Linking Theory of Mind and Central Coherence Bias in Autism
and in the General Population

Christopher Jarrold, David W. Butler, Emily M. Cottington, and Flora Jimenez
University of Bristol

Three experiments investigated whether 2 characteristic aspects of the psychological profile of autism, theory-of-mind deficits and weak central coherence, might be functionally related. Experiment 1 showed that in the general population, performance on a proposed test of theory of mind was inversely related to speed on the Embedded Figures Test, a measure of central coherence bias. Experiments 2 and 3 confirmed that poor theory-of-mind performance was linked to weak central coherence among typically developing children and among children with autism; however, the correlations between these measures were reliable only after accounting for differences in individuals' verbal mental ages. This pattern of results is interpreted in terms of a relationship between individual differences in theory of mind and central coherence bias, a relationship that is separate from any developmental differences in these domains.

In recent years, three separate psychological accounts have emerged as potential explanations of differing aspects of autism. It has been argued that autism is associated with, in turn, a theory-of-mind deficit (e.g., Baron-Cohen, 1995; Baron-Cohen, Leslie, & Frith, 1985), executive dysfunction (e.g., Ozonoff, 1995; Pennington & Ozonoff, 1996: Russell, 1997), and a weak drive for central coherence (Frith & Happé, 1994). Arguably the most influential of these accounts is the theory-of-mind deficit hypothesis, which suggests that individuals with autism are severely delayed in their acquisition of a folk psychology—the ability to ascribe beliefs and desires to others in order to predict their behavior. Evidence for this delay comes from a multitude of empirical studies, including those that have assessed individuals' ability to attribute false beliefs to others (e.g., Baron-Cohen et al., 1985; Perner, Frith, Leslie, & Leekam, 1989). The majority of individuals with autism fail such tasks, despite having mental ages that would typically be sufficient to ensure success (cf. Perner, Leekam, & Wimmer, 1987; Wimmer & Perner, 1983). Even the minority of individuals with autism who do pass these tests have been found to be impaired on more complex, or second-order, tests of false-belief understanding (Baron-Cohen, 1989), which has led to the suggestion that theory-of-mind development is markedly delayed in virtually every individual with autism. Strong support for this claim comes from an analysis of previous studies carried out in this area by Happé (1995). She showed that the probability of passing a standard theory-of-mind task was clearly related to level of development, as measured by verbal mental age, in both children with autism and children with mental handicaps (see also Sparrow & Howie, 1995; Yirmiya, Erel, Shaked, & Solomonica-Levi, 1998; Yirmiya, Solomonica-Levi, Shulman, & Pilowsky, 1996); however, for the children with autism, the mental age level required for likely success on the task was markedly higher than it was for the control children.

The strength of the theory-of-mind hypothesis is that it is well placed to explain many of the behavioral symptoms of autism. Diagnosis of autism is currently based on a triad of impairments, in socialization, communication, and imagination (American Psychiatric Association, 1994; Wing & Gould, 1979). It is argued that social withdrawal is an understandable consequence of having theory-of-mind problems, that is, of lacking the ability to explain otherwise confusing behavior in terms of underlying mental states (Baron-Cohen, 1992, 1995; Frith, 1989a, 1989b; see Frith, Happé, & Siddons, 1994). Similarly, one needs to understand that others have mental states that can differ from one's own in order to be motivated to communicate. Sperber and Wilson's (1986) analysis of communication as requiring an appreciation of another's knowledge and ignorance in order to be relevant indicates how a theory-of-mind deficit would lead to problems in this domain (Frith, 1989a; Happé, 1993). Finally, deficits in imagination, or in pretend play at least, would occur as a result of a theory-of-mind deficit if pretend play requires the same representational processes as does attributing beliefs to others. Leslie (1987) made exactly this argument, claiming that pretense, like the appreciation of beliefs, requires the representation of another's propositional attitude toward the state of the world (though see Harris, 1991; Jarrold, Carruthers, Smith, & Boucher, 1994; Lillard, 1993).

However, the theory-of-mind account struggles to explain some aspects of autistic behavior. In addition to problems in pretend play (Jarrold, Boucher, & Smith, 1993), individuals with autism show other deficits in imagination, as evidenced by repetitive behaviors...
and a preference for stereotyped routines. This aspect of the triad of impaired behaviors might best be explained in terms of executive dysfunction. This is the failure to guide behavior with reference to internally specified goals as opposed to external stimuli (Norman & Shallice, 1986), which is often seen in patients with frontal lobe damage. Individuals with autism are also thought to suffer from executive dysfunction, and empirical evidence for this suggestion comes from two areas. First, individuals with autism have been shown to perform poorly on the tests that are used to index executive problems in frontal lobe patients (e.g., Hughes, Russell, & Robbins, 1994; Ozonoff, Pennington, & Rogers, 1991; Prior & Hoffman, 1990). Second, more direct studies of executive control have indicated that individuals with autism may have problems in inhibiting prepotent responses to salient external stimuli (Hughes, 1996; Hughes & Russell, 1993; Ozonoff & Strayer, 1997; Russell, Mauthner, Sharpe, & Tidwell, 1991). Clearly, a deficit in self-control of action might well explain the repetitive and stereotyped behaviors seen in autism (Turner, 1997) as well as the problems in pretend play experienced by individuals with autism (Jarrold, 1997).

Individuals with autism also show characteristic strengths in certain areas that the theory-of-mind account cannot readily explain. These include relatively good performance on certain visuo-spatial or constructive subtests of the Wechsler scales (Wechsler, 1974, 1981), such as Block Design and Object Assembly (see Happé, 1994c; Shah & Frith, 1993). Frith and Happé (1994; Frith, 1989b) suggested that these peak abilities reflect a weak drive for central coherence associated with autism. They defined central coherence as the normal tendency to integrate local information in the search for global meaning, a tendency to focus on the whole rather than the parts of any stimulus. They suggested that individuals with autism, in contrast, are more likely to focus on the local rather than the global level.

Evidence for weak central coherence in autism comes from a number of sources. Individuals with autism have been shown to be faster than matched controls at the Embedded Figures Test (Witkin, Otman, Raskin, & Karp, 1971), in which individual shapes have to be found within a larger pattern (Jolliffe & Baron-Cohen, 1997; Shah & Frith, 1983; though see Brian & Bryson, 1996; Ozonoff et al., 1991). Shah and Frith (1993) showed that individuals with autism were faster at the standard block design task when the design to be completed was segmented into its constituent parts. Under these conditions, the speed of the control children increased to match that of the children with autism, which suggests that the individuals with autism "presegmented" the design in the standard task, focusing on the local parts rather than the whole. Other phenomena that have been ascribed to weak central coherence in autism include a failure to disambiguate homographs using surrounding word context (Frith & Snowling, 1983; Happé, 1997), a failure to enumerate canonical patterns when counting (Jarrold & Russell, 1997), and a tendency not to perceive visual illusions (Happé, 1996; though see Ropar & Mitchell, 1999).

