Are children with autism blind to the mentalistic significance of the eyes?

Simon Baron-Cohen*
Departments of Experimental Psychology and Psychiatry, University of Cambridge, Downing Street, Cambridge CB2 3EB, UK

Ruth Campbell
Department of Psychology, Goldsmiths' College, University of London

Annette Karmiloff-Smith and Julia Grant
MRC Cognitive Development Unit, London

Jane Walker
Department of Psychology, Goldsmiths' College, University of London

Previous work shows that children with autism manifest abnormalities in the use of gaze. They also have difficulties in the comprehension of mental states. The present paper explores whether these two abnormalities might be related. We report four experiments that test whether normal children ‘read’ the eyes, particularly eye-direction, as conveying information about a person’s mental states, and whether subjects with autism are specifically ‘blind’ to such information. The mental states assessed were desire, goal, refer, and think.

The results confirm that normal children do use eye-direction as a cue for reading these mental states, as do subjects with mental handicap (including those with William’s Syndrome). In contrast, subjects with autism failed to use eye-direction to infer the mental states. In addition, whilst normal children and children with mental handicap showed a preference for eye-direction over an unnatural cue when inferring these mental states, children with autism did not. These findings suggest that part of the explanation for the gaze abnormalities in autism may be a failure to comprehend that the eyes convey information about a person’s mental states.

That eye-contact by children with autism is abnormal has been noted ever since Kanner’s (1943) first descriptions of the condition: ‘When he is with other people, he does not look up at them’ (reprinted in Kanner, 1973, p. 26). This gaze abnormality was for a time thought to be simply avoidance (Bettelheim, 1968; Richer, 1978; Tinbergen & Tinbergen, 1983), but more recent studies rule this out. Children with autism seem to

* Requests for reprints.
make the same quantity of eye-contact overall (Sigman, Mundy, Ungerer & Sherman, 1986; Volkmair & Mayes, 1990), but not at the same times as children without autism. Rather, there is some consensus that it is the use of eye-contact by these children which is different (Baron-Cohen, 1988; Frith, 1989; Volkmair & Mayes, 1990; Tantam, 1992). For example, they do not seem to use it to regulate turn taking (Mirenda, Donnellan & Yoder, 1983) or for joint attention (Baron-Cohen, 1989a; Sigman et al., 1986) or for goal-detection (Phillips, Baron-Cohen & Rutter, 1992).

What might be causing this abnormality in the use of eye-contact? And how might this relate to the problems they also show in social development (American Psychiatric Association, 1987; Baron-Cohen, 1988; Rutter, 1983)? One hypothesis is that the eye-contact abnormalities reflect broader abnormalities in face processing (de Gelder, Vroomen & van der Heide, 1991). A number of studies have tested this possibility. Overall, however, rather than discovering face-processing impairments, these studies have mostly demonstrated intact face-processing abilities in autism. For example, children with autism can recognize identity and gender from photographs of the face alone (Baron-Cohen, 1991a; Campbell, Walker & Baron-Cohen, 1993; Hobson & Lee, 1988a; Langdell, 1978) and can recognize ‘simple’ emotional expressions in the face (Baron-Cohen, Spitz & Cross, 1993; Hobson, Ouston & Lee, 1988b). They are also able to distinguish speaking from non-speaking faces and recognize which of several phonemes a photographed face is likely to have been saying, from the shape of the mouth alone (Walker, 1992). However, they do seem to make less use of the eye region in making facial identity judgements (Hobson et al., 1988a; Langdell, 1978). This latter anomaly simply confirms that they are not making use of the eyes in the same way as do normal children. In sum, the notion that eye-contact problems in autism are caused by generalized abnormalities in face processing seems to be ruled out.

An alternative hypothesis, and one that we test in this paper, is that children with autism fail to understand something crucial about the eyes themselves (Baron-Cohen, 1993, 1994, 1995a,b,c; Hermelin & O’Connor, 1985; Hobson, 1990; Hutt & Ounsted, 1966). In this paper we report a series of experiments which test if children with autism understand the eyes in the same way as children without autism do.

Knowledge about the eyes

Perhaps the most basic piece of knowledge about the eyes is that eyes are for seeing. Young normal children, from at least 30 months of age, seem to understand what we might call the ‘laws of vision’: that for a person to see something, their eyes need to be open and their line of regard needs to be unobstructed (Lempers, Flavell & Flavell, 1977). There are several studies showing that children with autism also understand this basic fact (Baron-Cohen, 1989a,b, 1991b; Hobson, 1984; Leslie & Frith, 1988; Tan & Harris, 1991). In our first experiment we tested if children with autism could infer when someone could see them, since this is a crucial and salient aspect of eyes: so called ‘dyadic’

1 Some impairments in face processing have been found. For example, de Gelder et al. (1991) found an immunity to the McGurk effect in autism. It is unclear if this reflects a face-processing deficit specifically or a failure in cross-model processing.
eye-contact (Bakeman & Adamson, 1984; Gomez, 1991). We predicted that children with autism would have no difficulty on this task.

