Title: Journal of autism and developmental disorders.

Volume: 34  Issue: 2  Month/Year:  Apr 2004  Pages: 177-188

Article Author:

Article Title: Grant CM; Riggs KJ; Boucher J; Counterfactual and mental state reasoning in children with autism.

Counterfactual and Mental State Reasoning in Children with Autism

Cathy M. Grant,¹,⁴ Kevin J. Riggs,² and Jill Boucher³

The contributions of counterfactual conditional reasoning (CCR), belief understanding, and inferential reasoning to the performance of children with autism (CWA) on standard false belief tasks were investigated. To assess the roles of these three factors, we compared the performance of CWA on physical-state CCR tasks (which do not require either an understanding of belief or inferential reasoning); on Wellman and Bartsch's (1988) nonstandard tests of false belief (which require an understanding of belief, but not CCR or inferential reasoning), and on standard tests of false belief tasks. The CWA were impaired relative to controls on the physical-state CCR and standard false-belief tasks, but not on the nonstandard false-belief tasks, and the CWA's performance on the physical-state CCR and standard false-belief tasks correlated highly, even when the effects of verbal ability were partialed out. Finally, the CWA's performance on standard false-belief tasks was more impaired than their performance on the physical-state CCR tasks. We concluded that impaired performance on standard false-belief tasks in autism is associated with defective competence in CCR (or some of its component skills), plus defective competence in inferential reasoning and possibly generativity, but that impaired performance is not caused by an inadequate understanding of belief. The results are discussed in relation to other hypotheses concerning the cause or causes of impaired performance on standard false-belief tasks in children with autism.

KEYWORDS: Counterfactual reasoning; belief reasoning; false belief task; autism.

INTRODUCTION

A substantial body of research has reported that people with autism have an impaired understanding of mental states (see Baron-Cohen, 2000, for a review). In particular, it is well established that children with autism (CWA) have impaired understanding of false belief. This has led to the claim that an impaired theory of mind constitutes a major cognitive deficit in autism (Baron-Cohen, 1995; Leslie & Roth, 1993; Leslie & German, 1995).

A widely held view is that CWA fail standard false-belief tasks because they do not appreciate the representational nature of mental states, and hence they fail to understand that the world is not always truly reflected in people's beliefs (Baron-Cohen, 1995; Perner, 1991). In other words, it is claimed that CWA have a flawed conceptual understanding of belief. A similar explanation has been offered for the failure of typically developing young children on false-belief tasks (Perner, 1991). One difficulty with this account is that mental state understanding is only selectively impaired in autism. For example, the ability to understand simple desire is not impaired (Baron-Cohen, 1991; Tan and Harris, 1991). The ability to understand complex aspects of desire (in which it is not possible to use a goal-and-outcome-matching strategy) is impaired, but less
similar reasoning process is involved. Specifically, one must simulate Sally’s knowledge base by replacing one’s own knowledge or belief that the marble is in the box with Sally’s false belief that the marble is in the basket, and use the rule that people act on their beliefs to conclude that Sally will look for the marble in the basket. In this example, the individual temporarily adopts another person’s false-belief state, rather than a counterfactual physical state of the world. However, at least some of the cognitive requirements of the two types of task are the same, according to Peterson and Riggs, in that both types of task involve reasoning of the form “If X, then Y?” where X is a counterfactual proposition.

Data supporting the suggestion that physical-state CCR tasks and standard false-belief tasks share at least some cognitive requirements were reported in Riggs, Peterson, Robinson, and Mitchell (1998). In that study, the authors compared the performance of typically developing young children on a standard (Sally-Anne) false-belief task with their performance on physical-state CCR tasks of the kind described above. They found that not only did the majority of 3-year-old children fail the physical state tasks but also the performance on these tasks correlated significantly with performance on standard false-belief tasks, even with verbal mental age partialed out.

However, Riggs et al. (1998) also reported that typically developing children performed significantly worse on the false-belief questions compared with the physical-state questions. Two possible interpretations of this observation were offered. First, that it may be the case that CCR is necessary but not sufficient to pass false-belief tasks, and that a certain level of understanding of the representational nature of belief is also required (Perner, 1991). In this case, the discrepancy between typically developing children’s performance on standard false belief tasks and physical state tasks can be seen as resulting from an immature conceptual understanding of the nature of belief. Alternatively, the “If X, then Y?” reasoning required for physical state tasks may be easier than the “If X, then Y?” reasoning that is required for standard false-belief tasks because in the former type of task the counterfactual proposition “X” is provided explicitly (“If fish couldn’t swim . . .”). In contrast, in a standard false-belief task, the child has to infer Sally’s false belief that “the marble is in the basket” from information that is implicit in the action of the story.