An obvious and important question is whether these three psychological explanations are in any way related. There has been some discussion as to whether executive deficits might be a fundamental cause of theory-of-mind problems in autism, or vice versa (Carruthers, 1996; Ozonoff et al., 1991; Russell, 1996). Potential links between theory-of-mind deficits and central coherence bias have also been considered (Frith, 1989b; Happé, 1994b), but these domains are typically viewed as being separate from one another. There are perhaps two theoretical reasons for this. First, the two accounts appear to explain very different sets of behaviors; theory-of-mind deficits appear to account for many of the triad features of autism, whereas weak central coherence explains a more narrow set of nontriad skills. Second, Baron-Cohen (1995) and Leslie (1987; Leslie & Roth, 1993; Leslie & Thaiss, 1992) argued that theory of mind is a modular ability, that is, an ability that rests on the functioning of a dedicated and fixed neural system (Fodor, 1983). Modular systems are domain specific by definition, and although their output is available to domain-general processes, the working of a modular system is automatic, encapsulated, and not open to top-down influences from other systems. Consequently, from this theoretical standpoint, theory-of-mind functioning should not be linked to ability in any other domain (though it can of course affect development of those abilities that require a functioning theory of mind). Frith and Happé (1994) argued that there are also empirical reasons for viewing a theory-of-mind deficit and weak central coherence as being distinct aspects of autism. They point to the fact that some individuals with autism pass relatively complex second-order theory-of-mind tasks (which involve attribution of beliefs about beliefs) yet still show weak central coherence (Happé, 1994c; 1997). Though this might seem to suggest that these two domains are separable, it should be noted that even these individuals with autism are likely to have acquired theory-of-mind ability at a delayed rate given the arguments outlined above.

In fact, potential evidence for a link between theory-of-mind deficits and weak central coherence has emerged from recent studies that have used more complicated theory-of-mind tasks. Baron-Cohen and Hammer (1997b) presented 30 adults from the general population with a task in which intentions had to be read from pictures of individuals' eyes (Baron-Cohen, Jolliffe, Mortimore, & Robertson, 1997). In this task, participants had to judge, for example, whether an individual was concerned or unconcerned, or calm or anxious, on the basis of a photograph of the eye region of the face. Baron-Cohen and Hammer's sample consisted of equal numbers of men and women, and the 15 men were significantly more accurate on this test than the 15 women. In contrast, the men were faster than the women on the Embedded Figures Test. Baron-Cohen and Hammer therefore observed an opposite pattern of sex differences on their two tasks, one thought to be tapping theory-of-mind ability and the other known to be linked to central coherence bias. This finding implies that those individuals performing well on one task perform less well on the other, and Baron-Cohen and Hammer noted that their results suggested that "weak central coherence may go hand in hand with impaired mindreading" (Baron-Cohen & Hammer, 1997b, p. 530).

Some of the other work of Baron-Cohen and colleagues appears to support this claim. Baron-Cohen, Jolliffe, et al. (1997) reported a similar sex difference (women superior) on the same "eyes" task among a separate sample of 50 adults. Baron-Cohen and Hammer (1997b) also included a sample of 30 parents of children with Asperger's syndrome in their study, and they found the same pattern of sex differences on the eyes task and the Embedded Figures Test in this group. Much of this work is summarized in an article by Baron-Cohen and Hammer (1997a) in which they suggested that rapid Embedded Figures Test performance (and, by
implication, weak central coherence) and relatively poor theory of mind may reflect a male brain type. According to their argument, individuals with autism, who show poorer theory-of-mind ability and more rapid Embedded Figures Test performance than any of these groups (Baron-Cohen, Jolliffe, et al., 1997; Jolliffe & Baron-Cohen, 1997), represent an extreme form of this male brain profile.

The results of these studies are exciting because they suggest that two important psychological aspects of autism, which have previously been viewed as distinct from one another, may instead be linked. This in turn has important implications for the proposed modularity of theory-of-mind abilities and raises intriguing questions about the nature of any link between these two areas. For example, one might ask whether weak central coherence produces theory-of-mind deficits, or vice versa. However, the findings from Baron-Cohen and colleagues’ studies provide only relatively indirect support for an association between these domains, because correlations between levels of performance on the Embedded Figures Test and on eyes tasks were not examined for any of these samples. Consequently, in our first experiment we aimed to replicate and extend the work of Baron-Cohen and colleagues by assessing and correlating performance on the eyes task and performance on the Embedded Figures Test in a large sample of adults from the general population.

Experiment 1

Method

Thirty male and 30 female undergraduate university students, all between the ages of 18 and 25 years, took part in this experiment. Each individual was given two tasks, Baron-Cohen’s eye-reading task (Baron-Cohen, Jolliffe, et al., 1997) and the adult version of the Embedded Figures Test (Witkin et al., 1971). The eyes task used 24 stimuli,1 which were shown to the participant in succession. For each item, the participant was given a forced choice between two alternative mental state attributions, for example, between concerned and uninterested or between calm and anxious. The order in which the correct response was presented in the forced choice was counterbalanced across trials (see Baron-Cohen, Jolliffe, et al., 1997, for full stimuli details). The Embedded Figures Test consisted of 12 items. In each trial the participant was first shown a complex figure for 10 s, and then a card displaying the target shape was placed on top of the complex figure for a further 5 s. Next, the card displaying the target was removed, and the participant had to locate the target shape within the complex figure as rapidly as possible. The time taken (in seconds) to find the embedded target figure was recorded, although if participants took longer than 3 min on any particular trial they were credited with a score of 180 s for that trial, in line with the standard test procedure. Time taken to locate the target was used as the dependent variable for statistical analysis because it provides a much greater range in performance scores than does simply coding the number of items correctly identified. Although allowing a score of 180 s on trials when the target was not found does potentially introduce a bias into this measure, it is unlikely to alter the results to any real extent. It is rare for individuals not to find the target in this time period, and as Table 1 shows, the average time taken to find items was well below this 180-s cutoff.2

