When can children handle referential opacity? Evidence for systematic variation in 5- and 6-year-old children's reasoning about beliefs and belief reports

I.A. Apperly[a,*] and E.J. Robinson[b]

[a] *School of Psychology, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK*
[b] *Keele University, UK*

Abstract

Five- and 6-year-olds ($N = 51$) heard stories in which a character sorted items into two locations. Either the character had a false belief about one of the items (e.g., thought a tin contained biscuits, not Lego), or was only partially informed of an item's dual identity (e.g., did not know that a tie was a present). Children found it easier to reject a report of the character's belief that described the true state of affairs when the character had a false belief (e.g., Is Fred's uncle thinking "where shall I put this Lego?"), than to reject one in which an object known to the character was described using a term of which she was ignorant (e.g., Is Mum thinking "where shall I put this present?"). Similarly, children found it easier to predict the character's incorrect sorting of the target items for false belief (with food not toys) than for dual identity (in the wardrobe not with things to take on a visit). Correct reasoning about beliefs and reports of beliefs that misrepresent an object does not imply mastery of the fact that beliefs represent an object in a particular way.
© 2003 Elsevier Science (USA). All rights reserved.

Keywords: Theory of mind; Referential opacity; Intensionality; Belief reports; False belief

---

* Corresponding author. Fax: +44-0121-414-4897.
E-mail address: I.A.Apperly@Bham.ac.uk (I.A. Apperly).

Introduction

If we are told that there is a ball in the box, and also that the ball is a gift, it follows that there is a gift in the box. In contrast, if we are told that John is thinking that there's a ball in the box, it does not follow from the fact that the ball is gift that John is thinking that there's a gift in the box. This logical property of belief reports is described as ''referential opacity'' (see e.g., Quine, 1953), in contrast with the normal case of referential transparency, in which terms describing the same object (such as ball and gift in this case) may be substituted without changing the truth of the sentence. The behaviour of opaque contexts turns on the fact that the person whose mental state is described may not share our knowledge and beliefs about the world. For this reason, the transparency/opacity distinction has been used to illustrate what is gained as children learn to understand people's behaviour in terms of mental states. For example, Gopnik (1993) argues that ''...for the 3-year old, all serious psychological states [perceptions, desires, and beliefs as opposed to dreams and pretences] are 'transparent'.'' However, once children understand ''the possibility of misrepresentation'' [that a person's beliefs may be false] they will possess ''the intuitions captured by philosophical notions such as 'opacity''' (Gopnik, 1993). Similarly, Mitchell (1996) states ''If we have a working understanding of... *referential opacity*, (original italics) then we have a fundamental basis for an understanding of the mind as an organ that interprets and represents reality.'' Thus, the relation between children's conceptions of beliefs and their handling of the peculiar logical properties of belief reports is a topic of continuing interest to developmental researchers (e.g., de Villiers & de Villiers, 1999; de Villiers & Pyers, 2002; Gopnik, 1993; Hulme, Mitchell, & Wood, 2003; Kamawar & Olson, 1999; Leslie, 1987; Mitchell, 1996; Olson & Kamawar, 1999; Perner, 1991; Russell, 1992). We shall show that in this research, the handling of beliefs and of belief reports tend to be treated as single problems, that such treatment can be misleading, and that a more fine-grained analysis sheds light on 3- to 7-year-olds' developing understanding about representations.

If children's understanding of beliefs and of linguistic substitution in belief reports are as intimately related as authors such as Gopnik (1993) and Mitchell (1996) suggest, we should expect children to show similar patterns of performance on tests of these abilities. Consistent with this prediction, some studies have indeed found a relation between children's belief reasoning performance (as measured by false belief tasks) and their handling of substitution in belief reports (e.g., de Villiers & de Villiers, 1999; de Villiers & Pyers, 2002; Kamawar & Olson, 1999). For example, Kamawar and Olson (1999) gave children aged 3–7 years three story-based substitution problems. In one story a character gives a policeman some keys, but does not know that the policeman is a friend's dad. The substitution test question was ''Does Mark know that he gave the keys to Sue's dad?'' (correct answer NO). As an independent measure of children's belief reasoning they were given false belief tasks in which children had to take into account a character's false belief in order to predict how he or she would act (Perner, Leekam, & Wimmer, 1987; Wimmer & Perner, 1983). A significant, though small, correlation was found between summed performance on substitution questions and performance on false belief tasks.