According to the latter interpretation, CCR physical-state tasks can be seen as resembling Sparrevoorn and Howie’s nonstandard false belief tasks.
The first of these nonstandard tasks is described as an "inferred belief" task. In this task, a protagonist's behavior must be predicted on the basis of a false belief, which can be inferred from explicit information provided within the story. For example:

This is Jane. This morning Jane saw her colored pencils on the desk, not on the shelf. Now Jane wants her colored pencils. Where will she look for them?

The second nonstandard task is described as a "not own belief" task. In this task, a protagonist's behavior must be predicted on the basis of a false belief that is explicitly provided, but that is in opposition to the child's own guess as to the true state of affairs. For example:

This is Sam. Sam wants to find his puppy. It might be hiding in the house or in the garden. Where do you think Sam's puppy is hiding? (Child answers e.g., "garden"). That's a good guess. Sam thinks his puppy is in the house (different location than the child's answer). Where will Sam look for his puppy?

The third task is described as an "explicit false belief" task. In this task, a protagonist's behavior must be predicted on the basis of a false belief that is explicitly provided, but that is in opposition to the true state of affairs, which is also explicitly described within the story. For example:

This is Mary. Mary wants to find her kitten. Mary's kitten is really in the bedroom. Mary thinks her kitten is in the kitchen. Where will Mary look for her kitten?

In these tasks, as in physical-state CCR tasks, the counterfactual proposition X (which is equivalent to the protagonist's false belief) is either stated explicitly (e.g., "Sam thinks his puppy is in the house"; "Mary thinks her kitten is in the kitchen") or, in the case of inferred belief, the event underpinning the protagonist's false belief is explicitly described (e.g., "Jane saw her colored pencils on the desk"). However, nonstandard false-belief tasks do not, we would argue, require counterfactual conditional reasoning. Rather, the child has to accept the described situation as a starting point ("Sam thinks his puppy is in the house") and use the rule that people's behavior derives from their beliefs to reason that Sam will look for his puppy in the house. In contrast, in a physical-state CCR task, the child must take the counterfactual proposition as a starting point ("fish can't swim"), and run a simulation of what would happen if a nonswimming fish tried to live in the sea to reason that nonswimming fish could not live in the sea. No standard rule exists that allows the child to bypass this reasoning process in the case of physical-state CCR tasks.

According to our analysis, therefore:

- Physical-state CCR tasks require CCR, but not an understanding of belief or an ability to infer counterfactual states of affairs from implicit information.
- Nonstandard false-belief tasks require an understanding of belief, but not CCR or an ability to infer counterfactual states of affairs from implicit information.
- Standard false-belief tasks require all three skills.

The aim of the study reported here was to assess the roles of CCR, belief understanding, and inferential reasoning in contributing to the delayed acquisition of theory of mind in CWA and, in particular, to their delayed ability to pass false-belief tasks. To assess the roles of these three factors we compared the performance of CWA on physical-state CCR tasks, on standard tests of false belief, and on Wellman and Barsch's (1988) nonstandard tests of false belief. We hypothesised (1) that CWA have problems associated with the complex cognitive demands of counterfactual conditional reasoning, and we predicted that CWA would therefore be impaired relative to controls on physical-state tasks and on standard false-belief tasks; (2) that CWA do not have impaired understanding of beliefs and, in particular, that they know that people act on their beliefs even when they are false—we based this hypothesis on Sparrevohn and Howie's findings and, therefore, predicted that CWA will show normal performance, relative to controls, on nonstandard false belief tasks; (3) that CWA have difficulty in inferential reasoning, and we predicted that CWA will perform worse on standard false-belief tasks than on physical-state CCR tasks for this reason, rather than because they have an immature understanding of belief (see hypothesis 2).

**METHOD**

**Participants**

Three groups of children took part in the study. The first was a group of 25 children and adolescents with autism, all of whom had been previously diagnosed as having autism using established criteria for autistic disorder (Diagnostic and Statistical Manual of Mental Disorder, Fourth Edition, DSM-IV; American Psychiatric Association, 1994). Clinical psychologists/psychiatrists
Table I. Participant’s Mean Chronological Age (CA), Verbal Mental Age (VMA), and Verbal IQ (VIQ: standard deviations are shown in brackets)

<table>
<thead>
<tr>
<th>Group</th>
<th>CA</th>
<th>VMA</th>
<th>VIQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Autism (n = 25)</td>
<td>11.5</td>
<td>7.9</td>
<td>74.24</td>
</tr>
<tr>
<td></td>
<td>(2.91)</td>
<td>(1.64)</td>
<td>(15.81)</td>
</tr>
<tr>
<td>Mental retardation (n = 25)</td>
<td>11.2</td>
<td>7.42</td>
<td>73.95</td>
</tr>
<tr>
<td></td>
<td>(2.51)</td>
<td>(1.45)</td>
<td>(16.6)</td>
</tr>
<tr>
<td>Typically developing (n = 25)</td>
<td>7.6</td>
<td>7.5</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>(1.2)</td>
<td>(1.2)</td>
<td>(5.01)</td>
</tr>
</tbody>
</table>

who were not involved in this study had made the diagnosis. These participants were all attending special schools for autism. The second group comprised 25 children and adolescents with moderate mental retardation (MMR) of unknown origin who attended special schools. None of these participants had a diagnosis of autism or displayed autistic features. The third group comprised 25 typically developing children and adolescents who attended mainstream schools.