Results

The mean number of correct responses in the eye-reading task and the average time taken to complete an item of the Embedded Figures Test are shown in Table 1 for each of the two gender groups. Independent t tests showed that men and women did not differ significantly in performance on the eyes task, t(58) = 0.69, p = .52, or on the Embedded Figures Test, though the men showed a tendency to be faster than the women, t(58) = 0.98, p = .33. (P values for these and all subsequent tests are based on one-tailed tests of significance unless otherwise noted.) Across the whole sample of 60 individuals, there was a significant correlation between performance on the eyes task and average time taken on an embedded figure, r(58) = .30, p = .01, which implies that those individuals who performed better at the eye-reading task took longer to locate each embedded figure, and vice versa. When the gender groups’ performances are considered separately, the relationship between performance on the eye-reading and embedded figures tasks was significant for the women, r(28) = .50, p = .03, but not for the men, r(28) = .17, p = .13. Although the relationship between these measures was stronger among the women than among the men, the correlation for women was not significantly stronger than that observed for men (z = 0.64, p = .51, two-tailed; Howell, 1997, p. 261).

Discussion

Experiment 1 failed to replicate one aspect of Baron-Cohen and colleagues’ previous work in that no sex differences were observed on the eyes task. Our two gender groups in fact performed at very similar levels on this test, in contrast to the findings reported by Baron-Cohen, Jolliffe, et al. (1997) and Baron-Cohen and Hammer (1997b). A comparison of the current data with those presented in these two previous studies shows that the men assessed here performed at roughly the same level as the men in the previous studies (means of 19.23 vs. 18.8 and 19.5, respectively) but that the women performed less well than the women observed in the previous studies (means of 19.67 vs. 21.8 and 22.1, respectively). The absence of an effect in the current study cannot be attributed to a change in the task procedure, because the test was presented in the prescribed manner using the very same materials that were used previously, or to a particular lack of power, because the sample employed here was larger than any used in previous work. Though the result may therefore appear puzzling, it is worth noting that another of Baron-Cohen and colleagues’ studies failed to find a significant sex difference on a related task in which simple and

---

1 The stimuli used in the current study were identical to those used by Baron-Cohen and Hammer (1997b) and Baron-Cohen, Jolliffe, et al. (1997).

2 The same is true for subsequent experiments (see Tables 2 and 4).
The pattern of results seen on this task—poorer performance among a group of individuals with autism and Asperger's syndrome—matched that seen on another proposed high-level test of theory of mind, a "strange story" task (cf. Happé, 1994a). However, the controls used in this study were at ceiling on this latter task, and even if there was good evidence of an autism-specific deficit on these two tasks it would not prove that they assess the same ability.

There is also a clear theoretical reason for questioning the validity of the eyes task, and that is that the items employed in the test need not necessarily require the attribution of beliefs to others. Some of the test items require only appreciation of emotion (e.g., having a sad thought), whereas the majority involve the attribution of what might be termed states of mind (e.g., attraction, reflection, being nervous). Baron-Cohen, Joliffe, et al. (1997) called these complex mental states, but although this may be a reasonable description, it is not obvious that inferring these mental states requires the same representational processes involved in attributing false beliefs, for example. As noted above, in its strictest sense, a theory of mind involves constructing second-order (or meta-)representations—one needs to represent how another represents the world as being (Leslie & Roth, 1993; Perret, 1991). This involves representing the propositional attitude that another person takes toward a certain content (Leslie, 1987), as in, for example, "Jill believes (attitude) that the chocolate is in the cupboard (content)." The eyes task may well require the representation of an agent's attitude (e.g., she's anxious about something, she's certain), but it does not require the representation of the content of that attitude (e.g., she is anxious about her exams, she is certain that it is Monday). Without the representation of both attitude and content, one is not properly representing how another represents the world, and consequently this does not amount to metarepresentation (Perret, 1991; see also Currie, 1995; Jarrold et al., 1994). Put more simply, it is inferring the content of another's belief, and not just that they believe something, that forms the crucial test for a theory of mind (Dennett, 1978). Under this analysis, the eyes task cannot be viewed as a strict test of theory-of-mind (or, more accurately, metarepresentational) ability. One might still argue that it measures theory of mind in the broadest sense or that it taps theory-of-mind-related skills, though it is not clear how this would square with modularist conceptions of theory-of-mind acquisition (e.g., Baron-Cohen, 1995; Leslie, 1987).

The data from the present experiment therefore suggest that performance on the eyes task is linked to central coherence bias, but this in itself does not necessarily imply that theory of mind and central coherence are linked. Indeed, there is a more prosaic and arguably more parsimonious explanation of the negative correlation observed here. In developing the eyes task, Baron-Cohen, Joliffe, et al. (1997) argued that it had no central coherence component. However, one might readily argue that success on this task requires integrating a set of subtle visual cues, such as direction of the eyebrows, degree of eye closure, and direction of gaze, in order to obtain a global visual impression of the individual's mental state. If this is the case, then the task would benefit individuals with a global visual processing style. This would immediately explain the negative correlation with performance on the Embedded Figures Test, on which a local visual processing style is beneficial.

These arguments indicate that a proper test of a link between central coherence and theory-of-mind ability can best be carried out using theory-of-mind tasks that are (a) accepted tests of...
metarepresentational ability and (b) nonvisual in nature. A number of such tasks exist, but they are suitable for young children only. Consequently, in our second experiment we investigated theory-of-mind ability and central coherence bias in a sample of young children, using accepted tests of theory-of-mind ability. In this instance two central coherence tasks were used—the Embedded Figures Test and a block design test—to add convergent validity to the hypothesis under examination. One additional test that was also included was a measure of individuals’ verbal mental age. We assessed verbal mental age because, as discussed, previous research has shown a strong positive association between verbal mental age and theory-of-mind success in children (see Yirmiya et al., 1998). Similarly, speed on the Embedded Figures Test has been shown to increase with development and to be related to children’s level of verbal ability (Coates, 1972; Witkin, Goodenough, & Karp, 1967; Witkin et al., 1971). The importance of these points is that the links between age and general intelligence on the one hand, and theory of mind and Embedded Figures Test speed on the other, might mask the hypothesized inverse relationship between theory of mind and central coherence unless individual differences in age and intelligence are accounted for in a developmental sample. In addition, one might expect individuals who are older and/or functioning at a higher developmental level to be better at both block design and theory-of-mind tasks. Clearly, such superiority might lead to an artificial inflation of the correlation between theory of mind and block design performance. Consequently, in Experiment 2, age and verbal mental age were partialed out of the relationships between theory-of-mind ability and performance on the Embedded Figures Test and the block design task.