However, the age at which children successfully handle substitution problems in belief reports varies widely, from 3 to 4 years in de Villiers and de Villiers (1999; see also de Villiers & Pyers, 2002), 4–5 years in Kamawar and Olson (1999) and Robinson and Mitchell (1992, 1994) to 6 or even 7 years in Apperly and Robinson (1998, 2001) and Russell (1987). If success on false belief tasks is taken as evidence that children understand about beliefs from around 4 years of age, then we must ask why children's success with handling substitution problems in belief reports is so varied.

Perhaps methodological differences can account for all the observed variation in performance. The methods used by de Villiers (e.g., de Villiers & Pyers, 2002) differ markedly from those in other studies. De Villiers (de Villiers & Pyers, 2002) simply asked children to repeat a belief report that had just been given by the experimenter. For example, "He thought he found a ring [in first picture], but (second picture) it was really a bottle cap. What did he think? (pointing back at first picture)." In contrast, the other studies used longer scenarios, and importantly, required children to make an inference, rather than simply repeat what had just been said to them. For example, in a study by Russell (1987) children were told a story in which a man with curly red hair steals George's watch while George is asleep. George wakes, finds that his watch has gone, and sets out to find the thief. In the crucial test question children are asked "Can we say that George is thinking 'I must find the man with curly red hair who stole my watch'?" Although it would be correct to say that George was thinking he must find the thief who stole his watch, it is wrong to substitute "man with curly red hair" for "thief" in the belief report, because although the thief *was* a man with curly red hair, George does not know this. To answer correctly children must follow the story, infer that George cannot know the thief's hair colour because he was asleep and actively notice that the belief report characterises George's belief in terms of the thief's hair colour. Might factors such as these account for the wide variability in children's performance on substitution tasks?

The relative simplicity of deVilliers' method, and in particular the fact that children did not need to make a belief inference, may indeed help explain why children performed particularly well in her studies. However, standard false belief tasks require the child to make a belief inference from a story, yet are within the capabilities of many 4-year-olds.

Another methodological factor that may contribute to children's difficulties with substitution problems in belief reports is the syntax of the test questions. For example, Russell's (1987) test questions involved multiple embedded clauses (e.g., "Can we say that George is thinking 'I must find the man with curly red hair who stole my watch'?"), and many 6- and 7-year-olds failed. In a study by Robinson and Mitchell (1992, 1994) the test questions were syntactically simpler. Children watched enacted stories in which a character with outdated knowledge made a request. For example, Mum and Jane tidy away two bags of different multi-coloured material, one into a red drawer and the other into a blue drawer. In Mum's absence Jane swaps the bags over. Mum, who is ignorant of the exchange, calls from another room, and says "I need some more material, it's the bag in the red drawer" and the child is asked which bag Mum wants. This utterance must be understood in terms of the beliefs of the speaker, not the real situation that obtains, that is, Mum wants

the bag she believes to be in the red drawer, but which is actually in the blue drawer. Many 4- to 5-year-olds answered correctly, consistent with the idea that simplifying the syntax of the test questions may reduce children's difficulties on substitution tasks.

Yet another factor, which may interact with such syntactic demands, is the pragmatic problem of interpreting belief reports. We have so far talked as if alternative descriptions of an object are never substitutable in belief reports, but this is not the case. In fact, belief reports may often be interpreted in such a way that these substitutions do not change the truth of the report, and the choice of interpretation is governed entirely by the context in which the report is made. Thus, in the example provided at the beginning of the paper, on seeing John's reaction to the ball, his mother might say to his father "John thinks that Jimmy's gift is exciting," and father might legitimately reply "he'll be upset when he finds out it's Jimmy's gift." Although researchers in this field have attempted to make the intended, substitution-sensitive interpretation compelling, it is always difficult to be certain that children will make the intended interpretation. Perhaps more importantly, it is difficult to be sure that the pragmatic demands of reaching the intended interpretation are similar across different ages and experiments. Thus, Perner (1991) and Russell (1992) both suggest that these pragmatic difficulties may be a key source of children's errors when handling belief reports.

Because of these methodological variations it is very difficult to compare the results from existing studies. It is therefore difficult to evaluate the proposed relation between children's belief reasoning (for example, being able to predict that someone will act according to what they believe) and their handling of substitution problems in belief reports (i.e., appreciating that a person's beliefs constrain the terms in which that belief is described). However, it is clear that, if all of the studies are essentially asking the same question of children, the differences in difficulty should be eliminated if the methodological factors are adequately controlled. In contrast, if other important differences exist between these studies, controlling methodological factors should not eliminate differences in difficulty. We now turn to consider one such important difference: that there is systematic variation in the *kind* of beliefs used in the various studies.