The two clinical groups were matched pairwise on chronological age (CA) and verbal mental age (VMA), the latter of which was calculated using the long form of the British Picture Vocabulary Scale (BPVS) (Dunn, Dunn & Whetton, 1982). The fact that the two clinical groups were matched for both CA and VMA entailed that they were also matched for verbal IQ (VIQ) on the BPVS. The typically developing group was matched pairwise on VMA only. Table I summarizes mean CA, VMA, and VIQ for each group.

Materials

False-Belief and Physical-State Tasks

Four stories and four nonnarrative scenarios were prepared. For each of the four stories and the four scenarios, a false-belief question, a physical-state question, and a memory question were prepared. Thus, each of the stories and the scenarios could be used as either the basis of a test of false-belief understanding or the basis of a test of CCR about changed physical states rather than about another person’s beliefs. The scenarios were included in addition to the story-based material to rule out the possibility of a relationship between false-belief and physical-state task performance occurring as a result of children’s ability or willingness to follow a narrative.

The four stories were based on those used by Riggs et al. (1998) and described, respectively, a chocolate bar being moved in the protagonist’s absence, a child’s painting being blown into a tree in her absence, a shop assistant leaving work to go to the doctor in the absence of a colleague, and a poorly husband leaving the house to help put out a fire while his wife was out buying medicine. Each story was acted out using playmobil figures (plastic figures of about 3 inches in height made by Lego) and various props (e.g., miniature tables and chairs, a bed, a bedside cabinet, and a tree). One of the stories is reproduced below, with the false-belief question, the physical-state question, and the memory question that were used with that story (the three other stories can be found in Appendix A).

Chocolate Story

John and Mary are in the kitchen. They have some chocolate. They put the chocolate in the fridge. Then John leaves the house to go and visit a friend. Mary wonders to herself, “What will I do today? I know,” she says, “I’ll bake a cake.” She goes to the fridge and gets the chocolate. She makes a cake with some of the chocolate and puts the rest of it away in the cupboard. Then she eats the cake, “yum yum.” Does John know where the chocolate is? John has finished visiting his friend now. He’s hungry and wants to eat some chocolate.

False-belief question: Where does John think the chocolate is?

Physical-state question: If Mary had not baked a cake, where would the chocolate be?

Memory question: Where was the chocolate in the beginning?

The four nonnarrative scenarios all involved the unexpected transfer of an object or the unexpected change of an object’s property. They were acted out using a cuddly pink hippo, a green frog hand puppet, plasticine, coloring pens, a card, toy bricks, a toy dog, a miniature chair, a book and a box. One of the scenarios is outlined below, with the false-belief question, the physical-state question, and the memory question that were used with that scenario (the three other scenarios can be found in Appendix B).

Coloring Task

This is my friend George (a pink cuddly hippo). We’ve brought some things to school. We’ve got a square (1 x 1 cm) and a pen. What color is the square? (yellow). What color is the pen? (Blue). George is going to do some coloring in a bit later on, but right now he is going to have a sleep in my bag. (George placed in bag.) I know, I could color in the square with my coloring pen (square colored in).

Does George know that we’ve colored in the square?

Now, can you tell me:

False-belief question: What color does George think the square is?
Physical state question: If I had not colored in the square with my coloring pen, what color would the square be?
Memory question: What color was the square in the beginning?

Inferred Belief, Not Own Belief, and Explicit False-Belief Tasks

The three stories quoted in the Introduction were used. The stories were taken from Wellman and Bartsch (1988), and each involved a protagonist looking for a desired object or animal that is in one of two locations. Each story was illustrated in a comicstrip fashion, using three drawings per story, with the story written underneath as captions (see Fig. 1).

Procedure

Participants were individually tested in a quiet room within school after a minimum of two visits: one for familiarization and one in which language ability was assessed. The experimental tasks were given across two sessions that were not more than 1 week apart.