Experiment 2

Method

Twenty-four 5-year-olds (mean age = 67.63 months) drawn from a single-year group in a school took part in this experiment. The group consisted of 12 boys and 12 girls.

Individuals’ theory-of-mind ability was assessed with a series of tasks that differed in difficulty (adapted from Sparrevohn & Howie, 1995). These tasks included inferred belief, not-own belief, explicit false belief (cf. Wellman & Bartsch, 1988), own false belief (cf. Perner et al., 1987), other’s false belief, and second-order false belief (cf. Perner & Wimmer, 1985) tasks. Full procedural details of these tasks are given in the Appendix (see also Sparrevohn & Howie, 1995). The studies from which these tasks were taken demonstrated a clear developmental progression in the likelihood of children’s successfully completing each of these tasks, indicating that the tasks increase sequentially in difficulty in the order in which they are presented in the Appendix (see also Sparrevohn & Howie, 1995). Consequently, theory-of-mind performance was scored as the number of tasks completed by each participant (giving a possible range from 0 to 6).

We assessed central coherence bias using two separate tasks. First, all individuals were given the preschool version of the Embedded Figures Test (Coates, 1972). In this test, which mirrors the adult version, children have to locate a target figure in a series of more complex line drawings. In the preschool version, the target, a triangle, is the same for each trial, and the figures are less complex than those used in the adult version. Pilot work showed that children of this age found the task easy to understand if they were told to find the triangle in each figure and then trace it with their fingers while counting its sides out loud. There were 24 trials in the test, following a number of practice items, and children were given 30 s to locate the triangle on each trial. Their times to find the target were recorded, although if they took longer than 30 s, they were credited with a score of 30 for that item, in line with the standard test procedure. Second, individuals were given a block design test, the Pattern Construction subtest of the Differential Ability Scales (Elliot, 1990). In this task individuals are presented with a two-dimensional target figure, made up of black and yellow segments, which they have to recreate using blocks with faces consisting of different arrangements of black and yellow designs. Performance is measured as an ability score, which is derived on the basis of the number of items correctly completed coupled with the time taken to complete each item. Because of unavoidable absences from school at the time of testing, 4 individuals could not be assessed on this block design test. Data are therefore available for only 20 of the children in the sample for this measure.

Finally, all children were given a measure of verbal mental age, which was in fact a test of receptive vocabulary—the long form of the British Picture Vocabulary Scale (BPVS; Dunn, Dunn, Whetton, & Pintilie, 1982).

Results

Table 2 presents summary descriptive statistics for the measures assessed in this experiment, split by gender group. Independent t tests showed that the boys and girls did not differ significantly in age, t(22) = 1.16, p = .26 (two-tailed), in their performance on the battery of theory-of-mind tasks, t(22) = 0.70, p = .25, or on the Embedded Figures Test, t(22) = 1.28, p = .11. Performance on the block design test did not differ significantly between groups, although there was a tendency for superior scores among boys, t(18) = 1.56, p = .07. However, the two groups did differ significantly in verbal ability, as assessed by the BPVS, t(22) = 2.11, p = .05 (two-tailed). Given this result, a number of analyses of covariance were carried out to determine whether the gender groups differed in performance on any other measure when the potentially confounding effects of verbal ability were accounted for. These analyses confirmed that the two gender groups did not differ in performance on theory-of-mind measures, F(1,
Measures in Experiment 2

Table 3

Pearson Correlations for Interrelationships Between Measures in Experiment 2

<table>
<thead>
<tr>
<th>Measure</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Age</td>
<td>-.57</td>
<td></td>
<td>-.42</td>
<td></td>
<td>-.57</td>
</tr>
<tr>
<td>2. Verbal mental age</td>
<td></td>
<td>-.61</td>
<td>-.56</td>
<td>.55</td>
<td></td>
</tr>
<tr>
<td>3. Theory-of-mind score</td>
<td></td>
<td>-.12</td>
<td></td>
<td>.04</td>
<td></td>
</tr>
<tr>
<td>4. Embedded Figures Test time</td>
<td></td>
<td></td>
<td>-.99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Block design score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. df = 22, except for relationships with block design, where df = 18. All correlations greater than .42 are significant at the .05 level for a two-tailed test.

21) = 0.44, p = .26, the Embedded Figures Test, F(1, 21) = 0.04, p = .42, or the block design test, F(1, 17) = 0.50, p = .25.

Given the relatively smaller sample sizes used in Experiment 2 than in Experiment 1, correlational analyses were performed only across the whole sample and not for each gender group separately. Table 3 presents the intercorrelations among the measures of interest and shows that theory-of-mind performance was not reliably related to performance on either the Embedded Figures Test or the block design task. However, as predicted, verbal mental age correlated positively with theory-of-mind performance and block design score and negatively with Embedded Figures Test time. In addition, chronological age was positively correlated with theory-of-mind score, although it was not significantly related to either Embedded Figures Test time or block design score. Performance on the Embedded Figures Test was highly correlated with performance on the block design test.

These correlations indicate that with verbal development, and to some extent with increasing age, individuals get faster at the Embedded Figures Test and perform better on both block design and the theory-of-mind tasks. Consequently, the absence of the predicted inverse relationship between theory-of-mind performance and Embedded Figures Test speed may be due to a confounding effect of differences in verbal mental age and/or chronological age within the group. To control for verbal mental age and chronological age differences, we carried out a series of partial correlations. These indicated that theory-of-mind performance was related to Embedded Figures Test time when, first, verbal mental age was partialed out, r(21) = .71, p < .001, and second, both verbal mental age and chronological age were partialed out of the relationship, r(20) = .71, p < .001. Similarly, theory-of-mind performance was negatively related to block design score when verbal mental age was partialed out of this relationship, r(17) = -.57, p < .01, and also when both verbal mental age and chronological age were accounted for, r(16) = -.49, p = .01. A final partial correlation confirmed that performance on the Embedded Figures Test and performance on the block design test remained highly correlated when age and verbal mental age were accounted for, r(16) = -.98, p < .001 (two-tailed).