A second, related piece of knowledge about the eyes that young normal children possess is that people use their eyes to communicate their interest in something (Argyle, 1972). This kind of eye-contact is often called ‘triadic’ (Bakeman & Adamson, 1984) in that it involves two individuals sharing a focus of attention on a third object. It is clear that by 14 months of age, toddlers engage in gaze monitoring (Butterworth, 1991; Butterworth & Jarrett, 1991; Scaife & Bruner, 1975) and that by 18 months old toddlers recognize that the object an adult is looking at, while naming, is the object the adult is referring to (Baldwin, 1991). An even clearer case of the role of the eyes in communication is when no spoken language is used. Thus, imagine trying to get someone to look at something without moving your head or your hands or producing any vocalization. In all likelihood, you will first try to ‘catch’ their eye by making eye-contact; then you will look across at the object of interest, and then look back at them. In such examples of non-verbal communication, the observer treats the person’s eye-direction as one way in which a person can point, or refer, to things, or request something. Eye-movement is interpreted, rightly, as an ostensive act and thus the observer searches for plausible intentions behind the act (Argyle, 1972; Austin, 1962; Baron-Cohen, 1988; Grice, 1975; Leslie & Happé, 1989; Rutter, 1984; Sperber & Wilson, 1986).

Given that intentions, goals and desires are key mental states that children with autism fail to comprehend (Phillips, 1993; Phillips, Baron-Cohen, & Rutter, 1995), we predicted that they would also have difficulty in understanding that the eyes communicate information about a person’s intentions, goals and desires. This was tested in Expts 2 and 3. This prediction receives further support from clinical anecdotes from researchers who describe able adults with autism who complain of being unable to understand how other people talk to each other with their eyes or what they are silently ‘saying’ to each other (Hobson, cited in Frith, 1989).

Finally, a related piece of knowledge about the eyes that normal 3–4-year-olds possess is that, if a person is looking upwards and away, this is seen as a sign that that person is thinking (Baron-Cohen & Cross, 1992; Flavell, Green & Flavell, 1995). Presumably, the eyes are interpreted in this way because, in the absence of an external object out there being looked at, the observer assumes that the person is attending to an internal (mental) object. Given that children with autism have difficulty in understanding mental state concepts such as thinking (Baron-Cohen, 1989c, 1990; Baron-Cohen, Leslie & Frith, 1985; Perner, Frith, Leslie & Leekam, 1989), we predicted that they would also have difficulty in understanding that the eyes convey information about this mental state. This was tested in Expt 4.

Subjects

We tested three groups of subjects who all took part in all four experiments. The first group (N = 20) all had a diagnosis of autism, using established criteria (DSM-III-R, APA, 1987; Rutter, 1978), and were attending special schools for autism. The second

---

2 This thought-experiment comes from Baron-Cohen & Cross (1992).
group (N = 30) all had mental handicap and were attending special schools for pupils with pervasive learning disabilities. This latter group was chosen in order to disentangle what might be due to the effects of mental handicap, separate from autism, using a research design derived from earlier work (Baron-Cohen et al., 1985; Frith, 1989; Hermelin & O’Connor, 1970). While 20 of these subjects had mental handicap of unknown aetiology, 10 of them had a clear-cut diagnosis of William’s Syndrome (DSM-III-R, APA, 1987). This subgroup was included on grounds of theoretical interest; they may in some respects be the inverse of children with autism, in terms of the profile of psychological skills that they show (Karmiloff-Smith, 1992a,b). The third group (N = 20) were normal 4-year-olds, all attending a nursery school.

All subjects in the two clinical groups had a verbal mental age (MA) above 4-years-old, as assessed using the Test of Receptive Grammar (Bishop, 1983). This tests comprehension of syntax and is held to be a more conservative test of language comprehension than the more commonly used British Picture Vocabulary Scale (BPVS), while employing a similar choice format. They were also tested for non-verbal MA, using Raven’s Coloured Progressive Matrices (Raven, 1956). One subject with autism refused the Raven’s testing and one subject with mental handicap refused both Raven’s and TROG testing. Whereas the group with mental handicap did not differ significantly on verbal and non-verbal MA, the group with autism had a non-verbal MA significantly higher than their verbal MA, consistent with previous reports of their uneven profile (Frith, 1989). The subject details are shown in Table 1. The two clinical groups did not differ significantly from each other in terms of either chronological age (CA) (t(48) = 1.72, p = .09) or verbal MA (t(47) = 3.36, p = .07). Regarding non-verbal MA, the group with autism were functioning at a significantly higher level than the group with mental handicap (t(46) = 18.5, p = .0001).

Table 1. Mean CA and MA of the subjects in Expt 1

<table>
<thead>
<tr>
<th></th>
<th>CA</th>
<th>Verbal MA&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Non-verbal MA&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Normal</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>M</em></td>
<td>4:3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>0:2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>4:0–4:8</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Mental handicap</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>M</em></td>
<td>11:1</td>
<td>5:3</td>
<td>5:4</td>
</tr>
<tr>
<td>SD</td>
<td>3:0</td>
<td>0:9</td>
<td>1:2</td>
</tr>
<tr>
<td>Range</td>
<td>5:0–21:5</td>
<td>4:0–7:0</td>
<td>3:2–8:5</td>
</tr>
<tr>
<td><strong>Autism</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>M</em></td>
<td>13:5</td>
<td>4:9</td>
<td>7:2</td>
</tr>
<tr>
<td>SD</td>
<td>3:3</td>
<td>0:5</td>
<td>1:7</td>
</tr>
<tr>
<td>Range</td>
<td>8:0–18:2</td>
<td>4:0–5:8</td>
<td>4:5–10:0</td>
</tr>
</tbody>
</table>

<sup>a</sup> Test of Reception of Grammar (TROG).