Participants were tested on two story-based tests of false belief and two scenario-based tests of false belief. They were also tested on two story-based “physical state” tasks (i.e., tasks that required counterfactual conditional reasoning but did not require any understanding of belief) and two scenario-based physical-state tasks. All conditions, stories and orders of presentation were counterbalanced across sessions and across participants. Following the presentation of a story or a scenario, either a false-belief question or a physical-state question was asked. In either case, a memory question was also asked. All participants answered the knowledge and memory questions satisfactorily, and no participants were excluded from the analysis.

The inferred belief, not own belief, and explicit false-belief tasks were given at the end of the second test session. The order of presentation of these tasks was counterbalanced across participants. Participants either listened to the stories being read by the experimenter or read along with the experimenter. Attention was drawn to the color drawings that illustrated the key events in each story.

RESULTS

Initial data analysis found no effect of specific story in the narrative tasks ($\chi^2(1) = 0.99, n = 75, p > .05$) or of specific scenario in the scenario-based tasks ($\chi^2(1) = 0.98, n = 75, p > .05$). To rule out the possibility of interference effects between false-belief and physical-state questions, a between-subjects analysis of physical-state task performance of the participants with autism who received the task first as compared to those who received the task second was undertaken. No significant difference was found [$t(23) = 0.41, p > .05$]. A comparison was also made between the participants with autism who received the false-belief task first and those who received it second. Again, no significant difference was found [$t(23) = 0.18, p > .05$].

Fig. 1. The explicit false-belief task (after Wellman & Bartsch, 1988).
Table II. Mean Scores per Group for False-Belief and Physical-State Tasks (standard deviations are shown in brackets)

<table>
<thead>
<tr>
<th></th>
<th>False-belief task</th>
<th>Physical-state task</th>
</tr>
</thead>
<tbody>
<tr>
<td>Autism</td>
<td>2.08 (1.85)</td>
<td>2.56 (1.50)</td>
</tr>
<tr>
<td>Mental retardation</td>
<td>3.80 (0.50)</td>
<td>3.80 (0.58)</td>
</tr>
<tr>
<td>Typically developing</td>
<td>3.96 (0.20)</td>
<td>3.96 (0.20)</td>
</tr>
</tbody>
</table>

For the story tasks, each participant gained two scores between 0 and 2 according to the number of correct answers to the test questions in the false-belief and physical-state tasks. The same scoring procedure was used for the scenario tasks. Initial analysis showed that there was no effect of format type (story or scenario; t(74) = −.92, p > .05), and the data from the two formats was combined. Each participant, therefore, had two scores between 0 and 4: one for false belief and one for physical state. Scores are shown in Table II.

One 3 (Group) × 2 (task type: false-belief/physical state) analysis of variance was used to compare the participants' ability to pass false-belief and physical-state tasks. There was a significant effect of group [F(2,72) = 20.61, p < .0001] and a significant effect of task type [F(1,72) = 7.58, p < .01]. There was also a significant task-by-group interaction [F(2,72) = 7.58, p < .0001]. A Tukey test revealed that effect of group was caused by the group with autism scoring significantly lower than the MMR group and the typically developing group (p < .05). There were no differences in performance for either the MMR group or the typically developing group across the false-belief and the physical-state tasks. No further analysis of the MMR group and the typically developing groups was undertaken.

Table III illustrates the pattern of performance for the CWA on the physical-state tasks against the false-belief tasks. Eleven of the participants either passed all or failed all of the false-belief and physical-state tasks. A further four participants who did not pass all the tasks passed at least three false-belief and three physical-state tasks. Seven participants passed one false-belief task and one physical-state task. Three participants had a more variable performance between tasks: two passed two or three physical-state tasks but no false-belief tasks, whereas one passed two physical-state tasks and one false-belief task. Overall, 10 participants performed better on the physical-state tasks than false-belief tasks, and one child performed better on the false-belief tasks. This difference is significant (sign test, 10:1; p < .05). There was a significant correlation between performance on the false-belief and physical-state tasks for CWA [r(75) = .90, p < .001]. This correlation remained significant when VMA, VIQ, and CA were partialled out [r(20) = .90, p < .001].

To rule out the effect of chance responding on the present results, the data were also scored and analyzed using a criterion that only four out of four correct responses constituted a pass for each of the tasks. Using this criterion, the probability of scoring four out of four correct by chance alone (or, similarly, zero of zero) is only 6.25%. Table IV shows the number of participants who had perfect scores (4/4) and who were therefore classified as having "passed" the false-belief task and the physical-state task. Using the strict "perfect score" criteria, it was found that all of the groups performed better than chance (binomial, p < .05). A significant correlation was still found between the performance of the CWA on the false-belief and physical-state tasks [r(25) = .76, p < .0001]. There were no differences between group frequencies on the number of participants passing the inferred-belief, not-own-belief, and explicit false-belief tasks (Fisher's exact, p > .05). The group with autism performed near ceiling level on all three tasks (23 [92%] passing inferred belief; 24 [96%] passing not own belief; and 23 [92%] passing explicit false belief). Both the MMR group and the typically developing group were at ceiling on each of the tasks (25 [100%] passing).