Discussion

The clear absence of any sex difference on the Embedded Figures Test in this experiment contrasts with the results of the previous study, where a trend for male superiority on the task was observed, and with the claims made by Baron-Cohen and Hammer (1997a). As discussed above, any sex effect found on the Embedded Figures Test has been shown to be relatively weak, and the available evidence suggests that any such effect is even less marked among young children. Voyer et al. (1995) suggested that sex differences in performance are not apparent before the age of around 14 years. Once again, the current finding is therefore consistent with the existing literature. Given the absence of any sex differences in performance on the Embedded Figures Test among this age group, it is perhaps not surprising that there is also no evidence of a sex difference in block design performance. Finally, the absence of any female superiority on the theory-of-mind tasks used here is also at odds with the suggestion that poor theory-of-mind ability is associated with a male brain type. Other work suggests that sex differences in theory-of-mind ability in children of this age are again, at best, weak effects (Charman, Ruffman, & Clements, 1999) and unlikely to be observed in samples of these sizes.

Of more interest is the question of whether poor theory-of-mind ability is associated with weak central coherence in this sample. This would be shown by a positive correlation between theory-of-mind score and Embedded Figures Test time (because faster Embedded Figures Test performance reflects weaker central coherence) and by a negative correlation between theory-of-mind ability and block design score. Neither of these two correlations approached significance here. However, as discussed, one might expect an inverse relationship between Embedded Figures Test time and theory of mind to be masked by the positive associations between both of these measures and level of development. In this sample, although individuals varied little in actual age, they did differ in intellectual level as indexed by receptive vocabulary ability. Partiaing out these developmental factors reveals the predicted relationships between theory-of-mind and central coherence measures.

This aspect of the results of Experiment 2 therefore replicates those of Experiment 1. A point made in the discussion of Experiment 1 that deserves repeating is that the evidence of inverse associations between theory-of-mind scores and performance on the central coherence tests amounts to relatively good evidence for a direct association between these two sets of tasks. Superior performance on both block design and embedded figures is linked to lower theory-of-mind scores, and vice versa. Although the analysis used here aimed to control for differences in intelligence, one might argue that partialing out age and verbal ability alone might not equate individuals for all aspects of intelligence. However, an inverse relationship cannot be explained in terms of some unaccounted-for third factor such as nonverbal intelligence or motivation, which might otherwise result in shared variance in performance.

This experiment also represents an improvement over Experiment 1. One advantage is that the current experiment included two separate measures of central coherence, both of which showed the predicted association with theory-of-mind score once developmental factors were accounted for. In addition, performance on these two measures was highly correlated, which suggests that they tap similar underlying processes. In fact, the correlation between these measures was so high, even when differences in age and verbal ability were accounted for, that it seems likely that performance on these tasks might be linked in a number of ways (see the General...
Discussion). A second advantage of the current approach is that the theory-of-mind tasks used were all essentially presented verbally. The link between theory-of-mind score and central coherence performance cannot therefore be explained away in terms of visual cognitive style (global or local) in this case. A further criticism of Experiment 1 was that the theory-of-mind task used there may not have actually tested theory-of-mind ability. An additional benefit of the tasks used in Experiment 2 is that they are not so open to that criticism. Two of the tests, the own and other’s false belief tasks, are paradigmatic theory-of-mind tests. The other tasks are either developmentally easier (inferred belief, not-own belief, explicit false belief) or more difficult (second-order false belief) than these standard tasks. Clearly, these developmental differences imply that the tasks must differ in the demands they make on individuals. A modular view of theory-of-mind development (e.g., Baron-Cohen, 1995) should therefore lead to concerns about exactly what this range of tasks is measuring, because according to this view, mind reading ability should not really develop once the theory-of-mind “module” has come on line (Fodor, 1983, 1992). In this view, any increase in the difficulty of a task must be due to general increases in the processing demands of that test, rather than to additional mind reading demands (cf. Leslie & Thaiss, 1992). However, not all theorists are committed to this view (e.g., Harris, 1992; Hobson, 1993; Perner, 1991; Wellman, 1990), and all of the tasks used here are well-accepted tests of theory-of-mind ability despite this theoretical concern. The important point to emphasize is that in order to succeed on any of these tasks, one needs to answer with reference to another’s (or one’s own) attitude to a certain content, in other words, what someone believes.

Experiment 2 therefore replicates, strengthens, and extends the results of Experiment 1 by indicating that poor theory of mind is related to weak central coherence in young children once developmental differences in performance are accounted for. This is a result that clearly has potential implications for our understanding of autism. These implications are considered fully in the General Discussion, but it is worth noting here that although relevant to autism, these findings were obtained from a typically developing population. Given this fact, we felt that a third experiment should be conducted in an attempt to directly replicate the findings of Experiment 2 among a sample of children with autism.

Experiment 3

Method

Seventeen individuals, all with a formal diagnosis of autism, took part in this study. They ranged in age from 7 years 4 months to 12 years 11 months, with a mean age of 9 years 11 months (SD = 21 months). There were 13 boys and 4 girls in the sample, a sex ratio that is consistent with accepted estimates of the sex ratio for autism (see Bailey, Phillips, & Rutter, 1996). Each individual was assessed on the series of theory-of-mind tests described in Experiment 2 (see the Appendix), with performance measured in terms of the number of tasks successfully completed. In addition, central coherence bias was indexed by block design performance and by the average time taken to complete the items of the children’s version of the Embedded Figures Test (Witkin et al., 1971). In this version of the Embedded Figures Test, two target shapes are used, a triangle for the first 11 test items and a house shape for the remaining 14 items. In accordance with the specified testing procedure, participants were given 1 min to find the target shape within each test item and were given a score of 1 min if they failed to find the target in that time. As in Experiment 2, the block design test was the Pattern Construction subtest of the Differential Ability Scales. However, in a departure from the procedure used in Experiment 2, this test was not scored on the basis of a combined measure of items completed and time taken to complete those items. An alternative, unspeeded scoring procedure is provided for the test that is recommended for use with children with learning disabilities or developmental delays, and this scoring procedure was preferred in this instance. The verbal mental age of each individual was assessed with the long form of the BPVS. Although this test measures only receptive vocabulary, there is reasonable evidence that it provides a representative measure of verbal ability among individuals with autism (Jarrold, Boucher, & Russell, 1997).

Results

Table 4 presents details of the sample’s verbal mental age, theory-of-mind level, Embedded Figures Test performance, and block design score. Given the small number of girls in this sample, the data are not split by gender, and sex effects are not examined in detail in this study. The mean verbal mental age of the group was significantly lower than its mean actual age, t(16) = 4.51, p < .001 (paired t test).