In the textbook referential opacity examples typically cited previously, the relation that exists between alternative descriptions is of a specific kind. For example, the Morning Star and Evening Star, or Oedipus' wife and Oedipus' mother are truthful current descriptions of the same entity; Venus and The Queen of Thebes, respectively. In such cases opacity arises when we describe the mental state of someone who knows only one of two salient identities for an object or person. The experiments of Apperly and Robinson (1998, 2001), Kamawar and Olson (1999), and Russell (1987) were designed to create just such a situation by allowing the protagonist to learn one identity for an object but not the other. So while in the example previous it was true that George knew that the thief had stolen his watch, it was not true that he knew that the thief was a man with curly red hair, even though the thief was in fact a man with curly red hair (Russell, 1987). These dual identity substitution problems were particularly difficult for children in studies by Apperly and Robinson (1998,

2001) and by Russell (1987), and significantly harder than false belief tasks used by (Kamawar & Olson, 1999).

In contrast, in the studies of de Villiers and de Villiers (1999; de Villiers & Pyers, 2002) and Robinson and Mitchell (1992) (in which children performed relatively well) the story characters were not partially informed about legitimate descriptions for an object. Rather, because the characters held a false belief, they were committed to the validity of a description that excluded the others in question. For example, in the Robinson and Mitchell study, "The bag of material in the red drawer" was not an alternative true description for the bag of material in the blue drawer; in de Villiers' studies "a ring" was not another equally true way of describing the bottle cap. In both cases, one was the description that the story character mistakenly thought was correct ("bag in red drawer;" "ring") while the other was the correct description ("bag in blue drawer;" "bottle cap"). In this respect, the relation between alternative descriptions in these belief reports was different from that in the substitution examples involving genuinely co-referential terms.[1] Might this be a reason why questions in the de Villiers (e.g., de Villiers & Pyers, 2002) and Robinson and Mitchell (1992, 1994) studies were easier than Apperly and Robinson's (1998, 2001) and Russell's (1987)?

One reason for thinking that the relation between alternative descriptions in belief reports might affect children's performance comes from studies by Apperly and Robinson (1998, 2001), who compared children's performance on different kinds of belief reasoning problem. In the Apperly and Robinson (1998) study 4- to 6-year-old children completed standard false belief reasoning tasks (e.g., Gopnik & Astington, 1988) where they had to predict what an uninformed protagonist would think was in a sweet box that (unknown to them) contained pens rather than sweets. Children also completed similar reasoning tasks where the protagonist knew one identity for an object but not another. For example, the protagonist knew that the object in location A was an eraser, and that the object in a location B was a die, but crucially did *not* know that it was also an eraser. Children's task was to predict where the protagonist would search when he wanted to find an eraser. Either location would satisfy his desire for an eraser, but he only knew there was an eraser in location A. Four-, 5- and 6-year-old children did not choose location A significantly more often than location B, suggesting that they were failing to take the protagonist's knowledge into account. However, these children performed significantly above chance on the false belief task, correctly predicting that a naïve protagonist would think there were sweets in the box.

Apperly and Robinson's (1998, 2001) studies suggest that children find dual identity belief reasoning problems significantly harder than false belief reasoning problems. From this perspective, it would not be surprising if children's ability to handle substitution problems concerning belief reports reflected this difference in their handling of the beliefs themselves. However, the literature does not typically

---

[1] Indeed, for this reason, some might question whether these were genuine cases of referential opacity. However, the current concern is not about terminology, but about whether the belief reports used in the different studies vary systematically.

distinguish between dual identity and false belief reasoning problems, and this distinction has never been applied in empirical investigations of substitution problems in belief reports. Furthermore, with the tasks used in the existing studies, it remains possible that methodological differences were solely responsible for the difference in children's performance on the false belief and dual identity tasks.

A strong test of the hypothesis that the kind of belief is an important factor in children's performance requires tasks that are very closely matched, so that the methodological variables remain as stable as possible while the kind of belief that the children are asked to consider is varied. We developed such false belief and dual identity tasks for the experiment described here. Children were asked to predict a story character's behaviour, based on his or her belief, and to evaluate linguistic substitution in a report of that belief. The expectation based on the argument above was that children would perform better in false belief conditions than in comparable dual identity conditions. In contrast, if the reported variation in children's difficulty with substitution problems was due to linguistic, pragmatic or other, methodological, factors, any variation in children's performance in the current study should be unsystematic with respect to the kind of belief involved.[2]

Method

Participants

We tested 56 children. Five were excluded for failing to attend to one or more stories and thus data from 51 children (24 boy and 27 girls) aged between 5;5 and 6;4 (mean age 5;11) were entered into the analysis. All children attended the same junior/infant school within a lower middle class catchment area in Birmingham UK, and spoke English as their first language.

