Russell, J., Saltmarsh, R. and Hill, Elisabeth L.

What Do Executive Factors Contribute to the Failure on False Belief Tasks by Children with Autism?

Originally published in Journal of Child Psychology and PsychiatryThe publisher's version is available at: http://dx.doi.org/10.1111/1469-7610.00504

J. Child Psychol. Psychiat. Vol. 40, No. 6, pp. 859–868, 1999
Cambridge University Press

0021–9630/99 $15.00 + 0.00

What Do Executive Factors Contribute to the Failure on False Belief Tasks by Children with Autism?

James Russell, Rebecca Saltmarsh, and Elisabeth Hill
Cambridge University, U.K.

As children with autism have pervasive executive difficulties it is necessary to determine whether these contribute to their often-reported failure on the false belief task. Failure on this task is frequently taken to diagnose the lack of a "theory of mind". We report two studies using two tasks that make similar executive demands to the false belief task. The first experiment showed that children with autism are significantly challenged by a "conflicting desire" task, which suggests that their difficulty with the false belief task is not rooted in difficulty with grasping the *representational* nature of belief. In the second study children with autism were also found to be impaired on a novel version of the "false photograph task". A parsimonious reading of these data is that their difficulty with all three tasks is due to commonalities in the tasks' executive structure.

Keywords: Autism, executive functions, conflicting desires, false belief task, false photograph task, diagnostic tests, cognition.

Abbreviations: BPVS: British Picture Vocabulary Scale; MLD: moderate learning difficulties; TROG: Test for the Reception of Grammar; VMA: verbal mental age.

Introduction

It is a well-established finding that children with autism are impaired on the false belief task, in which participants must judge that a protagonist will search for something on the basis of a false belief about its location rather than on the basis of what they themselves know about its true location (Baron-Cohen, Leslie, & Frith, 1985, for the initial demonstration in autism; Wimmer & Perner, 1983, for the original report of the task). Children with autism also perform at a significantly lower level than controls when they have to explain why the story protagonist visited the wrong location (Tager-Flusberg & Sullivan, 1994a). Because appreciating that false beliefs drive erroneous actions is fundamental to understanding belief per se and because understanding belief is fundamental to a conception of mind, such results have been taken to form the cornerstone of the hypothesis that the core impairment in autism is in "theory of mind" (Baron-Cohen, 1995).

However, these findings do not in themselves demonstrate that children with autism fail the false belief task because they have an absent or weak understanding of belief. This is because the task, while gauging the understanding of the link between belief and action, also makes significant executive demands. "Executive" can mean a number of things (Rabbitt, 1997; Roberts, Robbins, & Weiskrantz, 1996); so what does it mean in this context? Although there is no universally acknowledged definition of the term "executive functions" there is quite a broad consensus that an executive task has the following two components (Diamond, 1991; Pennington et al., 1997). First, the participant has to suppress a prepotent but incorrect response; second, he or she must retain action-relevant information in working memory while doing so. Thus, in the false belief task it is necessary to suppress one's own true belief whilst simultaneously holding in mind the requirement to answer a question about what the protagonist will do. When we couple this with the fact that children with autism have executive impairments that are not only severe but arguably unique in nature (Pennington & Ozonoff, 1996; papers in Russell, 1997), it can plausibly be argued that failing a task like false belief and its cognates may be due, at least in part, to executive difficulties[1].

This argument has equal, if not stronger, force when it is applied to second-order false belief tasks (e.g. "What does John think that Mary will think?") on which persons with autism are impaired (Baron-Cohen, 1989). A second-order false belief task may not in fact be testing for mental concepts at a deeper level than the first-order task. Most normal adults would struggle with, say, a fourth- or

Requests for reprints to: Dr James Russell, Department of Experimental Psychology, University of Cambridge, Downing Street, Cambridge CB2 3EB, U.K.

[1] This analysis is somewhat less persuasive when the subject has to explain erroneous search rather than predict which location will be visited (Tager-Flusberg & Sullivan, 1994a). But in this case there is still a competition set up between the subject's true belief and the requirement to frame an answer in terms of their representation of another's belief state held in working memory.

Table 1
Parallels between the Executive and Narrative Structures of the Tasks Used Here

False belief	False photograph	Conflicting desire
1. Child sees protagonist place object at location A. 2. Protagonist departs and the object is moved to place B. 3. Protagonist returns and child is asked where protagonist will look for the object (or where he thinks it is).	1. Child sees camera take a photograph of situation A. 2. While the film is developing the situation is changed to B. 3. Before the photograph is revealed the child is asked what situation will be pictured when the photograph is turned over.	1. Child and opponent both want to win card A. 2. The opponent fails to win card A but child succeeds; so he now needs card B. 3. Before they begin to play again the child is asked which card the opponent wants to win.
Correct answer: The original location A.	*Correct answer*: The original situation A.	*Correct answer*: The originally desired card A.
Executive demands: Inhibit reference to one's current and true belief while framing an answer in terms of protagonist's false belief—false because out-of-date.	*Executive demands*: Inhibit reference to current situation while framing an answer in terms of false photographic representation of it—false because out-of-date.	*Executive demands*: Inhibit reference to one's own current desire while framing an answer in terms of the opponent's current desire—an out-of-date desire from one's own perspective.

fifth-order false belief task, but this would surely not be because their understanding of mental life was insufficiently deep. Rather, the second-order task might only assess the ability to embed representations in working memory, something that is tested, in an executive context, through the use of planning tasks like the Tower of Hanoi and the Tower of London, tasks on which autistic participants are impaired (Hughes, Russell, & Robbins, 1994; Ozonoff, Pennington, & Rogers, 1991). Crucially, Tager-Flusberg and Sullivan (1994b) have demonstrated that when second-order false belief tasks are presented to autistic individuals in such a way that this kind of holding-in-mind is not required, those who pass the first-order tasks not only pass the second-order ones but can also justify their answers appropriately.