Table 5 presents the results of correlations examining links between the various measures assessed in this experiment. The table shows that theory-of-mind score did not relate significantly with individuals’ performance on either the Embedded Figures Test or the block design task. However, as expected, verbal mental age was positively associated with theory-of-mind performance (p = .04) and with block design score. Verbal mental age was also negatively related to Embedded Figures Test time (p = .04). In addition, chronological age was reliably correlated with block design score and was negatively associated with Embedded Figures Test time. One other important result to note is that performance on the Embedded Figures Test and performance on the block design task were highly correlated.

As in Experiment 2, these results reflect the fact that as individuals develop, they become faster at the Embedded Figures Test and perform better on block design and on theory-of-mind tasks. This fact in turn implies that an inverse relationship between theory-of-mind score and Embedded Figures Test speed might be masked by the codevelopment of these measures along with increasing verbal ability and, to a lesser extent, age in this sample. To account for any confounding effects of verbal mental age and for chronological age differences between individuals, we examined partial correlations. These showed that theory-of-mind score was positively related to Embedded Figures Test time when verbal mental age was accounted for, r(14) = .61, p < .01, and also when verbal mental age and age were partialed out of the relationship.

Table 4

<table>
<thead>
<tr>
<th>Measure</th>
<th>M</th>
<th>SD</th>
<th>Range</th>
</tr>
</thead>
<tbody>
<tr>
<td>Verbal mental age (in months)</td>
<td>94.53</td>
<td>31.85</td>
<td>57-175</td>
</tr>
<tr>
<td>Theory-of-mind score (no. of tasks passed)</td>
<td>3.12</td>
<td>1.87</td>
<td>1-6</td>
</tr>
<tr>
<td>Embedded Figures Test time (in seconds)</td>
<td>16.49</td>
<td>5.03</td>
<td>7.80-27.80</td>
</tr>
<tr>
<td>Block design (ability score)</td>
<td>137.53</td>
<td>25.21</td>
<td>104-191</td>
</tr>
</tbody>
</table>
Measures in Experiment 3

Pearson Correlations for Interrelationships Between Measures in Experiment 3

<table>
<thead>
<tr>
<th>Measure</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Verbal mental age</td>
<td></td>
<td>.72</td>
<td>.05</td>
<td>-.58</td>
<td>.70</td>
</tr>
<tr>
<td>3. Theory-of-mind score</td>
<td></td>
<td></td>
<td>.43</td>
<td>-.44</td>
<td>.58</td>
</tr>
<tr>
<td>4. Embedded Figures Test time</td>
<td></td>
<td></td>
<td></td>
<td>.31</td>
<td>-.00</td>
</tr>
<tr>
<td>5. Block design score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>-.65</td>
</tr>
</tbody>
</table>

Note. df = 15. All correlations greater than .48 are significant at the .05 level for a two-tailed test.

Discussion

In Experiment 3 we aimed to replicate the findings of Experiment 2 among a sample of individuals with autism. This replication was almost entirely successful, because the results presented here are similar to those observed in Experiment 2. Once again, the two relationships between theory-of-mind score and the central coherence measures were not significant when considered in isolation. However, when verbal mental age was accounted for, theory-of-mind score was positively related to Embedded Figures Test time and negatively related to block design score. Both of these results imply that those individuals with autism who display poorer theory-of-mind ability show weaker central coherence, and vice versa. In addition, the association between theory-of-mind score and Embedded Figures Test time remained significant when both verbal mental age and age were controlled for, r(13) = .53, p = .01. Theory-of-mind performance and block design score were also significantly related when verbal mental age was partialed out of this relationship, r(14) = -.35, p = .05. However, this relationship was not significant when both verbal mental age and age were controlled for, r(13) = -.18, p = .13. Finally, the relationship between Embedded Figures Test time and block design score approached significance when differences in age and verbal ability were accounted for, r(13) = -.42, p = .06.

General Discussion

Taken together, the three experiments presented here provide evidence for a link between performance on tests of theory-of-mind ability and performance on accepted tests of central coherence bias. Experiment 1 showed that performance on the Embedded Figures Test was inversely related to performance on Baron-Cohen, Jolliffe, et al.'s (1997) eyes task in an adult sample drawn from the general population. The correlation between these measures was not particularly strong but was reliable in the sample of 60 individuals. Using developmentally appropriate theory-of-mind tests, we replicated these results in Experiments 2 and 3 among typically developing children and among children with autism. Experiments 2 and 3 also had the advantage of including a block design test, another task thought to be influenced by central coherence bias. Although theory-of-mind ability was not directly related to performance on either the Embedded Figures Test or the block design test in these two experiments, significant relationships were observed between these measures when verbal mental age, or both verbal mental age and age, were accounted for. The fact that these relatively high correlations were observed when two different central coherence tasks were used and in two separate experiments with quite different participant groups suggests that they are unlikely to have arisen purely by chance.

It is therefore worth considering what this pattern of findings implies. A potentially helpful way of thinking about these results is to focus on the fact that partialing out verbal mental age allows one to examine the extent to which other variables share variance that is not related to verbal mental age. In other words, the partial correlations presented here assess the extent to which factors other than age or verbal mental age combine to determine performance on these tasks. One might therefore view these partial correlations as a test of the relationship between individual rather than developmental differences in theory-of-mind and central coherence measures. This is a distinction that is particularly relevant for the study of central coherence bias. As previously discussed, children become faster at the Embedded Figures Test as they develop. One would not want to say that this faster performance reflects individuals developing a weaker central coherence bias. Instead, the notion of weak or strong central coherence really applies only at the level of individual differences; it makes sense to talk about relative bias only within, and not across, developmental levels. Consequently, one possible reason why the relationship between central coherence measures and theory of mind becomes apparent only when differences in verbal mental age are accounted for is that this process removes the variance in central coherence task performance that is not accounted for by central coherence bias.

At the same time, one needs to remember that age and verbal mental age do account for a significant proportion of the overall variance in theory-of-mind scores and in performance on the Embedded Figures Test and the block design test in Experiments 2 and 3. Verbal mental age accounts for about 37% of the variance in theory-of-mind performance in Experiment 2 and 18% of the variance in theory-of-mind scores in Experiment 3. Similarly, verbal mental age explains between 19% and 34% of the variance in central coherence measures in both of these experiments. Consequently, it must be emphasized that the partial correlations presented in these experiments do not show that all of the variance in theory-of-mind ability is shared with central coherence measures. Instead, they indicate that individual differences, but not developmental differences, in both theory-of-mind and central coherence task performance are related to each other.