Procedure

Children were tested individually on four story tasks each illustrated with pictures. Two stories involved a character knowing only one of an object's two identities and two involved a character with a false belief. In a similar way to standard false belief reasoning tasks (e.g., Perner et al., 1987; Wimmer & Perner, 1983) in each story children were asked to predict a character's action, based on his/her belief. In all four stories the character had to sort items into two locations. In one dual identity story the key item to be sorted was a "tie" that was also a "present" (gift) whereas in the other it was an "old cloth" that was also a "ningy" (child character's name for comfort blanket). The story character whose thoughts and sorting behaviour the child had to

---

[2] Of course, false beliefs, like all representational mental states, are held under some particular descriptions and not others, and so have the potential to cause dual identity problems. By referring to different kinds of belief problem we mean that in false belief questions it is the falseness of the belief that is relevant to solving the problem, not the *particular* (true or false) description under which it is held.

predict only knew the one, visually apparent label in each case (tie; old cloth). Hence their mental states corresponded to those in the traditional opacity examples of other experiments (Apperly & Robinson, 1998, 2001; Kamawar & Olson, 1999; Russell, 1987). In the false belief stories the key item to be sorted had a deceptive appearance: A biscuit tin actually contained Lego, and a "sweet" was really a trick. The story character had a false belief based upon the appearance of each item (biscuits; sweet), and so had a mental state very similar to that in standard false belief experiments (e.g., Gopnik & Astington, 1988;, Perner et al., 1987; Wimmer & Perner, 1983), and in the studies of children's handling of substitution problems conducted by de Villiers and de Villiers (1999), de Villiers and Pyers (2002), and Robinson and Mitchell (1992).

The substitution problem for each story required children to judge the acceptability of a belief ascription: "Is the character thinking where shall I put this X?" In the dual identity stories, X was the description about which the character was ignorant (present; old cloth). In the false belief stories, X was the object's true identity (Lego; trick) when the character had a false belief.

One dual identity story follows and the second is in the Appendix.

> This is a story about a very little girl called Judy. She's only two and she has an old cloth that she carries around with her everywhere. She calls it her ningy and no one else is allowed to take it away. She gets very upset when she loses it! One evening Judy was playing in the spare room. Suddenly she felt very tired and fell asleep right where she was, on top of a pile of old clothes. Eventually her mum found her and picked her up and put her to bed. The next day Judy's auntie came to stay, and she helped Judy's Mum tidy up the house. She went into the spare room and saw all the things on the floor in a mess. She picked up Judy's bottle and put it safely back in Judy's room, next to the cot. She picked up some of the old clothes and put them in the bin because they were dirty and old.

Questions were asked to ensure that children understood the basis for sorting the items: "Where will she put Judy's rattle? Where will she put that old shirt?" If children answered either of these questions incorrectly, they were corrected and told the reason why the item should go in the other location. They were then reminded explicitly about the character's partial knowledge, using a substitution insensitive (referentially transparent) statement: "Now, Judy's aunt doesn't know that the cloth is Judy's ningy. She sees this on the floor (pointing to cloth). Look, she's picked it up (picture of aunt holding the cloth)." The test questions followed:

Substitution question: "In this picture, what is Judy's auntie thinking? Is she thinking, 'where shall I put Judy's *ningy?*' .... or is that *not* what she's thinking?" (correct answer "No"). The form of this question was rather complex but was chosen to allow the false belief equivalent to be asked in the same way (see below).

Action question: "Where will she put that (pointing to Judy's ningy): In Judy's bedroom with her other things or in the bin with the clothes?" (correct answer "In the bin"). Children could answer this question verbally, or by pointing to a picture of the location.

Check question: (pointing to child's response) "Is that where it should go?"

If a child answered the check question correctly we could be more confident that a correct answer to the Action question reflected an understanding of the character's belief, rather than a misunderstanding of the object's role in the story.

The order of the substitution and action questions was counterbalanced between children, and the check question was always last.

False belief stories took a very similar form, but with these, the character came to have a false, rather than a true but importantly incomplete belief about the identity of the object. One such story follows and the other is in the Appendix.

> This story is about a boy called Fred. Fred likes to play with Lego. He keeps his Lego safe in an old biscuit tin so he doesn't lose all the little bits. Now, Fred's family is moving house and that means they have to pack up all their things into boxes so they can carry them to the new house. This is Fred's uncle, and he is helping Fred's family with the packing. Look, he's got one box for food and things that should go in the kitchen, and another for games and toys. He's packing up the things on this table. Look, he's put the jam in here with the food, and the doll in here with the toys.