A possible objection to this line of argument is that children with autism are not specifically impaired on a task that makes similar executive structure to the false belief task but which makes no mentalising demands: Zaitchik's (1990) false photograph task (Leekam & Perner, 1991; Leslie & Thaiss, 1992). If this is so—the argument runs—they cannot be failing the false belief task because of their executive difficulties. In the false photograph task the child initially sees a photograph taken of a certain physical situation, such as a doll in a red dress (as in the Leekam & Perner study). Whilst the photograph is developing the situation is changed in some way (doll is put into a green dress). Then the child is asked what colour dress the doll will be shown to be wearing in the photograph. Alternatively, a toy animal moves from the chair to the bed (in the Leslie & Thaiss study) and the question concerns where the photograph will show the toy to be located. Note that this kind of task not only has a similar narrative structure to the false belief task but also requires a grasp of representation: representation in a photograph, however, rather than in a mind. See the first two panels of Table 1 for the parallels between the tasks in narrative and executive features.

The problem with this argument, however, is that it is not clear that the false photograph task makes executive demands *of a similar severity* to those made by the false belief task. In the case of the Leekam and Perner study, the child has to refer to what is in a photograph of an earlier state of the model rather than refer to the updated state of the model. If we assume that the child knows what is in the photograph then the requirement reduces to that of referring to what is known about a *two-dimensional* physical state rather than to what is known about a *three-dimensional* physical state. (Note that dolls too are representations—of girls.) We cannot assume that the former is necessarily less salient than the latter, and so we cannot assume that the task encourages a strong prepotent response. In the Leslie and Thaiss procedure the child has, similarly, to refer to a two-dimensional representation of an animal on a chair rather than a three-dimensional representation of an animal on a bed. Although the former is a representation *of* a representation, we cannot assume that its representational status is any way diluted for the child. In fact we cannot even safely assume that a two-dimensional representation of X will be less salient than X itself. In any event, we will assume for the time being that the implications of the false photograph data for an executive account of false belief failure in autism are at present unclear. We will return to this question after our first experiment has been reported.

In the first study we shall report on, the contribution of executive difficulties to failure on the false belief task was based on the following rationale. The false belief task requires the child to understand that two individuals (self and other) can have *conflicting beliefs* about reality. A belief can be analysed as (1) the mental orientation of holding true to (2) a particular representation of reality. Thus, if children are to understand what it is to believe something, they need to understand something about the representing relation: what it means for X to be a representation of Y (Perner, 1991). The situation is different, however, for desire. Although some have taken desire to be representational (Moses & Chandler, 1992, p. 289), and while the verb to desire can take a "that-clause" (Perner, 1991, pp. 116–117 for discussion), desiring is *not* fundamentally a matter of *representing the world as being a certain way*. Thus, one can simply desire an object without any attendant beliefs about the way the

world is currently configured[2]. Thus, in English at least, it is possible to say "I want it to be the case that I eat a cheeseburger", but we can of course simply say "I want a cheeseburger". We cannot by contrast say "*I believe a cheeseburger", and the fact that we cannot is more than a matter of linguistic convention: it is due to the fact that in believing we take the world to be a certain way (that Y is the case) whereas in desiring we generally do not[3].

On this analysis, understanding how two people can have different and conflicting desires—self and other in this case—is not a matter of understanding conflicting mental *representations*. It is not a matter, in other words, of understanding the representational nature of mind. However, understanding conflicting desire sufficiently well to answer questions about it *makes the same kind of executive demands as understanding conflicting beliefs*. On an executive analysis one's own occurent mental states (beliefs, desires, etc.) are arguably stronger than those of others, and therefore more likely to determine behaviour (e.g. the framing of a verbal judgement). The assumption here that one's actually *being* in a certain mental state is a fundamentally different matter from one's representing the mental state of another, either by theory-application or by simulation (Harris, 1991): actually being *subject* to it renders it more salient.

Using this argument, Moore et al. (1995) compared the performance of groups of 3-year-old children on tasks in which either their own beliefs or their own desires conflicted with those of another (a doll). In addition to being given a false belief task children were given one of two conflicting desire tasks. In the first, they had a choice of two stickers, one of which was much more attractive. The assumption was that they would chose the more attractive one. They were also told that the doll had an unfortunate experience that had made him frightened of the attractive sticker. They were then asked which sticker the doll would choose. The correct answer was taken to be the sticker that was not the children's preferred one. In the second conflicting desire task children played a game against a doll in which they had to win pieces of a jigsaw in an attempt to be the first to complete a picture of a frog. They turned over coloured cards which each represented a different jigsaw-piece. At test, matters were arranged so that the child and the doll needed cards of different colours; the child was asked which card the doll wanted. The correct answer was of course to say which card the doll needed rather than which card they themselves needed. See the third panel of Table 1 for the executive parallel between this task and the false belief task.

Moore et al. found that these two conflicting desire tasks were of a similar level of difficulty to the standard false belief task. From this, they concluded that what was responsible for the similarity in performance was the fact that the two tasks made equivalent executive demands. If the false belief task had proved more difficult than the conflicting desire it would have provided evidence for Perner's (1991) view that the false belief task challenges younger children because it taps an understanding of the representing relation between a belief and what it is a belief about—"metarepresentation".

The purpose of our first experiment was to compare conflicting desire with conflicting belief performance in children with autism. We already know, of course, that conflicting (i.e. false) belief tasks present these children with significant difficulties. If conflicting desire proves to be equally difficult then we will have evidence consistent with the view that it is the executive nature of the tasks that is challenging them. If, however, conflicting desire is easier to understand we will have produced evidence consistent with Leslie's (1987) hypothesis that the understanding of metarepresentation is specifically impaired in autism.

