell, P. (1996). *Acquiring a conception of mind*. Hove: Psychology Press.
- Moses, L. J., & Flavell, J. H. (1990). Inferring false beliefs from action and reactions. *Child Development*, **61**, 929–945.
- O'Neill, D. K., Astington, J. W., & Flavell, J. H. (1992). Young children's understanding of the role that sensory experiences play in knowledge acquisition. *Child Development*, **63**, 474–490.
- Perner, J. (1991). *Understanding the representational mind* (1st ed.). Cambridge, MA: MIT Press.
- Pillow, B. H. (1989). Early understanding of perception as a source of knowledge. *Journal of Experimental Child Psychology*, **47**, 116–129.
- Pillow, B. H. (1993). Preschool children's understanding of the relationship between the modality of perceptual access and knowledge of perceptual properties. *British Journal of Developmental Psychology*, **11**, 371–389.
- Pratt, C., & Bryant, P. (1990). Young children understand that looking leads to knowing (so long as they are looking into a single barrel). *Child Development*, **61**, 973–982.
- Quine, W. V. (1953). *From a logical point of view* (1st ed.). Harvard Univ. Press.
- Riggs, K. J., Peterson, D. M., Robinson, E. J., & Mitchell, P. (1998). Are errors in false belief tasks symptomatic of a broader difficulty with counter-factuality? *Cognitive Development*, **13**, 73–90.
- Robinson, E. J., & Beck, S. R. (2000). What is difficult about counterfactual reasoning? In P. Mitchell & K. J. Riggs (Eds.), *Children's reasoning and the mind* (pp. 101–119). Hove: Psychology Press.
- Robinson, E. J., & Mitchell, P. (1995). Masking of children's early understanding of the representational mind: Backwards explanation versus prediction. *Child Development*, **66**, 1022–1039.
- Robinson, E. J., & Robinson, W. P. (1982). Knowing when you don't know enough: Children's judgments about ambiguous information. *Cognition*, **12**, 267–280.
- Robinson, E. J., Thomas, G., Parton, A., & Nye, R. (1997). Children's overestimation of the knowledge to be gained from seeing. *British Journal of Developmental Psychology*, **15**, 257–273.

- Robinson, E. J., & Whittaker, S. J. (1987). Children's conceptions of relations between messages, meanings and reality. *British Journal of Developmental Psychology*, **5**, 81–90.
- Ruffman, T. K., & Olson, D. R. (1989). Children's ascriptions of knowledge to others. *Developmental Psychology*, **25**, 601–606.
- Russell, J. (1987). "Can we say . . . ?" Children's understanding of intensionality. *Cognition*, **25**, 289–308.
- Russell, J. (1992). The theory theory: So good they named it twice. *Cognitive Development*, **7**, 485–519.
- Russell, J. (1996). *Agency its role in mental development* (1st ed.). UK: Erlbaum.
- Searle, J. (1983). *Intentionality* (1st ed., Vol. 1). Cambridge: Cambridge Univ. Press.
- Siegal, M. (1997). *Knowing children* (2nd ed.). Hove, UK: Psychology Press.
- Siegal, M. (1999). Language and thought: The fundamental significance of conversational awareness for cognitive development. *Developmental Science*, **2**, 1–34.
- Taylor, M. (1988). Conceptual perspective taking: Children's ability to distinguish what they know from what they see. *Child Development*, **59**, 703–718.
- Wellman, H. M., Hollander, M., & Schult, C. A. (1996). Young children's understanding of thought bubbles and of thoughts. *Child Development*, **67**(3), 768–788.
- Wimmer, H., & Mayringer, H. (1998). False belief understanding in young children: Explanations do not develop before predictions. *International Journal of Behavioural Development*, **22**(2), 403–422.