Questions were asked to ensure that children understood the basis for sorting the items: "Where will he put the tomato sauce? Where will he put the playing cards?" As before, incorrect responses were corrected and reasons given why it should go in the opposite location. "Now, do you remember this? (pointing to closed biscuit tin full of Lego) Fred's uncle hasn't seen inside here before, so he doesn't know that the biscuit tin really has Lego inside. Look, he's picked it up, but he hasn't opened it." Thus, as with the dual identity stories, children were explicitly told of the protagonist's lack of knowledge before they were asked the test questions.

Substitution question: "In this picture, what is Fred's uncle thinking? Is Fred's uncle thinking 'where shall I put this *Lego?*' . . . . or is that not what he's thinking?" (correct answer "No").

Action question: "Where will Fred's uncle put this: in the box with the food or in the box with the toys?" (correct answer "In the box with the food").

Check question: (pointing to child's response) "Is that where it should go?"

The two false belief stories were always presented together as were the two dual identity stories, with the order of the two stories of a type counterbalanced between children and crossed with whether false belief or dual identity stories came first. Crossed with both was the order of the suggestions in the action question (for example "In Judy's bedroom with her rattle and bottle or in the bin with the old clothes?" versus "In the bin with the old clothes or in Judy's bedroom with her rattle and bottle?").

Results

Table 1 gives the incidence of correct answers for each of the test questions in each of the four stories. There were 8 errors on the check question for one of the false belief stories (Sweet/trick). However, a number of these children gave the impression that they thought the trick sweet should go in the food tin so that the joke could continue. Because this suggested that these children had understood the story no children were excluded for failing this question. Within action and substitution test questions, there was no sign of significant differences between the two dual identity stories (tie/present and cloth/ningy) or between the two false belief stories (biscuit tin/

Table 1
Five and 6-year-olds' correct responses to the action, substitution, and check questions in the false belief and dual identity tasks

Question type	False belief tasks			Dual identity tasks		
	Biscuit tin/ Lego	Sweet/Trick	Combined score Mean (SD)	Tie/Present	Cloth/Ningy	Combined score Mean (SD)
Action	40/51 (78%)	44/51 (86%)	1.65 (.59)	29/51 (57%)	32/51 (63%)	1.20 (.85)
Substitution	30/51 (59%)	33/51 (65%)	1.24 (.89)	21/51 (41%)	21/51 (41%)	0.82 (.91)
Check	51/51 (100%)	43/51 (84%)	1.84 (.37)	51/51 (100%)	51/51 (100%)	2.0 (.00)

Lego and sweet/trick). Each child was given 4 scores out of 2 based on the number of correct answers to the false belief action and substitution questions and to the dual identity action and substitution questions.

Given the correlation reported by Kamawar and Olson (1999) between children's performance on false belief tasks and referential opacity tasks (dual identity substitution problems) we began by examining the relation between children's performance on our four types of question. As can be seen in Table 2, there are significant partial correlations (controlling for age) between children's performance on all four question types. Indeed, the correlation between False Belief-Action scores and Dual Identity Substitution scores, although the smallest obtained here, was larger than that obtained by Kamawar and Olson (who also had a larger sample). This is consistent with the idea that our closely matched tasks have succeeded in reducing incidental variability in children's performance.

Our primary interest was in the possibility that our question types would differ in difficulty for children (see Table 1). We computed a repeated measures analysis of variance[3] with belief type (false belief versus dual identity) and question type (action versus substitution) as within-subject factors. Question order (action or substitution question first) and story order (false belief or dual identity first) were entered as between-subject factors.

There was a significant main effect of belief type, $F(1, 47) = 25.6$, $p < .001$. Collapsing over question type children performed better on false belief tasks (mean score 1.45/2) than on dual identity tasks (mean score 1.02/2). There was also a significant main effect of question type $F(1, 47) = 15.5$, $p < .001$. Collapsing across belief type children performed better on action questions (mean score 1.43/2) than on substitution questions (mean score 1.04/2). There was a significant main effect of question order, $F(1, 47) = 5.39$, $p = .025$ with children making more errors on both substitution and action questions when the substitution question was asked before the action question. All other effects, including the interaction between belief type and question type, were non-significant, all Fs < 3.41, all ps > .071.

---

[3] All significant main effects were checked with appropriate non-parametric statistics.

Table 2
Partial correlation coefficients ($N = 51$) between the 4 question types

	False belief substitution	Dual identity action	Dual identity substitution
False belief action	.54**	.55**	.38*
False belief substitution	–	.44*	.69**
Dual identity action	–	–	.60**

* $p < .01$.
** $p < .001$.