We gave the children only the second form of conflicting desire task used by Moore et al.—the frog jigsaw task. This was because we had found in pilot studies that children with autism did not tend to regard the relation between fear of an object and shunning it in a way that was necessary for the procedure to work. They were also no more likely to pick the sticker that we had assessed to be more attractive[4].

Experiment 1

The desire in our conflicting desire task was an individual's need for a particular object in the context of a game. The children had to appreciate that because they have already obtained the card they initially needed, their own desire has changed (they want a different card) whereas the other player will continue to want the card he or she initially wanted. (See the third panel of Table 1.)

As in the false belief task, an error consists of failing to acknowledge that another's mental state (desire/belief) does not change when the child's own mental state (desire/belief) has changed. In the conflicting desire task the change is brought about by the satisfaction of the subject's initial desire and by the lack of satisfaction of the other's. In the conflicting belief (or false belief) task it is caused by the transferring of an object from one place to the other, given that this changes the subject's belief without changing the other's belief.

Method

Participants. A group of 32 children with autism and a group of 32 children with moderate learning difficulties (MLD) participated in the study. Their verbal mental ages (VMAs) ranged between 3 and 8 years. In an initial session, all potential participants (from a pool of 78 children) were assessed using the British Picture Vocabulary Scale (BPVS) (Dunn, Dunn,

[2] One may say that certain background beliefs are in play when one says, for example, "I want a cheeseburger"—about cheeseburgers being food and about them being tasty—but no occurrent beliefs are. In contrast, the beliefs that the false belief task concerns are occurrent beliefs.

[3] In a recent House of Commons Select Committee hearing an MP (David Willats) defended his interpretation of events by trading on an ambiguity in desire. This was the ambiguity between "want" in the sense of being desirous of and "want" in the sense of being in want of (e.g. "I want for nothing"). Such an ambiguity cannot occur in belief. Belief is representational through and through, and so we cannot be "in belief of".

[4] A report of this study is available from the first author on request.

Table 2
Participant Details for Experiment 1

Age (months)	Autism group ($N = 32$)	MLD group ($N = 32$)
CA		
Mean	135.0	124.0
SD	37.5	22.7
Range	82–229	86–170
VMA (BPVS)		
Mean	69.1	69.2
SD	17.4	17.3
Range	36–99	39–100
VMA (TROG)		
Mean	66.8	70.0
SD	18.3	19.6
Range	48–120	48–132

CA: chronological age.

Whetton, & Pintilie, 1982) and the Test for the Reception of Grammar (TROG) (Bishop, 1982). Matching for the experiment was carried out solely on the basis of the BPVS scores. Participant details are given in Table 2. It can be seen from the table that VMA estimates gleaned from the TROG were very similar to those from the BPVS.

All the children with autism were drawn from schools or units specifically for children with autism. They were selected if they had been formally diagnosed as having autism or Asperger's syndrome. If a child had no firm diagnosis, then a questionnaire based on the diagnostic criteria of DSM-II-R (American Psychiatric Association, 1987) was completed by their teachers as an informal check on the children's symptoms and behaviour[5]. Of the 32 children assessed, 26 had a diagnosis of autism, 3 a diagnosis of Asperger's syndrome, and 3 were included on the basis of the questionnaire[6]. The children with MLD were attending special schools for children with a variety of learning difficulties.

Materials and design. Each child received two tasks: a conflicting desire task and a false belief task, with order of presentation being counterbalanced. A "Daffy Duck" glove puppet was used in both. In addition, two cardboard boxes, one red and one blue, were used in both tasks. In the conflicting desire task a set of coloured cards and two identical frog jigsaw puzzles were also used. The puzzles were made up of three pieces: the frog's body, its head, and its eyes. The eyes fitted onto the head, which in turn fitted onto the body.

Procedure. This was identical to that used by Moore et al. (1995). (The protocols are to be found in the Appendix). In the false belief task the child watched as Daffy put his money in the red box and left the scene. The experimenter moved Daffy's money to the blue box and subsequently asked the child where Daffy thinks his money is, after control questions had been asked concerning the original location of the money, its current location, and whether the puppet was aware of the transfer.

In the conflicting desire task, children played a game against the puppet. The aim of the game was to be the first to complete a frog jigsaw puzzle. Both players started with a body piece, and subsequently they had to first win the frog's head and then its eyes. The head pieces for both players were placed in the red box and the eye pieces in the blue box. It was explained to the children that the coloured cards in the pack before them corresponded to the colours on the box. Consequently, if a player drew a red card he won the head. Having done so, if he drew a blue card he would win the eyes. A yellow card was neutral and, if drawn, the player had to allow their opponent to turn over a card. Initially, therefore, both players wanted a red card. If either drew a red card to win a head, their desire would be fulfilled and changed to that for a blue card. The pack of cards was stacked so as to ensure that the child always drew a red card before the puppet. Once children had drawn the red card and had fitted the head to the frog's body they were asked the three control questions that paralleled those used in false belief tasks. They were then asked a question about Daffy's current desire: "Which colour does Daffy want now?".

Results

In the false belief task, every child correctly answered the control questions concerning where the object had been placed initially and where it was currently located. Eleven children with autism and one child with MLD failed the control question that concerned whether the protagonist knew where the object was currently located. On the assumption that this question about the another's knowledge state is a question about mental representation, and because there has been found to be a strong statistical contingency between success on the knowledge question and on the belief questions (Lumb & Russell, 1996), we counted the 12 children who failed the knowledge question as failing the whole task. Accordingly, 34% of the children with autism were counted as passing the false belief task (11 from 32), as against 84% of the children with MLD (27 from 32). This difference was statistically significant [$\chi^2(1) = 8.57$, $p < .01$]; see Table 3.