Table III. Number of Physical-State Tasks Passed Compared with Number of False-Belief Tasks Passed by Children with Autism

<table>
<thead>
<tr>
<th>Number of false belief tasks passed</th>
<th>Number of physical-state tasks passed</th>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>9</td>
<td></td>
</tr>
</tbody>
</table>

Table IV. Number (and percentage) of Participants in Each Group Achieving a Perfect Score (four of four) on False-Belief and Physical-State Tasks

<table>
<thead>
<tr>
<th></th>
<th>False-belief</th>
<th>Physical-state</th>
</tr>
</thead>
<tbody>
<tr>
<td>Autism (n = 25)</td>
<td>10 (40%)</td>
<td>11 (44%)</td>
</tr>
<tr>
<td>Mental retardation (n = 25)</td>
<td>21 (84%)</td>
<td>22 (88%)</td>
</tr>
<tr>
<td>Typically developing (n = 25)</td>
<td>24 (96%)</td>
<td>24 (96%)</td>
</tr>
</tbody>
</table>
DISCUSSION

The aim of this study was to assess the roles of counterfactual conditional reasoning, belief understanding, and inferential reasoning in contributing to CWA's difficulties in passing false belief tasks. We hypothesised (1) that CWA will be impaired relative to controls on physical-state tasks involving CCR as well as on standard false-belief tasks that, on our analysis, also involve a form of CCR; (2) that CWA will show normal performance, relative to controls, on nonstandard false-belief tasks; (3) that CWA will perform worse on standard false-belief tasks than on physical-state CCR. Hypothesis 1 was supported. CWA performed worse than the control groups on the physical-state tasks as well as on the standard false-belief tasks. Moreover, a significant correlation was found between the ability to pass physical-state CCR tasks and performance on standard false-belief tasks in CWA, confirming that the cognitive demands of the two types of task are importantly related. The nature of the relationship is, however, unclear. It is unlikely that narrative skills underlie the observed correlation, as not all the tasks had a narrative format. It is also unlikely that the relationship is caused by memory demands, as all the children answered the memory questions correctly. It remains possible that some other nonspecific factor underlies the correlation. However, it is not obvious what this factor might be, especially given that the correlations remain high when VMA is partialled out.

If the correlation is not an artifact of nonspecific factors shared by the two types of task, the question arises as to what task-specific factors might underlie the observed correlations. It could be suggested, for example, that CWA have difficulty inhibiting their own true beliefs/knowledge of the world in both types of task. Roth & Leslie (1998) showed that CWA are able to ignore a marble placed first in a basket and then moved to a box when asked a question about the location of an identical marble that was previously placed in an identical basket and then placed behind a screen. Roth and Leslie conclude from this that CWA do not have a problem inhibiting a salient perceptual cue. Other evidence, however, indicates that CWA do have difficulty in inhibiting responses to salient perceptual information (see Ozonoff, 1997, for a review). Clearly, the details of individual tasks (and the developmental level of the individuals tested) need to be taken into consideration if one is to argue that impaired inhibition is involved in the difficulties that CWA (and typically developing young children) have with CCR.

Alternatively, it might be suggested that CWA have difficulty in imagining a counterfactual state of the world and in running the simulation, "what would happen if" (cf. Harris, 1995). Harris (1995) suggested that CWA may have difficulty in forming mental models, and there is plentiful evidence of a lack of imaginative fluency in CWA (see, e.g., Craig & Baron-Cohen, 1999; Turner, 1999). However, in an experiment that is relevant both to the suggestion that CWA have a problem with counterfactual imagining, and to the suggestion that CWA have a problem in inhibiting their own factual knowledge, Jarrold, Boucher, & Smith (1994) showed that CWA have no problem using everyday objects counterfactually in pretend play (e.g., using a rectangular eraser as a book, or a pencil as a toothbrush). We conclude, therefore, that it is unclear which of the primitive cognitive processes involved in CCR might be defective in children with autism.

A recent study by Scott, Baron-Cohen, and Leslie (1999) reported that CWA do not have difficulties with counterfactual reasoning. These authors found that CWA performed very well (compared to verbal age–matched controls) on syllogistic reasoning tasks with false premises; for example, "All cats bark, Rex is a cat. Does Rex bark?" Scott et al. interpreted their findings as evidence that counterfactual reasoning ability is intact in CWA. However, there are grounds for doubting this interpretation. Moreover, even if Scott et al.'s interpretation of their findings were valid, we would argue that their study involved counterfactual deductive reasoning rather than counterfactual conditional reasoning. In counterfactual deductive reasoning, answers to questions follow necessarily from the information/promises given. In counterfactual conditional reasoning, in contrast, the answer we give is at best likely or probable (Sally believes the marble is in the basket, because from her point of view, it is most likely the case—though by no means certain—that it stayed put).