<sup>b</sup> Raven’s Coloured Matrices.
EXPERIMENT 1

The ‘Which one is looking at you?’ task

In these experiments we used cartoon faces as stimuli, in order to impose maximum experimental control over the information being presented and processed. The order of presentation of Expts 1–4 was randomized, though for clarity of reporting they are described in a particular order here. In this first experiment, we made eye-direction the only available cue. Faces were either looking at the subject or looking away. We predicted that, since this only involved computing ‘dyadic’ eye-contact, children with autism would be unimpaired on this task.

Method and materials

We had six trials, randomly presented. In each trial, the subject was presented with a pair of faces and simply asked ‘Which one is looking at you?’. Examples of the stimuli used in this test are shown in Fig. 1.

![Figure 1. An example of a pair of cartoons used in Expt 1.](image)

In addition, in order to check that subjects were attending to the eyes, we included a ‘silly’ control condition in which we presented six pairs of cartoon faces, one of each pair displaying the eyes in impossible positions (i.e. moving independently of each other), and asked ‘Which is the silly face?’. We also employed an emotion control condition, to check that subjects were able to make judgements about a face when features other than the eyes were important. In this condition, we simply varied whether the mouth was smiling or frowning and asked of six such pairs, ‘Which is the happy face?’ or ‘Which is the sad face?’. Finally, since the experimental task involved the subject being able to trace imaginary straight lines, in order to check that any failure on this task was not due to this geometric component we included an arrow condition, in which the subject was shown a black arrow pointing to one of a random array of six dots and was asked, ‘Which dot is the arrow pointing at?’. There were 10 trials in the arrow condition. An example of each of these three control conditions is shown in Fig. 2.
Figure 2. An example of a pair of cartoons used in the three control conditions of Expt 1: the silly condition (a), the emotion condition (b), and the arrow condition (c).
**Results**

The mean scores of each group on each condition are shown in Table 2. There were no group differences on any of the conditions: silly control condition \((F(2,67)) = .44, p = .65\); emotion control condition \((F(2,67) = .53, p = .59)\); arrow control condition \((F(2,67) = 2.53, p = .09)\); look condition \((F(2,67) = .46, p = .63)\). An analysis in terms of the percentage of each group passing each condition overall, where a pass was defined as scoring equal to or more than five trials, revealed the same picture (see summary in Table 6). Note that since the probability of success on one trial by chance alone is .5, the probability of passing five or more trials by chance would be very small (.031).

**Table 2. Mean scores on each condition in Expt 1 by each group (maximum = 6)**

<table>
<thead>
<tr>
<th></th>
<th>Look</th>
<th>Silly</th>
<th>Emotion</th>
<th>Arrow (max. = 10)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Normal</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(M)</td>
<td>5.3</td>
<td>5.1</td>
<td>6.0</td>
<td>6.9</td>
</tr>
<tr>
<td>SD</td>
<td>0.9</td>
<td>1.1</td>
<td>0</td>
<td>2.3</td>
</tr>
<tr>
<td><strong>Mental handicap</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(M)</td>
<td>5.2</td>
<td>5.1</td>
<td>5.9</td>
<td>8.4</td>
</tr>
<tr>
<td>SD</td>
<td>1.0</td>
<td>1.5</td>
<td>0.7</td>
<td>5.2</td>
</tr>
<tr>
<td><strong>Autism</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(M)</td>
<td>5.0</td>
<td>5.4</td>
<td>5.9</td>
<td>5.8</td>
</tr>
<tr>
<td>SD</td>
<td>1.2</td>
<td>1.1</td>
<td>0.4</td>
<td>3.8</td>
</tr>
</tbody>
</table>

**Discussion**

Using cartoon faces, children with autism were not impaired in judging when someone is looking at them. Their good performance on the three control conditions also demonstrates that children with autism do attend to the eyes and the face and can trace imaginary straight lines. Their ceiling performance on the emotion control condition also replicates an earlier study (Baron-Cohen et al., 1993). Having demonstrated that, at least under the highly structured and simplified conditions of cartoon stimuli, they can judge where someone is looking, in the next experiment we tested if they could infer mental states from eye-direction, using the same materials.

**EXPERIMENT 2**

**The four sweets task**

In this experiment we explored whether for normal children the eyes communicate information about the simple mental states of desire, goal and intention to refer. In addition, we tested our prediction that subjects with autism would be impaired in recognizing these, while subjects with mental handicap would not be.
Materials and method

We used a display in which four sweets were colourfully depicted, using the normal wrappers from these sweets, each pasted on to its own quadrant of a sheet of white, A4 size card. The four sweets were Smarties, Mars Bar, Milky Way and Polo Mints, selected as highly recognizable to most children. The experimenter said to the subject, 'Here are four sweets. Do you know what they are?' All of our subjects named them correctly, with the exception of four of the normal 3-year-olds, who did not recognize Polo Mints. The name of this sweet was then supplied, and their ability to name this was retested. When it was clear that the subject knew the names of all four sweets, the experimenter asked, 'Which one is your favourite?', and noted which one the subject pointed to or named, in order to establish if later responding was simply due to a bias to point to the subject's favourite sweet. A second display was used in half of the trials, in which each sweet now had a new position, in order to establish if later responding was simply due to a bias to point to the same position.