However, does this necessarily show that individuals with weak central coherence have poorer theory-of-mind ability? In discussing Experiment 1 we noted that this conclusion follows only if the tests used are valid measures of theory of mind and of central
coherence bias. In that discussion we raised concerns about the nature of the eyes task as a measure of metarepresentational theory of mind. We suggested instead that performance on the eyes task may be linked to Embedded Figures Test performance simply because the former requires a global visual processing style. This may explain why a significant first-order correlation was observed between this theory-of-mind measure and Embedded Figures Test performance in Experiment 1, in contrast to the results of Experiments 2 and 3. Alternatively, this difference in findings across experiments may reflect the use of developing populations in the latter two experiments. As noted, in these samples, large developmental differences may mask more subtle effects at the individual-difference level. This may be less of an issue in adults, because although adults will vary in age and verbal ability, they are likely to have developed to a level where differences in these measures are unlikely to form a major constraint on task performance. The results of Experiment 1 may therefore be consistent with the results of Experiments 2 and 3, but clearly more weight can be placed on the findings from the experiments that used accepted theory-of-mind tests.

An equally important question to ask is whether the tests used to assess central coherence bias really require a local visual processing approach. As previously discussed, both the Embedded Figures Test and the block design test are accepted measures of central coherence bias that have been used extensively in many other studies. Indeed, the Embedded Figures Test, which requires the participant to locate a target embedded in a more complex design, was explicitly designed to test the "ability to break up an organized visual field in order to keep a part of it separate from that field" (Witkin et al., 1971, p. 4). Of course, this in itself does not prove that both tests tap visual processing bias. However, there is also considerable empirical evidence to support this view. Much of it comes from the literature on field-dependent or field-independent cognitive styles (Witkin & Goodenough, 1981). Many studies in this area have shown that individuals who are relatively rapid at "disembedding" the target figure in the Embedded Figures Test also adopt a local approach to other tests of field dependence-independence such as the rod and frame test and the body adjustment test (see Witkin & Goodenough, 1981, and Witkin et al., 1971, for reviews). A local approach on the Embedded Figures Test or on close variants of this task has also been shown to be related to enhanced ability to "disembed" items in tasks presented in nonvisual modalities (e.g., Axcelrod & Cohen, 1961; Lefever & Ehri, 1976; White, 1954). Similarly, Shah and Frith's (1993) study of block design performance provides direct evidence to show that success on this task is mediated by an ability to take a local visual approach. The fact that presegmenting designs reduces the time taken to complete test items indicates that successful task performance requires the participant to ignore the global form and focus instead on the local constituent parts of the design (see also Mervis, Morris, Bertrand, & Robinson, 1999).

In addition, the relationships observed between performance on the Embedded Figures Test and performance on the block design test in Experiments 2 and 3 lend support to the claim that these tests tap the same underlying process. However, the relationship between these measures was weaker in Experiment 3 than in Experiment 2. Although the same block design test was used in both experiments, the scoring procedure for the test was altered for use with the children with autism in Experiment 3. In this case, performance was not scored in terms of the time taken to complete each trial. In Experiment 2, where time to complete block puzzles did play a part in determining ability scores, the relationship between performance on this task and Embedded Figures Test performance might have been inflated by a shared speed component. Those individuals who had a relatively rapid speed of processing might have performed at a higher level on both tasks, because both of these measures are time dependent. This might account, in part, for the extremely high correlation seen between these measures in Experiment 2. Nevertheless, the first-order correlation between performance on these two tests in Experiment 3 was still −.65, which dropped to −.42 when age and verbal ability were partialled out. Although the latter value only approaches significance given the small sample size in Experiment 3, it appears that these two tests do share reasonable variance even when performance is not dependent on the time taken to complete a test. Consequently, although there may be multiple factors influencing performance on both the Embedded Figures Test and the block design task, we would argue that they do provide a reasonable index of individuals' relative bias toward either a global or local visual approach. Having said this, we note that evidence of a weak drive for central coherence does not imply that an individual will never adopt a global approach. Clearly, the extent to which a local or global approach is preferred will depend on the interaction between an individual's bias and the requirements of the particular task (Jarrold & Russell, 1997; Plaisted, Swettenham, & Rees, 1999).

The current experiments, particularly Experiments 2 and 3, therefore do suggest that poorer theory-of-mind ability is associated with a weaker drive for central coherence, and vice versa. The provision of evidence for this link represents an important step forward in an understanding of autism, because it indicates that two important domains of psychological functioning associated with the disorder that have often been seen as distinct from one another may in fact be related. In addition, the data suggest that the same link between domains is found in the typically developing population, which clearly broadens the scope of this result. The fact that Experiments 2 and 3 used slightly different versions of the Embedded Figures Test and adopted different scoring procedures for the block design test as appropriate precludes a direct comparison of the nature of these relationships in autism and in typically developing children. Without further research, one cannot therefore be certain that the same processes are operating in these two groups. Although the individuals with autism may have relatively poorer theory-of-mind abilities and show relatively stronger embedded figures and block design performance than the younger children assessed in Experiment 2 (compare Tables 2 and 4), there seems no reason to suspect that the association between these two domains is qualitatively different in these two groups. This, though, begs the question of what exactly underlies this observed link between theory of mind and central coherence bias. In the remainder of this article we attempt to offer some speculative explanations for this apparent relationship.

One thing that is clear from the current data, and also from previous literature reviewed above, is that it is unlikely that this link represents a male brain profile, as Baron-Cohen and Harmer (1997a) suggested. No sex differences were found on the eyes task in Experiment 1 or in the theory-of-mind performance of the young children in Experiment 2. The existing literature suggests...
that there may in fact be weak sex effects on theory-of-mind performance. However, it seems unlikely that sex differences alone could explain the pattern of findings observed here. The expected sex difference on the Embedded Figures Test approached significance in Experiment 1. This is in line with previous literature which indicated that there is a trend toward male superiority on this task among the adult population. However, there was no suggestion of a sex difference in Embedded Figures Test performance in Experiment 2, which is again consistent with the literature on the use of this test with children. There is very little evidence, therefore, to suggest that the performance of children with autism on these tasks reflects an extreme form of the male pattern of performance.