Turning to the conflicting desire task, 50% of the children with autism correctly answered the other-desire question (16 from 32), as against 78% of the children with MLD (25 from 32). This group difference was statistically significant [$\chi^2(1) = 6.22$, $p < .025$]. However, the question about the child's own desire was also failed by 5 of the 16 children with autism who had failed the other-desire question, and by 3 of the children with MLD. It could be argued that these 8 children could not have been failing the task because they had been influenced by the prepotent lure of their own desire (which they failed to report), so we reanalysed the group differences excluding the data from these children (see Table 3). The group difference on conflicting desire (other's belief) remained significant [$\chi^2(1) = 5.28$, $p < .05$].

Overall, the children did not find one of the tasks to be more difficult than the other [Cochrane's $Q(1) = .20$, n.s.]. This was also true at the level of the group. There was no difference in task difficulty in the children with autism [$Q(1) = 1.14$, n.s.] nor in the children with MLD [$Q(1) = .67$]. Finally, within each group there was no significant relationship between performance on the two tests [ϕ correlations both $p > .05$].

Discussion

These data show that a conflicting desire task that requires subjects to answer in terms of another's desire

[5] This questionnaire is available from the first author on request.

[6] Inspection of the data showed that the six children who did not have a diagnosis of autism performed similarly to those who did have this diagnosis.

Table 3
Performance of the Two Groups in Experiment 1 on the Two Tasks (32 Participants per Group)

	False belief task		Conflicting desire task		
Group	Passing	Failing	Passing	Failing (including)[a]	Failing (excluding)[b]
Children with autism	11 (34%)	21 (66%)	16 (50%)	16 (50%)	11 (34%)
Children with MLD	27 (84%)	5 (16%)	25 (78%)	7 (22%)	4 (12.5%)

[a] Including those failing the own-desire question.
[b] Excluding those failing the own-desire question.

rather than in terms of their own updated desire is nearly as challenging to children with autism as the false belief task. The false belief task can, as we have argued, equally be regarded as a conflicting belief task. Performance on the conflicting desire task by children with autism was significantly worse than that of VMA-matched children with mental handicap. These data parallel those of Moore et al. (1995) insofar as conflicting desire and false belief performance is similar in the two groups. The percentage success on the conflicting desire task in the children with autism (excluding those who failed the own-desire question) happened to have been identical to that of the normally developing 3-year-olds in Moore et al.'s study (34%).

How do these data relate to previous studies of the understanding of desire in children with autism? These studies have shown that children with autism are able to predict a protagonist's desire on the basis of his or her initial preference and that they are able to match desires with outcomes (Baron-Cohen, 1991a; Phillips, Baron-Cohen, & Rutter, 1995). Also, Tan and Harris (1991) have shown that children with autism will re-assert their original desire when what they have been given turns out not to correspond to it. Baron-Cohen (1991b) additionally reports that children with autism are no worse than mentally handicapped controls in predicting a protagonist's emotion on the basis of a fulfilled or unfulfilled desire. Children with autism are reported, moreover, to talk about desire in a relatively normal manner outside the laboratory (Tager-Flusberg, 1992). In short, the understanding of desire seems to be adequate in autism. With regard to the present study, if the understanding of desire is essentially normal in autism, then the case is strengthened for saying that the understanding of *conflicting* desire challenges these children by reason of the executive demands made in the task, not because of their poor understanding of that mental state.

While the children with autism were significantly worse than the MLD children on the conflicting desire task, it is also the case that the group difference was somewhat more marked on the false belief task. The children with autism seem to have been challenged more by the false belief task, whereas this was not true for the children with MLD. Does this not suggest that the metarepresentational nature of belief was causing the children with autism additional problems? This is possible. Indeed, we do not wish to insist that the false belief task challenges children with autism *solely* by virtue of its executive demands. However there is an alternative explanation for the more notable difficulty of the false belief task in the autism group. As was pointed out earlier, passing the false belief task required children to answer two mental-state questions correctly rather than one: the knowledge question (Does the protagonist know about the transfer?) and the belief question. This is tantamount—on the present argument—to increasing the executive demands.

We now consider an alternative interpretation of the data. It could be proposed that children with autism do indeed have a weak grasp of desire, and this is part of their generally weak grasp of *all* kinds of mental state. This difficulty with desire—on this argument—did not show up in the studies mentioned earlier because the tasks used were insufficiently challenging (e.g. matching desires with outcomes may be done by matching a sentence with an outcome). However, a task in which their own desire is pitted against that of another reveals how weak their grasp of desire really is. In other words, it is not the executive demands of the task per se that made it difficult for the children with autism but rather the fact that the conflict that had been set up was between two *mental* states, albeit nonrepresentational ones.

This alternative hypothesis makes the following prediction: children with autism will be mental-age-normal on a task that is executive in structure, that requires some understanding of misrepresentation, but which does not have a mental content. The false photograph task (described in the Introduction) fits this description and, as we have seen, two published studies have shown it to be correct (Leekam & Perner, 1991; Leslie & Thaiss, 1992). Children with autism do not appear to be challenged by the false photograph task. Therefore—the argument runs—they cannot be failing the false belief task for executive reasons. We have seen, however, that the assumption that the false photograph task makes executive demands similar to those of the false belief task can be challenged.

We earlier raised the question of whether the false photograph task does indeed encourage a prepotent response sufficiently strong to challenge participants who have executive difficulties. This is because it is not clear that the contents of a two-dimensional representation (e.g. that of a doll in a red dress) is necessarily less salient than that of a three-dimensional representation (the doll, representing a girl, in a green dress). We cannot even safely assume that a photograph of X is less salient than X itself or some updated state of X. Accordingly, our aim in the second experiment was to find out how children with autism perform when they are given a version of the false photograph task in which the prepotency is substantially boosted—in which assumptions about relative salience and what is a prepotent response are more solidly grounded. Our assumption was that, although judgements about prepotency are inevitably subjective to some degree, it is hardly debatable that *the presence of*

something is more salient than its absence. Thus, the arrival of a focal object onto a scene in which previously there was none would be expected to encourage executively challenged subjects to frame an answer that refers to it.