In the desire condition, the experimenter first asked, 'Which one do you want?', and noted the subject's choice. Usually this coincided with the one they had identified as their favourite. The experimenter then laid a plastic transparency over the display, on which was drawn a cartoon face identical to that used in Expt 1, such that the cartoon face fitted centrally between the four sweets. The eyes of the cartoon face pointed to one of the four sweets. Just as he or she laid it over the display, the experimenter said, 'This is my friend Charlie. Charlie wants one of these sweets. Which one does Charlie want?'.

The experimenter was always careful to select a transparency in which Charlie's eyes were pointing to one of the three sweets not identified as the one the subject wanted (see Fig. 3). If the subject pointed to the sweet that Charlie's eyes were pointing towards, this was scored as a pass (P). If the subject pointed to the sweet they had previously identified as the one that they wanted, this was scored as an egocentric (E) response. Finally, if the subject pointed to one of the other sweets (neither their favourite, nor the one Charlie was looking at), this was scored as a random (R) response.

Figure 3. The display used in Expt 2, with the transparency overlaid onto it.
Gaze comprehension in autism

In the goal condition, before the transparency of Charlie's face was placed over the display, the experimenter asked about the subject's own goal: 'Which one will you take?'. Then, having overlaid the transparency with a face looking (randomly) at one of the three other sweets, the experimenter asked, 'Here is Charlie again. He is going to take a sweet. Which one is he going to take?'. Again, a P, E or R response was scored, using the same criteria outlined previously.

In the refer condition, the experimenter introduced a new display, on which there were two nonsense shapes, one in each of the two lower quadrants of a sheet of A4 white card. The experimenter said, 'One of these is a beb [an arbitrarily selected nonsense word, designed to be a novel word for the subject], Which one is the beb?'. Interestingly, all the normal children and the subjects with autism pointed to one of the shapes when asked this question, as did most of the subjects with mental handicap, though some of this latter group also replied, 'Don't know'. In this case, the experimenter asked them to guess which was the beb. Then the experimenter laid the transparency over the display, such that Charlie's eyes were looking at the one the child had not chosen. The experimenter then said 'Charlie says, "There's the beb!" Which one does Charlie say is the beb?'. A pass was scored if the subject pointed to the one Charlie was looking at.

In addition to the three mental state conditions (desire, goal and refer), there was a size control condition, in order to check if subjects passing these experimental tasks were simply using a strategy of pointing to wherever Charlie's eyes were pointing, irrespective of the relevance of this strategy. This control condition comprised showing two different size boxes drawn on a page, with Charlie looking at the bigger of the two. The experimenter said, 'Charlie is going to put one box inside the other. Which box can go inside the other?'. Thus, the question was again about Charlie, but a correct answer required the subject to ignore where Charlie's eyes were pointing. Over different trials in this condition, the smaller box was randomly positioned above or beneath the larger box.

There were three trials of each condition, with the exception of refer and size, for which there were six trials. This was because, for all trials except refer and size, the probability of passing one trial by chance alone was .25, and thus the probability of passing three out of three trials by chance alone was very small (.016). On the refer and size conditions, the probability of passing one trial was .5 and thus a larger number of trials (six) was introduced, in order to minimize chance performance. The probability of passing five out of six trials by chance alone is .031. The other five nonsense words used in the refer condition were zab, reth, bor, tbo and sem. Trials in all conditions were mixed and presented in random order. Finally, in the refer and size conditions, just pass or fail was scored, as there were only two response options.

Results

The mean scores of each group on each of the conditions are presented in Table 3. ANOVAs revealed significant group effects on all three mental state conditions: desire (F(2,67) = 7.87, p < .009), goal (F(2,67) = 27.1, p < .001) and refer (F(2,67) = 31.61, p < .001). Post hoc Scheffé tests showed all of these to be due to the group with autism scoring significantly lower than the other two groups (all p < .05). The groups did not differ on the size control condition (F(2,67) = 16.0, p = .57). The summary in Table 6 presents the percentage of each group passing each condition, where a pass is defined as scoring three out of three on the desire and goal tasks, and equal to or more than five out of six on the refer and size conditions. Fisher Exact Probability Tests on these percentages mirrored the results from the analysis of raw scores: significantly fewer subjects in the group with autism passed overall on desire, goal and refer (all p < .05). The group with autism was significantly lower than chance on these tests, reflecting that their errors were not random but were largely in the egocentric category. Indeed, analysis of error types revealed that the vast majority of errors by subjects in each group were E errors (86.2 per cent in the group with autism, 100 per cent in the normal group and 80 per cent in the mental handicap group). These E errors were not accounted for by a positional bias in responding. The remaining errors comprised R errors.
Table 3. Mean scores on each condition in Expt 2 by each group (maximum = 3, for desire and goal all conditions; for refer and size conditions where maximum = 6)

<table>
<thead>
<tr>
<th></th>
<th>Normal</th>
<th></th>
<th>Mental handicap</th>
<th></th>
<th>Autism</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Desire</td>
<td>2.8</td>
<td>0.6</td>
<td>2.7</td>
<td>0.8</td>
<td>1.8*</td>
<td>1.2</td>
</tr>
<tr>
<td>Goal</td>
<td>2.7</td>
<td>0.6</td>
<td>2.8</td>
<td>0.5</td>
<td>1.3*</td>
<td>1.2</td>
</tr>
<tr>
<td>Refer</td>
<td>4.9</td>
<td>1.3</td>
<td>5.0</td>
<td>1.6</td>
<td>1.7*</td>
<td>1.7</td>
</tr>
<tr>
<td>Size</td>
<td>5.5</td>
<td>1.3</td>
<td>5.0</td>
<td>1.6</td>
<td>5.9</td>
<td>0.5</td>
</tr>
</tbody>
</table>

* Autism × other two groups, p < .001.