Hypothesis 2 was also supported. CWA performed as well as controls on the nonstandard false-belief tasks. This calls into question the assumption that CWA fail standard false-belief tasks because they do not understand that people act on their beliefs, even when these are false. However, as all three groups were at or near ceiling on the nonstandard false belief tasks, it

---

3 A correct answer could also be given using a process of anaphoric reference at a syntactic level (Harris, 1995: see also Harris & Leevers, 2000).
may be the case that performance on nonstandard belief tasks reflects a partial competence and that CWA's understanding of belief falls short of that of other children in some way that might emerge on a more sensitive test. Moreover, Perner (1989) argued that Wellman and Bartsch's inferred-belief and not-own-belief tasks could be passed using "think" and "see" as clues to reality. This may be the case. However, in the explicit false-belief task, the actual state of affairs is explicitly stated, and the child has to understand that people act on their beliefs, even if these beliefs are false, and to use this rule to predict the protagonist's behavior. The "explicit false-belief task" therefore provides the strongest test of CWA's understanding of belief, and they performed near ceiling on this task.

Hypothesis 3 was also supported: The children with autism performed worse on the standard false-belief tasks than on the physical-state CCR tasks. We therefore tentatively conclude, following the suggestion of Riggs et al. (1998), who reported a similar finding in typically developing children, that the critical difference is that in the physical-state tasks the counterfactual proposition is made explicit (e.g., "If Mary had not baked a cake"), whereas in the standard false-belief tasks the protagonist's false belief must be inferred. Here again, it is unclear exactly what is involved, cognitively, in working out for oneself explicit knowledge of Sally's false belief. Peterson and Bowler (2000) suggest that the underlying problem is a generativity deficit, arguing that in standard false-belief tasks, Sally's false belief "that the marble is in the basket" must be generated, whereas in physical-state CCR tasks the counterfactual proposition is explicitly given. Generativity may be involved in tasks in which the counterfactual proposition or false belief is not explicit. However, other cognitive processes are also involved, and it is not possible from the available evidence to pinpoint which of these processes might be the source of the CWA's difficulties. In particular, we suggest that inferential processing is required to generate the protagonist's (e.g., Sally's) false belief. The problem does not appear to relate to understanding that "seeing is knowing/believing," because the CWA in our study performed well on a task in which they had to infer a protagonist's false belief from the statement "Jane saw her colored pencils on the desk." Rather, the problem may lie before that, in inferring that Sally's action in placing the marble in the basket would lead to a belief that the marble was in the basket. One wonders how CWA might perform on a modified version of the Sally-Ann task in which the words "Sally sees her marble in the basket" were added at the point at which Sally places the marble in the basket.

In sum, it appears that autistic children's difficulties in passing standard false-belief tasks derive from the cognitive requirements of the tasks and not from a flawed conceptual understanding of belief. Moreover, the cognitive difficulties that CWA experience on standard false-belief tasks are associated both with the particular cognitive demands of counterfactual conditional reasoning and with drawing inferences/generating propositions where critical information is not made explicit.