Instead, it seems more likely that, across both sexes, weak central coherence must somehow be linked to relatively poor theory-of-mind performance. Although the correlational approach used here provided no evidence of causal priority, other data suggest that these two domains are most likely to be linked in this direction (as opposed to poor theory-of-mind causing weak central coherence), the reason being that central coherence appears to be a lower level process than theory of mind. The effects of central coherence bias have been seen in what are thought to be relatively high-level tasks such as reading for meaning (Frith & Snowling, 1983; Happé, 1997) or decoding communication (see Frith, 1989b). However, central coherence phenomena have also been observed in very low-level tasks, such as counting visual stimuli and locating embedded figures (Jarrold & Russell, 1997; Jolliffe & Baron-Cohen, 1997; Shah & Frith, 1983), and possibly even in susceptibility to visual illusions (Happé, 1996). It appears highly unlikely that something as "psychological" as theory of mind, however implicitly it might function, could explain low-level visual bias effects such as these.

Of course, this raises the question of how strong central coherence—in other words, a global processing style—is linked to theory-of-mind ability. At the outset of the article we noted that Frith and Happé (1994) had reached the conclusion that theory of mind and central coherence were distinctive domains of psychological functioning in autism. In fact, this is a relatively recent theoretical position for Frith, and in her 1989 book she argued that a theory-of-mind deficit would arise from a failure to integrate relevant information from a variety of sources. Her argument there is that a weak drive for a "meaningful" integration of information is the "central cognitive dysfunction" associated with autism. She suggests that a theory of mind "can be seen as a cohesive interpretative device par excellence: it forces together complex information from totally disparate sources" (Frith, 1989b, p. 174).

This suggestion may be consistent with other accounts of the early development of theory-of-mind ability that have emphasized the need to integrate visual information from external sources. Many would argue that joint attention—the ability to attend to what someone else is attending to—is a crucial step on the way to acquiring a full theory of mind, partly because of the evidence of joint-attention deficits in autism (see Charman, 1997, for a recent review). For example, Hobson (1990, 1993) suggested that the ability to "triaangular" attention between self, other, and the object of the other's attention is an important developmental precursor to mind reading proper. Baron-Cohen (1991, 1995) also argued for the importance of this kind of process, though from a very different, modular, theoretical perspective. Mundy and Sigman (1989) claimed that joint-attention deficits may be more fundamental than metarepresentational problems in autism, but they suggested that both deficits may reflect "difficulty with abstracting contingencies in social interaction" (p. 180).

One might therefore suggest, in line with Frith (1989b), that strong central coherence is important in theory-of-mind development because it biases the developing individual to take a global view of a situation and to integrate what the individual and what another person are attending to. An individual with weak central coherence might fail to integrate separate cues into a meaningful model of the global social situation. Clearly, this is a rather vague and broad hypothesis, which arguably does little more than rephrase existing developmental accounts. However, it does at least suggest a pathway by which individual differences in central coherence bias might conceivably give rise to individual differences in theory of mind. This hypothesis also casts some doubt on the modularity, or at the very least the domain specificity, of theory of mind because it suggests that aspects of theory-of-mind ability may rest on nonsocial perceptual biases. Finally, this account may lead to some testable predictions. If correct, then one would expect measures of joint-attention behavior in very young children to relate to central coherence bias if one could measure it at the appropriate age level. One might also predict that individuals with a local processing bias should be relatively poor at integrating social cues from either visual or auditory stimuli.

The results of the experiments described in this article suggest that individual differences in central coherence bias and theory-of-mind development are linked, but further work is clearly needed in order to properly explain the nature of this association. In addition to examining the impact of central coherence bias on potential precursors of theory-of-mind functioning, future researchers might also explore the component processes of central coherence task performance (cf. Plaisted, O'Riordan, & Baron-Cohen, 1998a, 1998b). Other questions that remain to be addressed are whether the relationships observed here among individuals with autism and among typically developing children are qualitatively similar and whether these links really do persist into adulthood, as might seem to be suggested by the results of Experiment 1.

References


THEORY OF MIND AND CENTRAL COHERENCE BIAS


Appendix

Protocols for Theory-of-Mind Tasks Used in Experiments 2 and 3

**Inferred Belief**

"This is Emma [show picture]. This morning Emma saw her pencil case on the desk [show picture of desk], not on the shelf [show picture of shelf]. Now Emma wants her pencil case. Where will she look for it?" [correct answer: desk]

**Not-Own Belief**

"This is Paul [show picture]. Paul wants to find his dog. It might be hiding in the house [show picture of house] or in the garden [show picture of garden]. Where do you think Paul’s dog is hiding?" [child answers]. "That’s a good guess. Paul thinks his dog is in the [other location]. Where will Paul look for his dog?" [correct answer: other location]

**Explicit False Belief**

"This is Andrew [show picture]. Andrew wants to find his kitten. Andrew thinks his kitten is in the kitchen [show picture of kitchen], but really it is in the bedroom [show picture of bedroom]. Where will Andrew look for his kitten?" [correct answer: kitchen]

**Own False Belief**

"Here is a box [show chocolate box]. What do you think is in here?" [child answers: chocolates] "No, look, it’s got Lego bricks in it [show Lego bricks and then close lid]. What is really in here?" [child answers: Lego] “When I first showed you the box, what did you think was inside it?” [correct answer: chocolates]

**Other’s False Belief**

"Here is Rachel [show doll]. She has got a sweet, look. She is going to put her sweet under the blue box. Now Rachel is going out; she will come back for her sweet later [act out with doll]. Now here comes Sarah [show second doll]. Sarah is naughty and is going to move the sweet. Look, she’s going to put it under the red box. Now Sarah goes away [act out with doll]. Ah, here comes Rachel, she wants to get her sweet now. Where does Rachel think the sweet is?" [correct answer: blue box]

**Second-Order False Belief**

"Here are Sarah and Michael in the park [show dolls and act out following scenario]. Along comes the ice-cream man. Michael would like an ice-cream, but he has left his money at home. He is very sad. “Don’t worry,” says the ice-cream man, “you can go home to get your money and buy some ice-cream later. I’ll be here in the park all afternoon.” So Michael goes home. Now the ice-cream man says to Sarah, “I am going to drive my van to the church to see if I can sell my ice-cream there.” The ice-cream man drives to the church, but on his way he passes Michael’s house. Michael sees him and asks him where he is going. The ice-cream man says, “I am going to sell my ice-cream outside the church” and then drives off. Now Sarah goes home to her house and then goes over to Michael’s house. She knocks on the door and asks Michael’s mum if Michael is in. “No,” says Michael’s mum, “he’s gone to buy an ice-cream.” Where does Sarah think that Michael has gone?" [correct answer: the park]