Discussion

Experiment 2 extends the results of the previous experiment in two ways. First, it demonstrated that both normal children and subjects with mental handicap use the direction of a person's eyes as a cue in inferring the mental states of desire, goal and refer. The role of eye-direction in inferring reference also replicates Baldwin's (1991) finding from normal children, though with an older age group and a new procedure. Second, Expt 2 suggests that the majority of subjects with autism fail to read eye-direction to infer the mental states of desire, goal and refer. Instead, they mostly responded in an egocentric fashion, pointing to their own favourite sweet. This deficit occurred despite being as good on the control (size) condition. In the next experiment we retested this effect and explored it further, by manipulating the available cues in the display and by including a further control condition to check if failure on the task might be due to the linguistic complexity of the test questions.

EXPERIMENT 3

Are the eyes the natural pointer?

In this experiment we were interested to test whether normal and mentally handicapped subjects saw the eyes as a natural cue in such tasks, and whether children with autism did not. The experiment involved presenting the subject with a choice of a natural cue (eye-direction) and an unnatural cue (arrow-direction) with which to infer a character's mental state. In addition, we included a colour control task to investigate the role of syntactic complexity in the failure of children with autism on the 'eyes' tasks.

Method and materials

We tested desire and goal only (in order to avoid overtesting these subjects) in exactly the same way as had been done in Expt 2, except that on the transparency was also drawn a bold, black arrow pointing to one of the four sweets. If, in response to the test question, the subject pointed to the sweet to which the arrow was pointing, this was scored as an arrow (A) response. Otherwise, as before, pass (P), egocentric (E) and random (R) responses were coded.

We also included a colour control condition in order to check several variables that might be affecting performance: first, we wanted to check if the subject was able to ignore both eye-direction and arrow-direction when required; and second, we wanted to check if the subject was able to understand verbal instructions of equivalent sentence length and complexity when the question did not relate to Charlie's eyes or mental states. It comprised showing Charlie's face located between four boxes, each of which was
coloured differently. Charlie’s eyes were depicted as looking at one of these boxes and an arrow was depicted pointing at another. The subject was first asked to name the four different colours and was then told, ‘Charlie is going to do a painting. These are his paints. Which one will Charlie use to paint [some grass]?’ With each trial in this condition, the correct answer involved pointing to one of the boxes which was neither indicated by Charlie’s eyes nor by the arrow. Again P, A and R responses were recorded but, instead of E responses, a new category of gaze (G) responses was scored if the subject selected the paint colour that Charlie was looking at. E responses could not be scored, since the experimenter did not ask which was the subject’s favourite colour. In the other two trials in this condition, the pictures named were the sun and the sky.

As in Expt 2, there were three trials in each of the desire, goal and colour conditions. These trials and conditions were presented in randomized order. A pass was again defined as passing three out of three trials on a condition. The probability of passing by chance alone was \( p = .016 \). Finally, three arrow control questions were asked at the end of the experiment, in order to confirm the earlier finding (from Expt 1) that all subjects understood the convention of arrows. This took the form of asking ‘Where is the arrow?’ and ‘Which sweet is the arrow pointing to?’. All subjects were able to answer these questions.

### Results

Table 4 shows the mean score of each group on each of the four conditions. As in Expt 2, ANOVAAs and post hoc tests again showed that the group with autism performed significantly worse than the other two groups on desire \( (F(2,62) = 11.99, p<.001) \), goal \( (F(2,62)) = 6.45, p<.002) \), but not on colour \( (F(2,67) = .13, p = .87) \). Note that for the desire and goal tests in this experiment, the group with mental handicap was reduced to \( N = 25 \) (rather than 30), as five subjects who were at ceiling on all of the previous experiments opted not to do these two tests. The summary in Table 6 also shows this group difference when comparing the percentage of each group passing each condition (autism \( \times \) other two groups, Fisher’s Exact Tests, all \( p<.005 \), except colour, which was \( p>.05 \)).

<table>
<thead>
<tr>
<th></th>
<th>Normal</th>
<th>Mental handicap</th>
<th>Autism</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>( M )</td>
<td>( SD )</td>
<td>( M )</td>
</tr>
<tr>
<td>Desire</td>
<td>2.5</td>
<td>0.8</td>
<td>2.7</td>
</tr>
<tr>
<td>Goal</td>
<td>2.7</td>
<td>0.7</td>
<td>2.5</td>
</tr>
<tr>
<td>Colour</td>
<td>2.9</td>
<td>0.5</td>
<td>2.8</td>
</tr>
</tbody>
</table>

* Autism \( \times \) other two groups, \( p<.004 \).