If, as we suggest, CWA's difficulties in passing standard false-belief tasks derive from the cognitive requirements of the tasks and not from a flawed conceptual understanding of belief, it is relevant to consider why CWA are impaired on physical-state CCR tasks but not on versions of the "false-photo" task (Leekam & Perner, 1991; Leslie & Thaiss, 1992). In the false-photo task participants are asked to predict what picture will appear in a polaroid photograph taken of a scene (e.g., a marble on a table), which is then altered (e.g., the marble is moved to a box). The false-photo paradigm, like the physical-state CCR paradigm, does not involve any understanding of belief. Leslie & Roth (1993) argued that the general problem-solving characteristics of false-photo and false-belief tasks are highly similar. However, it is doubtful whether false-photo tasks involve either counterfactual conditional reasoning or the ability to infer a counterfactual state of affairs. Instead, the false-photo task requires an understanding that Polaroid cameras generally take a picture of whatever they are pointing at when the button is clicked. To succeed on the task, participants with this knowledge simply have to remember and to report what the camera was pointing at when the button was clicked. They do not have to inhibit or ignore the fact that the marble has since been moved. Equally, they do not have to simulate a situation in which the marble was not moved. Evidence against Leslie and Roth's (1993) claim that false-photo and false-belief tasks require similar problem-solving abilities comes from a report by Perner, Leekam, Myers, Davis & Ojodgers (2000). These authors found no correlation between performance on false-belief and false-photo tasks, indicating—that for typically developing children at least—the two tasks have different information-processing characteristics. We would argue that physical-state CCR tasks also differ from the false-photo task in their problem-solving requirements, and that this is why CWA are impaired on the former but not the latter.
The claim that people with autism have problems engaging in the type of problem solving required to pass standard false-belief tasks is not a new one; suggestions of this type having been made by Frye et al. (1995), Tager Flusberg (2000), and Peterson & Bowler (2000). Frye and his colleagues (Frye et al., 1995; Frye, Zelazo, & Burack, 1998; Zelazo & Frye, 1998) argue that developmental changes in reasoning with embedded rules (if-if-then) account for developmental changes in perspective taking and thus in the ability to ascribe false beliefs or to consider counterfactual alternatives. These authors suggest that CWA have problems with “if-if-then” reasoning, and that this relates to their impaired performance on standard false-belief tests. Somewhat similarly, Tager Flusberg (2000), following de Villiers and de Villiers (1999) suggests that the ability to understand embedded sentential complements in natural language is closely related to the ability to pass standard false belief tasks, and that both abilities may be impaired in CWA. Recently, Smith, Apperly, & White (2003) have narrowed down still further the type of syntactic competence most closely related to success on standard false-belief tasks in typically developing young children. These authors identify as critical the ability to understand what they term “double-event clauses” in which the object of the superordinate clause becomes the subject of the embedded clause (e.g., “The girl kicked the man who jumped over the wall”). Smith et al. (2003) suggest that it is the need to view an event from multiple perspectives that links the acquisition of such double-event clauses with the development of false-belief reasoning.

Peterson and Bowler (2000) carried out standard tests of false-belief and physical-state tests of CCR tasks with CWA and reported results similar to ours. Specifically, the CWA were impaired on both the physical-state tasks and on the standard false-belief tasks, but were more impaired on the latter than on the former. Peterson and Bowler concluded that counterfactual conditional reasoning is necessary for success on standard false-belief tasks. Our conclusion, on the basis of similar evidence, is more cautious than this: namely, that physical-state and standard false-belief tasks both involve CCR, and that one or more of the primitive cognitive processes involved in CCR may be impaired in autism. Peterson and Bowler also conclude, as we do, that the difference in CWA’s performance on physical-state and standard false-belief tasks has to do with the fact that the counterfactual proposition “if X” is made explicit in physical-state tasks but not in standard false-belief tasks. This conclusion is strengthened in our study by the CWA’s excellent performance on nonstandard false-belief tasks in which the protagonist’s false belief was made explicit. The inclusion of nonstandard false-belief tests also enables us to rule out the role of flawed belief understanding as a cause of the discrepancy between the CWA’s performance on physical-state CCR tasks and standard false-belief tasks. However, although agreeing with Peterson and Bowler that the difference in CWA’s performance on physical-state and standard false-belief tasks has to do with whether or not the counterfactual proposition/false belief is made explicit, our interpretation of this observation is, again, a little more cautious than theirs.

It is unparsonious to argue that one cognitive deficit underlies CWA’s problems with counterfactual conditional reasoning and that some additional cognitive deficit underlies their problems in tasks in which counterfactual propositions or false beliefs are not made explicit. In the longer term, therefore, a single explanation of these two observations must be sought. It seems likely that the convergent approaches of Frye and his colleagues, Tager-Flusberg and the de Villiers, and most recently, Smith, will provide a theoretical understanding of the problems of counterfactual conditional reasoning demonstrated in this study and in the study by Peterson and Bowler. It is less obvious how problems of inference and generativity might be accommodated with such a theoretical framework. Nor is it obvious how the cognitive deficits that underlie CWA’s impaired performance on standard false-belief tasks may be seen to relate to executive dysfunctions, although problems of inhibition and generativity would bring them under this broad umbrella. Of the other mainstream theoretical explanations of autism, our findings are descriptively compatible with Minshew’s suggestion that autism is caused by problems of complex information processing (Minshew & Goldstein, 1998). In as far as we have implicated impairments in the extraction of meaning from narrative events, there may also be links with the weak central coherence theory of autism (Frith, 1989; Happe, 1997). Our replication of Sparrevoorn and Howie’s (1995) findings strengthens their conclusion that CWA’s impairment on standard false-belief tasks does not stem from a flawed conceptual understanding of belief. This may bring us into conflict with some authoritative theorists in the field. However, we would point out that we are claiming only that CWA who pass nonstandard false belief tasks have sufficient understanding of belief to pass standard false-belief tasks. We are not claiming that their understanding of belief is in all respects normal.
ACKNOWLEDGMENTS

This research was submitted as partial fulfilment for the degree of Ph.D. by Cathy Grant in the Department of Human Communication Sciences, University of Sheffield. The research was supported by studentship from the University of Sheffield. We thank the staff and students of the following schools for participating in this study: Abbots Lea, Liverpool; Foxwood, Nottingham; Peterhouse, Southport; Rosehill, Nottingham; St. Joseph’s Junior School, Kirkby; Wargrave House, Newton le Willows. We would also like to thank Dr. Guy M. Sutton for the use of his drawings.