Discussion

In the introduction we noted that clear conceptual parallels may be drawn between the development of children's handling of beliefs and of the logical properties of belief reports. Empirically though, estimates of the age at which children successfully handle belief reports vary from 4 to 7 years, while it is often supposed that children handle beliefs successfully from around 4 years, when they pass false belief tasks. We considered two explanations of this pattern of findings. The first explanation accepts that children who pass false belief tasks have a general capacity for belief reasoning, and attributes the variation in their success with handling belief reports to varying syntactic, pragmatic or other, methodological, factors. This account predicts that controlling these factors should reduce the variability in children's performance with belief reports. The second explanation identifies false belief problems and dual identity problems as two distinct classes of belief reasoning task, which differ in difficulty. If this is correct, children's difficulties with belief reports should vary systematically according to the kind of belief reported, even when other factors have been controlled. The aim of our study was to distinguish between these accounts.

Children heard four stories in which a character came to have a belief that differed from their own. At the end of each story, the child was asked a question about how the character would act (a belief reasoning problem) and about what the character was thinking (a substitution problem in a belief report). The wording used for the questions was the same for all four stories, thereby controlling syntax exactly and other methodological factors very closely. However, in two stories the character had a false belief, whereas in the other two s/he knew only one of two salient descriptions of the object.

Consistent with previous findings (e.g., Apperly & Robinson, 1998, 2001), 5- and 6-year-old children were more accurate at predicting how a character would act when that character had a false belief than when s/he knew one true description but was ignorant of the description necessary for correct action. Crucially, the same pattern was found on the substitution questions: children were more accurate at judging that a character was not thinking "where shall I put this X?" in the false belief condition than in the dual identity condition. In the false belief condition X was the true description of an object that the character believed falsely to be a Y. In the dual identity condition, X was a correct description of an object that the character

knew only under a second correct description, Y. Our design was intended to keep methodological variation to an absolute minimum, yet children's ability to predict action based on a belief, and to answer substitution problems about belief reports varied systematically, according to the kind of belief involved.

This finding is important because no previous empirical work has distinguished between substitution problems based on reports of false beliefs, and the more standard cases of referential opacity that are based upon differing knowledge of alternative true descriptions. As we argued in our introduction, applying this distinction to existing studies gives some handle on the wide variation in the reported performance of children. Finding that this distinction predicts performance in the current, well controlled, within-child study is strong evidence that, besides syntactic, pragmatic, or other, methodological considerations, *the kind of belief* (i.e., the relation between belief and reality that the child must represent) involved in a substitution problem is crucial and hitherto unrecognised factor in children's performance.

It is also the case that no previous study has made systematic, within-task comparisons of children's ability to reason about beliefs and about linguistic substitution in reports of those beliefs. In the current study, children made significantly more errors on the substitution questions than on the action questions, for both kinds of belief. This is consistent with the idea that substitution questions do pose difficulties (syntactic, pragmatic, or otherwise) over and above belief reasoning per se, and should make us cautious about the use of such questions as a simple index of children's belief reasoning ability. However, it should be noted that the need to have one form to present both dual identity and false belief questions resulted in our substitution questions being relatively complex. Therefore, they may be a rather conservative measure of the difficulty of substitution problems in general.

A further noteworthy feature of the data was that, despite the differences in difficulty, there were significant correlations between children's responses to action and substitution questions, regardless of belief type. These findings are clearly consistent with an alternative interpretation of our data from the one we have been offering, which says that all of these questions tap the same basic understanding of beliefs or of representations in general (see e.g., Kamawar & Olson, 1999). However, although this is one way of explaining the observed relationship between children's handling of beliefs and of belief reports, such an account cannot readily explain the systematic differences in difficulty that we predicted and observed. In contrast, our account can explain the pattern of consistency plus differences in difficulty by assuming that success on the easier, false belief, tasks forms a necessary but insufficient condition for success on the more difficult, dual identity, tasks.

The current data make it very unlikely that the difference in difficulty between dual identity and false belief tasks is the result of varying syntactic, pragmatic or other, methodological factors. What, then, is the basis for the systematic difference between dual identity and false belief tasks? Several studies using tasks quite different from ours suggest that understanding of the mind is still undergoing qualitative change well after mastery of false belief has been achieved. For example, children up to the age of around 5 or 6 years overestimate the informativeness of ambiguous verbal input (e.g., Flavell, Speer, Green, & August, 1981; Robinson, 1994; Robinson

& Robinson, 1982; Robinson & Whittaker, 1987). Children of this age also have related problems judging the informativeness of pictures (e.g., Chandler & Helm, 1984; Robinson & Robinson, 1982; Taylor, 1988), and in recognising the unpredictability of different people's interpretations of ambiguous visual or verbal input (e.g., Carpendale & Chandler, 1996).