An error analysis showed that while the group with mental handicap continued to make predominantly E errors (50 per cent) and some A errors (35.7 per cent) and R errors (14.4 per cent), the group with autism now made predominantly A errors (42.1 per cent), with some E errors (29.8 per cent) and R errors (21.1 per cent). The normal group’s errors resembled those of the children with mental handicap (A errors = 16.6 per cent, R errors = 25 per cent, E errors = 58.3 per cent). The difference in the proportion of arrow (A) responses by the group with autism was highly significant (autism \( \times \) mental handicap, \( \chi^2(1) = 44.19 \); autism \( \times \) normal, \( \chi^2(1) = 46.94 \), both \( p<.001 \)). On the colour control condition, just one subject in the group with mental handicap made any errors, these being all A errors, and just two subjects in the normal group made errors, these
being all G errors. In the group with autism, two subjects made one error each, these being one R and one A error.

**Discussion**

Experiment 3 replicates the results from Expt 2 in finding autism-specific deficits in using eye-direction to infer a character's desire or goal, but extends that result in showing that, when normal subjects or those with mental handicap approach this task, they show a spontaneous preference for the natural cue (eye-direction) over an unnatural cue (arrow-direction). In contrast, the subjects with autism tend to use the unnatural cue of the arrow-direction to guide their judgement, acting as if this is *more* important than eye-direction. Interestingly, in this study, both the normal children and the subjects with mental handicap virtually ignored the arrow in making their judgement: hardly any of them ever even asked the experimenter what the arrow was for.\(^3\) They simply acted as if the arrow was not there, or was at best irrelevant to the task. This cannot have been because they did not understand the graphic convention of arrows, since they all passed the arrow control questions.

We conclude that the performance by children with autism on this test reveals their 'blindness' to the mentalistic significance of the eyes. Their use of a salient but unnatural cue in this task suggests that, for them, the task must have been more like a test of intelligent guess work, whereas for the normal children and those with mental handicap it was a test of 'natural psychology' (Humphrey, 1983). In the final experiment we sought to test if children with autism would have similar difficulty in inferring one last mental state from information in the eyes, namely, when someone is *thinking*.

**EXPERIMENT 4**

**Thought detection**

Whereas Expts 2 and 3 tested children's understanding of mental states that, according to some theories (Baron-Cohen, 1993; Wellman, 1993), are the earliest to be understood in normal development (desire/goal and refer), in this final experiment we looked at their sensitivity to a more complex mental state, namely *thinking*. Given the earlier reported difficulties children with autism show in distinguishing the mental and the physical, and in understanding what it is for someone to have a thought (Baron-Cohen, 1989c, 1990), we predicted that they should also find it hard to identify from eye-direction alone when a person was thinking. We used a method that Baron-Cohen & Cross (1992) had used with normal 3-year-old children, as well as a cartoon version based on the technique used in Expt 1, to control for stimulus complexity.

**Materials and method**

In condition 1 we used 16 pairs of photographs of children's faces, comprising an equal number of male and female models of different ethnic groups. The photos were 3 square inches, black and white, and were reproduced from the Fairburn System of Visual References (1978). The expression of the mouth was either neutral or smiling. In each pair of photographs one face was looking to one side and in an *upwards* direction, and the other was looking directly forwards at the camera. The photos in which the

\(^3\) The exceptions to this were some of the older subjects with mental handicap, who did occasionally make arrow errors. We speculate that this age difference may relate to knowledge of the notational significance of the arrow symbol.
Gaze comprehension in autism

A child was looking away were all judged (by two independent raters) to depict 'thinking'. As the experimenter laid out each pair of photographs, the subject was simply asked, 'Which one is thinking?'.

In condition 2 we asked the same question of each child, but this time used six pairs of cartoon faces, indentical to those used in Expt 1, but where one face in each pair had its eyes centred (i.e. 'looking' directly forward), while the other had the black dot located in the upper left or right quadrant of the eyes (i.e. 'looking' upwards and away). Only one pair was laid out at a time and the position of the correct photo was randomly varied between left and right. A pass was scored if the child was correct on five out of six trials. An example of both the photo and cartoon stimuli pairs is shown in Fig. 4.

Figure 4. An example of a pair of stimuli used in Expt 4: photo condition (a) and cartoon condition (b).
Results and discussion

Mean scores by each group on each condition are shown in Table 5. ANOVAs revealed highly significant between-group differences on both conditions: photos \(F(2,67) = 49.57, p<.001\); cartoons \(F(2,67) = 19.92, p<.001\). Again, post hoc Scheffé tests showed these differences were between the group with autism and the other two groups \(p<.05\). Results from both conditions thus provided converging evidence for an autism-specific deficit in inferring when a person is thinking on the basis of eye-direction alone, as predicted, and this deficit was not explained by stimulus complexity. A similar picture emerged when the proportion of subjects in each group passing each test was analysed (see summary in Table 6). Of interest, while the same proportion of subjects with autism passed both the photo and the cartoons (15 per cent), a slightly larger proportion of the other groups passed on the photos than on the cartoons suggesting that, for the non-autistic subjects, the photos were slightly easier than the cartoons.