In our modified version of the false photograph task there was initially no object before the camera (apart from a screen), and the object was introduced while the film was developing. At test, the children had to refrain from saying that the camera would develop a picture of the newly arrived object. If executive factors have a significant role to play in the difficulty that children with autism have with false belief tasks then they should be less likely to answer this question correctly than control children.

Experiment Two

Method

Participants. Two groups of children took part in the study: 25 children with autism and 25 children with MLD. They were matched for VMA on the BPVS (see Table 4). Because the chronological ages of the children with autism were somewhat higher than those of the MLD group we also compared the verbal IQs of the two groups. These were not significantly different [$t(48) = 1.01$, $p > .05$]. The backgrounds of the children were similar to those of the children in our first experiment. Indeed, a number of them had served in that experiment. This time all of the 25 children with autism had received formal diagnoses of autism.

Design. The two groups of children and adolescents were given three tasks: a standard false photograph task, a modified (i.e. with increased prepotency) false photograph task, and a false belief task.

Table 4
Participant Details in Experiment 2

	Autism group ($N = 25$)	MLD group ($N = 25$)
CA (months)		
Mean	148.0	128.0
SD	32.3	22.2
Range	79–200	93–188
VMA (months)		
Mean	88.0	83.0
SD	25.0	26.4
Range	39–114	42–117

Apparatus. In both false photograph tasks a Polaroid camera was used. In addition, a coloured screen (40 × 30 cm) was set up as a backdrop for the photographs. Children were allowed to choose one of three toys to photograph in the training phase (a plastic frog, a wooden car, and a plastic horse). In all three tasks an Action Man and a Barbie doll were used. A marble and two cardboard boxes were also used in the false belief task.

Procedure. Participants were tested individually in a quiet room in their school. At the start of each session they were familiarised with the Polaroid camera. The experimenter erected the backdrop screen and allowed them to choose one of the three toys to place in front of it. The experimenter then helped each participant to take a picture of his or her chosen toy, after which they discussed the resulting photograph. The three tasks then followed, whose order of presentation was counterbalanced.

(1) *Modified false photograph task*. The screen remained in place, but without a toy in front of it. The experimenter said: "I'm going to take another photo. I'm looking through the little window and I'm pointing the camera over there (towards the screen). I've pressed the button and I've taken a photo. Here it is." Once the photograph emerged from the camera the experimenter placed it face down on the table and said, "Now we have to wait while the picture gets clear. While we are waiting let's get something out of my bag. Look I've got Barbie here. [Action Man was used for male subjects.] I know, while we are waiting, let's let Barbie sit here. The picture's coming out now." The experimenter indicated the back of the photograph and asked the following questions:

Control question: "When I took the photo, where was Barbie?"
(*Correct answer*: in the experimenter's bag.)
Test question: "What will we be able to see in the photo when it's ready?"
(*Correct answer*: pointing to screen, saying the colour of the screen, or saying something equivalent to "nothing" or "nobody".)

(2) *Standard false photograph task*. With the screen in place the experimenter said "Now look who I've got in my bag; it's Barbie. Barbie's going to sit over here [the experimenter placed the doll in front of the screen]. I'm going to take another photo. I'm looking through the little window and I'm pointing the camera over there. I've pressed the button and I've taken a photo. Here it is." Once the photograph emerged from the camera the experimenter placed it face down on the table and said, "Now we have to wait while the picture gets clear. While we are waiting let's put Barbie away. There, Barbie has gone back into my bag. Oh, look who else is in here, it's Action Man. Let's put Action Man on the table instead of Barbie. Now Action Man's sitting on the table. The picture's coming out now." The experimenter pointed to the back of the photograph and asked the following two questions in the following order:

Control question: "When I took the photo, where was Barbie?"
(*Correct answer*: in front of the camera.)
Test question: "What will we be able to see in the photo when it's ready?"
(*Correct answer*: Barbie.)

Table 5
Numbers and Proportions of Children per Group Passing the Tasks from a Total of 25 in Experiment 2

Group	False belief	False photograph	Modified false photograph (plus memory failers)	Modified false photograph (minus memory failers)	Q significance
Autism	10 (40%)	22 (89%)	10 (40%)	10 (62.5%)	$p < .001$
MLD	23 (92%)	22 (89%)	18 (74%)	18 (100%)	n.s.
Chi significance	$p < .001$	n.s.	$p < .05$	$p < .025$	

(3) *False belief task*. The experimenter said, "Now I'm going to tell you a story. This is Barbie and this is Action Man. They've got a red box and a blue box. [The child was told to point to each box in turn.] Look, Barbie's got a marble. Barbie's putting her marble into the red box and now she's going out to play. While Barbie's outside, Action Man's taken the marble out of the red box and is playing with it. Now he's putting it away. Look, he's putting it in the blue box. So the marble is in the blue box now, it has changed places. Barbie doesn't know that the marble has been moved. Here comes Barbie, she wants the marble." At this point the experimenter asked the following questions in the following order:

Reality question (control): "Where is the marble now?"
Memory question (control): "Where was the marble in the beginning?"
Belief question (test): "Where does Action Man think his marble is?"

Note that, in contrast to Experiment 1, we did not ask the "knowledge" control question. This was in order to reduce the difference between the number of questions asked in the false photograph tasks and in the false belief tasks[7].