Chandler and colleagues (e.g., Chandler & Sokol, 1999) argue that until 6 or 7 years children lack a full understanding of the representational nature of knowledge and beliefs, specifically that they fail to treat beliefs as interpretations that may vary among individuals with the very same informational access. This idea of important changes at 6–7 years fits with the current pattern of 5- to 6-year-old children experiencing difficulty reasoning about certain belief problems. However, the idea that children's difficulty is with understanding the mind as an interpreter of information does not appear to distinguish between the problem of understanding false belief on the one hand and true beliefs about only one of an object's possible descriptions on the other (see also, Apperly & Robinson, 1998, 2001). Nevertheless, a connection can be made between children's handling of interpretation and their handling of dual identity problems. Robinson and Apperly (2001) (see also Apperly & Robinson, 1998) report an empirical relation between children's handling of dual identity substitution problems and their understanding of ambiguous messages. We believe this is because there is a problem common to handling ambiguous utterances and mental states that truly but incompletely represent their referents. This problem is with handling the fact that representations (whether mental, linguistic or pictorial) both truly represent an object and describe it in a particular way. As predicted by such an analysis, 5- and 6-year-old children have difficulty handling a word's referential and descriptive functions in a metalinguistic task that does not require the child to reason about beliefs or different people's interpretations of information (Apperly & Robinson, 2002).

In summary, we suggest that research on children's handling of referential opacity has failed to distinguish between the different kinds of relation that might hold between the linguistic terms whose substitution the child is supposed to judge. The current finding is that children find it relatively easy to reject substitutions of true descriptions when the protagonist's belief is false, but relatively hard when the protagonist has true belief under an alternative description. This goes some way to explaining the substantial variation in the age at which children have been found to pass such tasks, and should be an important consideration in future research on this topic. On the action prediction questions, the close matching of the dual identity and false belief conditions, including the fact that both tasks have zero baselines, provides the strongest evidence yet for a change in children's belief reasoning ability *after* they pass false belief tasks. Does this mean that we should withhold our attribution of a representational understanding of mind until a later age, perhaps until children pass dual identity action tasks, or dual identity substitution problems? We would resist imposing any such criterion and suggest that the representational nature of mental states should be viewed as posing more than one distinct problem, which children may solve in distinct ways and at different points in development.

Appendix

Dual Identity Story 2

This is a story about a boy called Martin. Martin and his mum and dad are going to visit his cousin who is much older than Martin, and lives in a house on his own. Before they go, Martin goes shopping with his dad to buy a present for the cousin. Shall we see what they buy? Look, they buy a tie to give to his cousin for a present, and they buy some new school trousers for Martin to wear next week. When they get home Martin takes all the shopping upstairs and leaves it on his bed. As usual the bed is very messy with Martin's clothes.

This is Martin's mum. She doesn't see Martin and his dad when they get back from shopping, so she doesn't know that the tie is a present. She goes into Martin's room and says to herself, I must pack up the things we need to take to the cousin's house, and I'll tidy up the other things too! She picks up some clothes and puts them in the wardrobe. Then she finds Martin's anorak and puts that in the bag to take to the cousin's house. Look, there's Martin's jumper. His mum thinks, "He'll need to take that with him to the cousin's house." Where will his mum put it? Look, there are his new school trousers. They don't need to go with them to the cousin's house. Where will she put them? Remember, Martin's mum doesn't know that the tie is a present for the cousin. She sees this on the bed. Look, she's picked it up.

  Questions
(1) In this picture, what is Mum thinking? Is she thinking "where shall I put the *present*?"......or is that not what she's thinking?
(2) Where will mum put that [pointing to tie/present]: In the clothes cupboard or in the bag to go to the cousin's?
(3) Is that where it should go?

False Belief Story 2

This is a story about a girl called Jemma. Jemma likes to play tricks and has lots of different things that she uses to play tricks on people. She's very excited today because she has a new one–shall we have a look at it? Here's a picture of it. It's a really good trick because when you pick it up and look at it, it seems just like a sweet, but if you try to eat it, it tastes all yucky. It's disgusting! She leaves it on the table in the lounge hoping that someone will come along and try to eat it.

Here comes her mum. She goes into the room and she sees the trick sweet on the table with lots of other things. Look, there is some chocolate and some biscuits and lots of Jemma's toys. But she's not very hungry, so she doesn't eat anything. Instead

she decides to tidy up. She puts the teddy bear in the box where Jemma keeps her toys and tricks, and the biscuits in the tin with the food in. Where will she put the chocolate? Where will she put the car? Now remember, Jemma's mum hasn't tasted the sweet so she doesn't know it's really a trick. She sees this [pointing to sweet/trick] on the table but she doesn't open it. Look here she is, she's picked it up:

    Questions

(1) In this picture what is mum thinking? Is she thinking "where shall I put this *trick*?"…or is that not what she's thinking?