Table 5. Mean scores on each condition in Expt 4 by each group

<table>
<thead>
<tr>
<th></th>
<th>Photos (max. = 16)</th>
<th>Cartoons (max. = 6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(M)</td>
<td>15.2</td>
<td>5.1</td>
</tr>
<tr>
<td>(SD)</td>
<td>1.5</td>
<td>1.1</td>
</tr>
<tr>
<td>Mental handicap</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(M)</td>
<td>15.3</td>
<td>5.1</td>
</tr>
<tr>
<td>(SD)</td>
<td>1.4</td>
<td>0.8</td>
</tr>
<tr>
<td>Autism</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(M)</td>
<td>8.6*</td>
<td>3.5*</td>
</tr>
<tr>
<td>(SD)</td>
<td>4.2</td>
<td>1.2</td>
</tr>
</tbody>
</table>

* Autism × other two groups, \(p<.001\).

Effects of MA, CA and mental handicap subtype

In order to check if the autism-specific deficits in Expts 2–4 were in part due to either CA or MA, we compared the subjects with autism who passed on each task with those who failed, using \(t\) tests. None of these comparisons were significant at the \(p<.05\) level. In addition, in order to check whether the subgroup of individuals with William’s Syndrome was responsible for the group difference between the autism and mental handicap groups, the results from Expts 2–4 were reanalysed, this time excluding the 10 subjects with William’s Syndrome. Despite this, the differences between the autism and mental handicap groups remained significant at the \(p<.05\) level, even when the effects of MA were covaried.

Contingency analysis across different experiments

Because most subjects participated in all of the above experiments, we can report the contingencies between passing and failing different experiments. Of key interest is
Table 6. Summary table, showing percentage of each group passing on each condition in each experiment

<table>
<thead>
<tr>
<th></th>
<th>Normal</th>
<th>Mental handicap</th>
<th>Autism</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expt 1</td>
<td>Look$^a$</td>
<td>80</td>
<td>87</td>
</tr>
<tr>
<td></td>
<td>Silly$^a$</td>
<td>80</td>
<td>90</td>
</tr>
<tr>
<td></td>
<td>Emotion$^a$</td>
<td>100</td>
<td>97</td>
</tr>
<tr>
<td></td>
<td>Arrow$^b$</td>
<td>45</td>
<td>73</td>
</tr>
<tr>
<td>Expt 2</td>
<td>Desire$^c$</td>
<td>80</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>Goal$^c$</td>
<td>75</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>Refer$^a$</td>
<td>70</td>
<td>73</td>
</tr>
<tr>
<td></td>
<td>Size$^a$</td>
<td>85</td>
<td>73</td>
</tr>
<tr>
<td>Expt 3</td>
<td>Desire$^c$</td>
<td>85</td>
<td>80</td>
</tr>
<tr>
<td></td>
<td>Goal$^c$</td>
<td>80</td>
<td>76</td>
</tr>
<tr>
<td></td>
<td>Colour$^c$</td>
<td>90</td>
<td>88</td>
</tr>
<tr>
<td>Expt 4</td>
<td>Photos$^d$</td>
<td>95</td>
<td>93.3</td>
</tr>
<tr>
<td></td>
<td>Cartoons$^a$</td>
<td>70</td>
<td>66.6</td>
</tr>
</tbody>
</table>

$^a$ Autism × other two groups, Fisher's Exact Probability Test, $p<.05$.
$^b$ Pass = >5/6.
$^c$ Pass = >8/10.
$^d$ Pass = 3 out of 3.

whether the children with autism who pass in Expt 2 are the same individuals who pass Expts 3 and 4. Given that all of these three experiments require an understanding of mental states, one might expect that passing on one of these is related to passing on the other two experiments. This was in fact seen. Thus, eight subjects in the autism group passed the desire condition in Expt 2, six of whom also passed the desire condition in Expt 3. Of these, four also passed the goal condition in Expts 2 and 3. Finally, three of these same subjects passed in Expt 4, recognizing which face was 'thinking'.

In addition, 12 out of 20 of the subjects with autism had been given standard false belief tests as part of another study (Charman & Baron-Cohen, 1992), and in the case of the three subjects who passed Expt 4, all of these individuals also passed on the Sally Anne False Belief test (Baron-Cohen et al., 1985). None of the other nine (for whom we had relevant data) in this sample did so. In the final section, we turn to consider some possible interpretations of the results from these four experiments.

**GENERAL DISCUSSION**

Experiment 1 shows that children with autism are able to infer when a person is looking at them. This result is in line with findings that children with autism are unimpaired on tests of visual perspective taking (Baron-Cohen, 1989a, 1991b; Hobson, 1984; Leekam, Baron-Cohen, Perrett, Milders & Brown, 1993; Tan & Harris, 1991). In contrast, Expts
2–4 point to a severe deficit in children with autism in identifying the mental states of desire, goal, refer and think from the eyes, relative to the remarkable proficiency by young normal children and children with mental handicap to do so. Here we explore some possible reasons for this autism-specific deficit.

First, we can rule out either CA or verbal MA as explanations, given that the two clinical groups did not differ on these variables. We can also rule out non-verbal MA as an explanation, given that the group with autism was actually superior on this to the group with mental handicap. Problems in general language comprehension are also ruled out by the good performance on control questions by the group with autism. It should, however, be noted that, given the ceiling effects found in both of our control groups, it remains to be shown whether these two control groups would differ from each other if more complex tasks were employed. Such ceiling effects do not however affect the finding of an autism-specific deficit.