APPENDIX A

Three other stories (additional to the “chocolate story” given in the text) that were used to assess physical-state counterfactual conditional reasoning and false belief.

Post Office Story

Sally and Peter are in their house. Peter isn’t feeling very well, so he decides to go to bed. Sally tucks Peter in and makes sure he is warm and comfortable. Then Sally leaves the house to go to the shops. Peter is fast asleep—“zzzzz.” Suddenly the phone rings and the man from the post office says, “Please Peter, come quickly and help us—the post office is on fire!” Peter leaps out of bed and hurries to the post office. When he gets there he helps to put out the fire.

Does Sally know where Peter is?
Sally has finished shopping now. She’s got some medicine for Peter.

Physical-state question: If there had been no fire, where would Peter be?
False-belief question: Where does Sally think Peter is?
Memory question: Where was Peter in the beginning?

Picture Story

Jenny is in her garden. She’s painting a picture. Now she’s finished painting her picture and it’s time to go to school. She leaves the picture on the table and goes off to school. While she’s at school a big wind blows and blows her picture up into the tree—“wheeeosh.”

Does Jenny know where her picture is?

Jenny has finished school now. She’s coming back to see her picture.

Physical-state question: If the wind hadn’t blown, where would Jenny’s picture be?
False-belief question: Where does Jenny think her picture is?
Memory question: Where was the picture in the beginning?

Shop Story

Andrew and Rebecca are in their shop. They are working there all day. Rebecca decides to go home for lunch. She says “Bye bye Andrew, I’ll be back later.” Andrew stays in the shop to work and eats an apple—“crunch crunch.” But ohhh, it’s a bad apple and it gives him a sore tummy. So Andrew leaves the shop and goes to the doctor to get some medicine.

Does Rebecca know where Andrew is?
Rebecca has finished her lunch now. She’s coming back to work with Andrew.

Physical-state question: If Andrew had not eaten the apple, where would he be?
False-belief question: Where does Rebecca think Andrew is?
Memory question: Where was Andrew in the beginning?

APPENDIX B

Three other scenarios (additional to the “coloring scenario” given in the text) that were used to assess physical-state counterfactual conditional reasoning and false belief.

Bricks Scenario

This is my friend Tidy Teddy (green frog glove puppet). Can you see we’ve brought some things into school today (bricks, cup, and box). Tidy Teddy likes to tidy up. He’s tidying up the bricks and putting them into the box. Now, Tidy Teddy is tired and he’s going to go to sleep in my bag (Tidy Teddy placed in bag). I know. We could play with the bricks couldn’t we? We could play with the bricks and pour them into the cup.

Does Tidy Teddy know that we’ve played with the bricks?
Tidy Teddy is awake now, and he wants to play with his bricks.
Counterfactual and Mental State Reasoning in Children with Autism

Now, can you tell me:

**False-belief question:** Where does Tidy Teddy think the bricks are?

**Physical-state question:** If I had not played with the bricks, where would they be?

**Memory question:** Where were the bricks in the beginning?

---

**Dog Scenario**

This is my friend Tidy Teddy. Can you see we’ve brought some things into school today (book, toy dog, and chair). Tidy Teddy likes to tidy up. He’s tidying up and putting the dog onto the chair. Now, tidy teddy is tired and he’s going to go to sleep in my bag (Tidy teddy placed in bag). I know. We could play with the dog couldn’t we? We could pick up the dog and place him on the book.

Does Tidy Teddy know that we’ve picked up the dog?

Tidy Teddy is awake now, and he wants to play with the dog.

Now, can you tell me:

**False-belief question:** Where does Tidy Teddy think the dog is?

**Physical-state question:** If I had not picked up the dog, where would he be?

**Memory question:** Where was the dog in the beginning?

---

**Shape Scenario**

This is my friend George. Can you see we’ve brought some things to school? We’ve got some plasticine. What shape is the plasticine? (round). George is going to play with his plasticine a bit later on, but right now he is going to go into my bag to go to sleep (George placed in bag). I know. I could play with the plasticine. I could roll it out into this shape, what shape is the plasticine now? (sausage).

Does George know that we’ve played with the plasticine?

Now, can you tell me:

**False-belief question:** What shape does George think the plasticine is?

**Physical-state question:** If I hadn’t played with the plasticine, what shape would it be?

**Memory question:** What shape was the plasticine in the beginning?

---

**REFERENCES**