Results and Discussion

The control question in the modified false photograph task (about where the doll was when the photograph was taken) was almost invariably failed by saying that the doll had been in front of the screen. However children with autism were not significantly more likely to make this error (36%) than the children with MLD (28%) [χ^2, $p > .05$]. Note that, as in the previous study, more control questions had to be included in the false belief task.

There was no difference in task difficulty within the MLD group [$Q(3) = 4.94$, n.s.]. By contrast, there was a significant effect of task in the children with autism [$Q(2) = 15.16$, $p < .001$]. Table 5 shows this to be due to the fact that the autistic group found the false photograph task in its standard format much easier than the other two tasks.

It can also be seen from Table 5 that the same number of children in each group passed the false photograph task in its standard format, in which an object was initially photographed and then replaced. In both groups 89% of the children succeeded. This is in line with previous studies. It was also to be expected that the children with MLD would outperform children with autism on the false belief task, and this seems to have been the case. Performance on the modified false photograph task was also superior in the MLD children.

Hierarchical log linear analyses were carried out to test both for group differences and order effects, with these factors being entered in a saturated model. In the false belief task, group membership, but not order, significantly affected performance [$\chi^2(1) = 15.1$, $p < .001$]. In the modified false photograph task group membership also had a significant effect. This was true both when those who failed the memory control question were included within the failing group [$\chi^2(1) = 5.1$, $p < .05$] and when they were excluded from the analysis entirely [$\chi^2(1) = 6.13$, $p < .025$]—see Tables 5 and 6. With regard to order, children were also more likely to pass the modified false photograph task if it had been presented after the standard false photograph task [$\chi^2(1) = 4.72$, $p < .05$].

Table 6
Contingency Table of Modified False Photograph Performance (25 per Group)

Group	Passing	Failing (including)[a]	Failing (excluding)[b]
Group			
Children with autism	10	15	6
Children with MLD	18	7	0

[a] Including those failing the memory control.
[b] Excluding those failing the memory control.

The children with autism found the modified false photograph task significantly harder than the standard version [McNemar $\chi^2(1) = 6.67$, $p < .01$]. This was not the case in the children with MLD [McNemar $\chi^2(1) = 1.5$, $p > .2$]. The latter result has to be interpreted with caution, however, given that ceiling effects were probably operating in this group.

Finally, it is notable that when control question failers were counted as task failers, exactly the same number of children with autism (40% of the total) passed the modified false photograph task as passed the false belief task. There was, however, no correlation between performance on the two tasks within the autistic group [$\phi = 2.48$, $p > .05$]. Thus, although the two tasks were of similar difficulty for the children with autism it was not the case that the same children tended to pass or fail both.

Of the three tasks, then, the false belief task and the modified false photograph were the ones that challenged the children with autism, with the children with MLD performing well in all three. This pattern is consistent with the claim that the false belief task makes heavy executive demands upon children with autism. It might be objected, however, that the original version of the false photograph task is more "natural" than the modified version, given that people do not normally take photographs with only a backdrop before the camera, and given that children may find it unnatural to answer that a photograph shows nothing or shows only the backdrop. While this is true, one can also argue that the modified version of the false photograph task parallels the false belief task more closely than does the standard version in at least one respect. In the false belief task the child is faced with the choice of referring to a place at which there is something as against a place at which there is nothing. A comparable demand is present in the modified false photograph task insofar as the test question requires the subject to refer to the null contents of a photograph (instead of referring to a box without contents).

General Discussion

Our first experiment demonstrated that children with autism are challenged by a task in which they have to inhibit reference to their own current desire while framing

[7] It should be noted that the knowledge question was neither asked in the original false belief study of Wimmer and Perner (1983) nor in the original demonstration of autistic difficulty with the task (Baron-Cohen et al., 1985).

an answer on the basis of another's desire. Indeed their degree of difficulty was nearly equivalent to the difficulty they experience with the false belief task. Because one can regard the false belief task as requiring subjects to inhibit reference to their current true belief while framing an answer about another's, it is possible to conclude that the tasks challenge children with autism for the same reason. They make the classic executive demands of having to inhibit a prepotent response whilst keeping action-relevant information in mind.

However, the first experiment does not tell us that it is executive factors alone which make both conflicting desire and conflicting belief (false belief) tasks hard for children with autism. This is because both tasks are mental-state tasks. These data do, though, undermine the view that it is the *representational* aspect of mind, in particular, which is difficult to grasp in autism (as claimed by Leslie, 1987), given that desire tasks do not require subjects to understand what it means for the mind to represent the world as being a certain way. Therefore, in order to test the stronger claim that the false belief task challenges children with autism essentially through its executive structure it would be necessary to show that a non-mentalistic (but representational) task with a similar executive structure to that of the false belief task is equally as challenging as the false belief task.

The two published studies showing mental-age-normal performance by children with autism on the false photograph task (Leekam & Perner, 1991; Leslie & Thaiss, 1992) would seem to have undermined this claim. However, the assumption that the false photograph task clearly requires the inhibition of a prepotent response can be questioned on the grounds that a two-dimensional representation (e.g. photograph of a doll in a green dress) may not necessarily be less cognitively salient than a three-dimensional representation (a doll in a red dress). We argued that if the camera initially took a picture with no focal object before it and the child were asked, after an object had meanwhile been stationed before the camera, what the film will show, this would challenge any participants who had executive difficulties. This conjecture was confirmed: the children with autism performed significantly worse on the modified false photograph task than did the children with MLD; they were also worse on the false belief task; but they performed at exactly the same level as the children with MLD on the standard version of the false photograph task.