One possibility is that the difficulty in the group with autism was due to a failure to understand the specific mental state terms in the test questions. In support of this, Tager-Flusberg (1993) reports that children with autism hardly ever use the term think in their spontaneous speech. However, the term desire (‘want’) is one of the few mental state terms that children with autism do spontaneously produce; yet they failed the desire test in Expts 2 and 3. Furthermore, in Expts 2 and 3 the mental states of ‘goal’ and ‘intention to refer’ were not described with mental state terms. They were actually described in terms of behaviour (‘going to take’ and ‘say’, respectively), yet the children with autism failed the goal and refer tests. This makes it unlikely that their difficulty was due to the mental state terms themselves.

A more likely explanation is that the children with autism have difficulty in understanding the concepts of desire, goal, intention to refer and think, resulting in them not mapping these onto behaviour such as eye-direction. Consistent with this possibility, Phillips (1993; Phillips, Baron-Cohen & Rutter, 1995) has reported evidence that children with autism do have difficulties in understanding all of these concepts. Under this interpretation, it is the lack of mental state concepts that causes the failure to understand that eye-direction signifies this range of mental states.

Experiment 3 went on to demonstrate that normal children use the eyes as a natural cue to mentalistic information over and above other cues that could be used, such as the arrow in our display, while children with autism were more likely to use the arrow. One possible explanation for this might be an increased susceptibility to distraction on the part of the subjects with autism, given the problems in attention regulation that have been reported in this group (Dawson & Lewy, 1989). While this is possible, it still does not explain why they should have been distracted away from the eyes as cues to mental states, rather than towards them. Equally, the poor performance by the children with autism could have been due to an inability to resist the salience of the arrow stimulus, as some versions of the ‘frontal’ theory of autism have proposed (Hughes, Russell & Robbins, 1994). Against this, it should be noted that on the control condition they were (quite correctly) clearly able to ignore both the direction of the eyes and the arrow in making their judgement, suggesting at least under these conditions they are capable of inhibition. For our part, we interpret this, in the context of the other experiments, as showing that children with autism are blind to the mentalistic significance of the eyes.
The question we turn to next is how this might relate to other symptoms in autism. First, can this ‘mindblindness’ in relation to gaze interpretation explain the well-documented lack of joint attention in autism? If we accept that, in the normal case, infants monitor another’s gaze direction in order to check what the person is interested in and attempt to direct another’s gaze direction (by pointing, showing, etc.) in order to draw the person’s attention or interest to a shared focus, then given that interest and attention are basic mental states, the mindblindness hypothesis has some plausibility (Baron-Cohen, 1989b, 1991c, 1995c). The mental state of interest was not included in those tested here, but the prediction is that if children with autism were asked, ‘Which sweet is Charlie interest in?’, they again would fail to use eye-direction to solve this problem, since the present results show that they do not use this to infer other mental states.

Second, can the mindblindness hypothesis explain other findings in autism that relate to face processing, such as the relative insensitivity to affective expression (Yirmiya, Sigman, Kasari & Mundy, 1992)? We would not wish to overextend this explanation to abnormalities of affective sensitivity, since this latter phenomenon may be independent of the ability to read eye-direction in terms of the volitional and cognitive mental states that we have tested here. To clarify, we suggest that the mindblindness hypothesis has the power to explain a failure to read eye-direction in terms of the volitional mental states of goal, desire and intention and the cognitive mental states of interest, attention and thinking. (Refer is, we suppose, a particular instance of intention.)

Finally, the apparent ‘blindness’ to the significance of eye-direction on the part of the group with autism should be seen against what seems in some respects its inverse, in the subgroup with William’s Syndrome: these individuals performed almost exclusively at ceiling, despite their severe mental handicap. For them the tasks were trivially easy and their competence seemed to be mirrored in their real-life behaviour: they spent a great deal of time making eye-contact, as if reading the significance of the other’s eyes was for them an important part of social life. This may be a preserved islet of ability in this group, related to their development of a theory of mind (Karmiloff-Smith, 1992a, b; Karmiloff-Smith, Klima, Bellugi, Grant & Baron-Cohen, 1995), contrasting with the islet of severe deficit in the group with autism. It might be objected that to compare subjects with autism and William’s Syndrome is inappropriate, given the relative ‘sparing’ of social skills in the latter group (Reilly, Klima & Bellugi, 1991). However, this criticism is easily countered by the finding that the autism-specific deficits reported here were found when the subjects with William’s Syndrome were both included or excluded from the analysis. The value in their inclusion in this series of studies is to highlight those social–cognitive skills that may be relatively independent of general intelligence.

Acknowledgements

The first author was a member of the Departments of Child Psychiatry and Psychology at the Institute of Psychiatry, University of London, during the period of this work and was supported by grants from the Medical Research Council and the Mental Health Foundation. We are grateful to the following schools that helped us in this research: The Sybil Elgar School, Ealing; Griffin Manor School, Camberwell Grove Nursery, Camberwell; Rosemary School, Islington; the Idiopathic Infantile Hypercalcaemia Association, UK; and the Association du Syndrome du Williams, France; Annerley Primary School, Annerley; Burnt Ash Primary School, Bromley; and Churchfields Primary School, Beckenham. Figure
4 is reproduced from the Fairburn System of Visual References (1978). Unfortunately, the publishers of these no longer exist. Figure 3 is reproduced with permission from Nestlés. This work was presented by the first author at the European Developmental Psychology Conference, Seville, September 1992.

References


Gaze comprehension in autism


Received 29 September 1993; revised version received 4 October 1994