(2) Where will she put this [pointing to sweet/trick]: In the toy box or the food tin?

(3) Is that where it should go?

References

Apperly, I. A., & Robinson, E. J. (1998). Children's mental representation of referential relations. *Cognition, 63*, 287–309.

Apperly, I. A., & Robinson, E. J. (2001). Children's difficulties handling dual identity. *Journal of Experimental Child Psychology, 78*, 374–397.

Apperly, I. A., & Robinson, E. J. (2002). Five year olds' handling of reference and description in the domains of language and mental representation. *Journal of Experimental Child Psychology, 83*, 53–75.

Carpendale, J. I. M., & Chandler, M. J. (1996). On the distinction between false belief understanding and subscribing to an interpretive theory of mind. *Child Development, 67*, 1686–1706.

Chandler, M. J., & Helm, D. (1984). Developmental changes in the contribution of shared experience to social role-taking competence. *International Journal of Behavioral Development, 7*, 145–156.

Chandler, M. J., & Sokol, B. W. (1999). Representation once removed: Children's developing conceptions of representational life. In I. E. Sigel (Ed.), *Development of mental representation* (pp. 201–230). Mahwah, NJ: Lawrence Erlbaum Associates.

de Villiers, J. G., & de Villiers, P. A. (1999). Linguistic determinism and the understanding of false beliefs. In P. Mitchell & K. Riggs (Eds.), *Children's reasoning and the mind* (pp. 191–229). Hove, UK: Psychology Press.

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

Flavell, J. H., Speer, J. R., Green, F. L., & August, D. L. (1981). The development of comprehension monitoring and knowledge about communication. *Monographs of the Society for Research in Child Development*, Serial No. 192.

Gopnik, A. (1993). How we know our minds: The illusion of first-person knowledge of intentionality. *Behavioural and Brain Sciences, 16*, 1–14.

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.

Hulme, S., Mitchell, P., & Wood, D. (2003). Six-year-olds' difficulties handling intensional contexts. *Cognition, 87*, 73–99.

Kamawar, D., & Olson, D. R. (1999). Children's representational theory of language: The problem of opaque contexts. *Cognitive Development, 14*, 531–548.

Leslie, A. M. (1987). Pretence and representation: The origins of "Theory of mind." *Psychological Review, 94*, 412–426.

Mitchell, P. (1996). *Acquiring a conception of mind*. Hove, UK: Psychology Press.

Olson, D. R., & Kamawar, D. (1999). The theory of ascription. In P. D. Zelazo, J. W. Astington, & D. R. Olson (Eds.), *Developing theories of intention: Social understanding and self control* (pp. 153–167). Mahwah, NJ: Lawrence Erlbaum Associates.

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

Perner, J., Leekam, S. R., & Wimmer, H. (1987). Three year old's difficulty with false belief: The case for a conceptual deficit. *British Journal of Developmental Psychology, 5*, 125–137.

Quine, W. V. (1953). *From a logical point of view* (first ed.). Cambridge, MA: Harvard University Press.

Robinson, E. J. (1994). What people say, what they think and what is really the case: Children's understanding of utterances as sources of knowledge. In C. Lewis, & P. Mitchell (Eds.), *Children's early understanding of mind: Origins and development* (pp. 355–381). Hove, UK: Erlbaum.

Robinson, E. J., & Apperly, I. A. (2001). Children's difficulties with ambiguity and partial knowledge arise from problems with partial referential relations. *Cognitive Development, 16*, 595–615.

Robinson, E. J., & Mitchell, P. (1992). Children's interpretation of messages from a speaker with false belief. *Child Development, 63*, 639–652.

Robinson, E. J., & Mitchell, P. (1994). Young children's false belief reasoning: Interpretation of messages is no easier than the classic task. *Developmental Psychology, 30*, 67–72.

Robinson, E. J., & Robinson, W. P. (1982). Knowing when you don't know enough: Children's judgements about ambiguous information. *Cognition, 10*, 267–280.

Robinson, E. J., & Whittaker, S. J. (1987). Children's conceptions of relations between message, meanings and reality. *British Journal of Developmental Psychology, 5*, 81–90.

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.

Taylor, M. (1988). Conceptual perspective taking: Children's ability to distinguish what they know from what they see. *Child Development, 59*, 386–396.

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