Before commenting on the significance of these data for the study of autism, we need to confront the question of why the false photograph task is difficult for *normally developing* preschool children. In fact it is a little more difficult than the false belief task (Zaitchik, 1990). It is not possible, within our rationale, to say that the task is difficult for executive reasons, particularly if one also wants to claim that the false belief task challenges normally developing children because of the executive demands it makes (Moore et al., 1995). It is possible that false photograph tasks challenge normally developing children because they do not understand referential opacity. We know that their difficulties with the logical opacity created by mental verbs continues beyond the fifth year of life (Russell, 1987). Thus, when they hear questions such as "In the picture, what colour is Julie?" (Leekam & Perner) or "In this photograph who is sitting in this toy box?" (Leslie & Thaiss) they fail to appreciate —for whatever reason—that the reading should be logically opaque rather than logically transparent[8]. Their reading might be translated as "What colour is Julie (who has been photographed)?" or "Who is sitting in this toybox (which has been photographed)?".

Finally, what do these data imply, then, about the nature and severity of mentalising difficulties in autism? It is naive to conclude that they undermine the view that there *are* mentalising difficulties in autism. For even if one ignores the clinical picture and Kanner's original description of the syndrome, children and adults with autism are impaired on tasks with mental contents that do not make the kinds of executive demand described above. For example, they find it more difficult than controls to discriminate mental from non-mental verbs (Baron-Cohen et al., 1994, Expt. 1); they find it difficult to interpret the mental "language of the eyes" (Baron-Cohen, Wheelwright, & Jolliffe, 1997); and they find irony difficult to detect (Happé, 1993). That said, these data do not tell us that individuals with autism *lack* a theory of mind: they show us that individuals with autism do not operate with this "theory" (on a very loose definition of the word) very efficiently, or do not operate with it in the same way that we do. Moreover, it could be said that answering questions about figurative language—especially irony—requires the subject to inhibit reference to reality; and this is an executive demand.

It is necessary, however, to insert that caveat that Experiment 2 does not, by strict logic, demonstrate that children with autism fail executive tasks because of the executive demands these tasks make. This is because it is possible that children with autism fail the false belief task because of its mentalistic content and fail the modified false photograph because of its executive demands, with the standard false photograph task posing no problems because it makes neither kind of demand. However "strict logic" can sometimes be in tension with the requirement to prefer the more parsimonious explanation. The false belief task does indeed make executive demands (both a priori and as suggested by the results of Experiment 1), as does the modified false photograph task; and that is all they share. So it is difficult to insist that they may be difficult for different reasons. Further research is certainly needed to clarify the question.

What these data do achieve, at least, is the exerting of pressure on the idea that failing the false belief task is an acid test for possessing a theory of mind in a clinical group. The clearest manifestation of this way of thinking

[8] Logically opaque contexts can be created in a number of ways, but their creation by mental verbs is the easiest to appreciate. For example, in the sentence "Oedipus thought Jocasta was desirable" the truth-value changes from true to false if we replace "Jocasta" with "his mother", despite the fact that both terms refer to the same person. To think that such a replacement is permissible is to fail to appreciate logical opacity. Similarly, in the context of pictorial rather than mental representation, the opaque rather than transparent reading can be required. See Fauconnier (1985) for a discussion of the parallels between the two kinds of logical opacity—pictorial and verbal.

is to be found among those who feel that the theory-of-mind deficit hypothesis of autism is threatened by the fact that some autistic subjects pass the test whilst presenting as no less autistic than those who fail it (Happé, 1993). This apparent challenge is usually surmounted by saying that these individuals have, by virtue of their age and verbal intelligence, managed to work out the right answer to such puzzles by generalising from past instances without insight into the general principle—the so-called "hacking out" of a solution (Happé, 1993, pp. 57–58). But it can be argued that such a challenge to the theory-of-mind-deficit hypothesis is minor when set against the more general concern that the false belief task makes deeply ambiguous demands: that it can be failed for reasons that may have rather little to do with lacking a concept of belief.

Acknowledgements—This research was made possible by a grant from the Wellcome Trust to the first author. We are grateful to the head teachers, staff, and pupils of the following schools for their kind cooperation and support: Baskerville School, Birmingham; Harford Manor School, Norwich; Lady Adrian School, Cambridge; Llangynwydd School, Mid Glamorgan; Maytree Autistic Unit, Swansea; Priory Annexe, Stormhouse School, Doncaster; The Lindens School, Penarth; and Walnuts School, Milton Keynes.

References

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders* (3rd revised ed.; (DSM-III-R). Washington, DC: Author.

Baron-Cohen, S. (1989). The autistic child's theory of mind: A case of specific developmental delay. *Journal of Child Psychology and Psychiatry*, *30*, 285–298.

Baron-Cohen, S. (1991a). Do people with autism understand what causes emotion? *Child Development*, *62*, 385–395.

Baron-Cohen, S. (1991b). The development of a theory of mind in autism: Deviance and delay? *Psychiatric Clinics of North America*, *14*, 33–51.

Baron-Cohen, S. (1995). *Mindblindness*. Cambridge, MA: MIT Press.

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

Baron-Cohen, S., Ring, H., Moriarty, J., Schmitz, B., Costa, D., & Ell, P. (1994). Recognition of mental state terms: Clinical findings in children with autism and a functional imagining study of normal adults. *British Journal of Psychiatry*, *165*, 640–649.

Baron-Cohen, S., Wheelwright, S., & Jolliffe, T. (1997). Is there a "language of the eyes"? Evidence from normal adults, and adults with autism or Asperger Syndrome. *Visual Cognition*, *4*, 311–331.

Bishop, D. V. M. (1982). *Test for the reception of grammar*. Cambridge: MRC Applied Psychology Unit.

Diamond, A. (1991). Neuropsychological insights into the meaning of object concept development. In S. Carey & R. Gelman (Eds.), *The epigenesis of mind*. Hillsdale, NJ: Lawrence Erlbaum.

Dunn, L. M., Dunn, L. M., Whetton, C., & Pintilie, D. (1982). *British Picture Vocabulary Scale*. Windsor, U.K.: NFER–Nelson.

Fauconnier, G. (1985). *Mental spaces*. Cambridge, MA: MIT Press.

Happé, F. G. E. (1983). Communicative competence and theory of mind in autism: A test of relevance theory. *Cognition*, *48*, 101–119.

Harris, P. (1991). The work of the imagination. In A. Whiten (Ed.), *Natural theories of mind*. Oxford: Basil Blackwell.

Hughes, C., Russell, J., & Robbins, T. W. (1994). Evidence for executive dysfunction in autism. *Neuropsychologia*, *32*, 47–92.

Leekam, S. R., & Perner, J. (1991). Does the autistic child have a metarepresentational deficit? *Cognition*, *40*, 203–218.

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

Leslie, A. M., & Thaiss, L. (1992). Domain specificity in conceptual development: Neuropsychological evidence from autism. *Cognition*, *43*, 225–251.

Lumb, A., & Russell, J. (1996). *The role of understanding the ignorance question in false belief performance*. Poster presented at the annual conference of the Developmental Psychology Section of the British Psychological Society, Oxford, September.

Moore, C., Jarrold, C., Russell, J., Lumb, A., Sapp, F., & MacCallum, F. (1995). Conflicting desire and the child's theory of mind. *Cognitive Development*, *10*, 467–482.

Moses, L., & Chandler, M. (1992). Traveller's guide to children's theory of mind. *Psychological Inquiry*, *3*, 286–301.

Ozonoff, S., Pennington, B. F., & Rogers, S. J. (1991). Executive deficits in high-functioning individuals with autism: Relationships to theory of mind. *Journal of Child Psychology and Psychiatry*, *32*, 1081–1105.

Pennington, B. F., Rogers, S. J., Bennetto, L., Griffith, E. M., Reed, D. T., & Shyu, V. (1997). Validity tests of the executive dysfunction hypothesis. In J. Russell (Ed.), *Autism as an executive disorder*. Oxford: Oxford University Press.

Pennington, B. F., & Ozonoff, S. (1996). Annotation: Executive function and developmental psychopathologies. *Journal of Child Psychology and Psychiatry*, *37*, 51–87.

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

Phillips, W., Baron-Cohen, S., & Rutter, M. (1995). To what extent can children with autism understand desire? *Development and Psychopathology*, *7*, 151–169.

Rabbitt, P. M. A. (1997). *Methodology of frontal and executive functions*. Hove, UK: Psychology Press.

Roberts, A., Robbins, T. W., & Weizkrantz, L. (Eds.). Executive and cognitive functions of the prefrontal cortex. *Philosophical Transactions of the Royal Society: Biological Sciences*, *351*, No. 1346.

Russell, J. (1987). "Can we say…?" Children's understanding of intensionality. *Cognition*, *25*, 289–308.

Russell, J. (1997). *Autism as an executive disorder*. Oxford: Oxford University Press.

Tager-Flusberg, H. (1992). Autistic children's talk about psychological states: Deficits in the early acquisition of a theory of mind. *Child Development*, *63*, 161–172.

Tager-Flusberg, H., & Sullivan, K. (1994a). Prediction and explaining behaviour: A comparison of autistic, mentally retarded, and normal children. *Journal of Child Psychology and Psychiatry*, *35*, 1059–1075.

Tager-Flusberg, H., & Sullivan, K. (1994b). A second look at second-order belief attribution in autism. *Journal of Autism and Developmental Disorders*, *24*, 577–586.

Tan, J., & Harris, P. L. (1991). Autistic children understand seeing and wanting. *Development and Psychopathology*, *3*, 163–174.

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

Zaitchik, D. (1990). When representations conflict with reality: The preschooler's problems with false beliefs and "false" photographs. *Cognition*, *35*, 41–68.

Manuscript accepted 27 November 1998

Appendix

Experiment 1: False Belief Task

Children are shown two empty boxes, one red and one blue. Daffy enters and the experimenter says, ''Here comes Daffy. Daffy's been working in a shop all day and has earned this money. Look, one pound! Daffy is going to put his money away in the red box [Daffy opens the red box, places his money inside and closes the lid]. Now he's going out to play. He will come back for his money later [Daffy is put into the experimenter's bag]. Let's play a trick on Daffy. Let's move his money to the blue box [The experimenter takes the coin out of the red box, puts it in the blue box and closes both boxes]. Daffy is coming back now and wants to find his money.''

As Daffy is returning, the following three control questions are asked:

Initial state question: ''Where did Daffy put his money at the start?''
Current state question: ''Where is the money now?''
Protagonist's state question: ''Does Daffy know that we've moved the money?''

Children were then asked the belief question:

Belief question: ''Where does Daffy think the money is now?''

Experiment 1: Conflicting Desire Task

Children are told how to play the game: ''Let's play a game with Daffy. Look at these frogs. Let's take one apart and put it back together again [The experimenter shows the child how the puzzle fits together]. Now this is how to play the game. You both start off with just a body and we put both of the heads in this red box and both sets of eyes in this blue box. Then you and Daffy take it in turns to pick up a card. If you pick a red card, like that, then you get a head from the red box and put it on the frog's body like this. Then you pick up cards until… you get a blue card like that. When you get a blue card, you can get some eyes from the blue box and you have won. Remember, you can't get the eyes from the blue box until you have the head from the red box. Do you understand? Daffy thinks he's good at this game. Do you want to play with him?''

The child and Daffy begin to play. Once the child has drawn a red card and fitted the head onto the body of the frog, three control questions are asked:

Initial state question: ''Which colour card did you get last time?''
Current state question: ''Which colour card do you want now?''
Protagonist's state question: ''Has Daffy won a head yet?''

The child is then asked the following test questions:

Other desire question: ''Which colour card does Daffy want now?''
Own outdated desire question: ''Which colour card did you want last time?''