Impaired sadness recognition is linked to social interaction deficit in autism

Zillah Boraston a, Sarah-Jayne Blakemore b,*, Rebecca Chilvers a, David Skuse a

a Behavioural and Brain Sciences Unit, Institute of Child Health, University College London, London WC1N 1EH, UK
b Institute of Cognitive Neuroscience, University College London, London WC1N 3AR, UK

Received 27 June 2006; received in revised form 21 November 2006; accepted 21 November 2006
Available online 29 December 2006

Abstract

Can autistic individuals use motion cues to identify simple emotions from 2D abstract animations? We compared emotion recognition ability using a novel test involving computerised animations, and a more conventional emotion recognition test using facial expressions. Adults with autism and normal controls, matched for age and verbal IQ, participated in two experiments. First, participants viewed a series of short (5 s) animations. These featured an ‘emotional’ triangle, interacting with a circle. They were designed to evoke an attribution of emotion to the triangle, which was rated both in terms of anger, happiness, sadness or fear from its pattern of movement, and how animate (“living”) it appeared to be. Second, emotion recognition was tested from standardised photographs of facial expressions. In both experiments, adults with autism were significantly impaired relative to comparisons in their perception of sadness. This is the first demonstration that, in autism, individuals can have difficulties both in the interpretation of facial expressions and in the recognition of equivalent emotions based on the movement of abstract stimuli. Poor performance in the animations task was significantly correlated with the degree of impairment in reciprocal social interaction, assessed by the Autism Diagnostic Observation Schedule. Our findings point to a deficit in emotion recognition in autism, extending beyond the recognition of facial expressions, which is associated with a functional impairment in social interaction skills. Our results are discussed in the context of the results of neuroimaging studies that have used animated stimuli and images of faces.

© 2006 Elsevier Ltd. All rights reserved.

Keywords: Autistic; Asperger syndrome; Emotion; Animations; Endophenotype; Social cognition

1. Introduction

Autism is a pervasive developmental disorder, characterised by a triad of impairments: verbal and non-verbal communication problems, difficulties with reciprocal social interactions, and unusual patterns of repetitive behaviour (DSM-IV; American Psychiatric Association (APA), 1994). These impairments vary in their severity across the autism spectrum. For example, communication problems range from a complete absence of language in the most severe cases, to more subtle pragmatic language difficulties. Similarly, difficulties in reciprocal social interaction can manifest themselves as a complete absence of peer relationships, or problems such as unusual patterns of eye contact during social interaction. Individuals with autism vary considerably in IQ, with some suffering severe learning disability, while others have average or above average IQ. Asperger syndrome, a disorder on the autism spectrum, is characterised by an absence of language delay or learning disability (Frith, 2004). Autism is now widely regarded as a disorder of brain development, and numerous brain imaging studies have demonstrated both structural and functional anomalies in the autistic brain (for a review, see DiCicco-Bloom et al., 2006).

One of the most common observations about people with autism and Asperger syndrome is that they have problems interpreting the emotional states of other people. Empirical studies have shown emotion recognition problems in autism using tasks such as matching appropriate gestures, vocalisations, postures, contexts and facial expressions of emotion (Hobson, 1986a, 1986b), matching facial expressions of emotion across different individuals (Hobson, Ouston, & Lee, 1988), labelling facial expressions of emotion (Tantam, Monaghan, Nicholson, & Stirling, 1989; Bormann-Kischkel, Vilsmeier, & Baude, 1995) finding the ‘odd one out’ from a range of facial expressions (Tantam et al., 1989), and matching facial or prosodic
expressions of emotion with verbal and pictorial labels (Lindner & Rosen, 2006). Caregivers commonly report that children with autism fail to recognise emotions, and empirical studies have shown that they fail to react to signs of distress in adults (Sigman, Kasari, Kwon, & Yirmiya, 1992; Bacon, Fein, Morris, Waterhouse, & Allen, 1998).

This evidence of impairment in emotion recognition is supported by functional imaging studies demonstrating differences in the patterns of brain activity in autistic and control participants when they perform emotion recognition tasks. When matching a set of photographs according to facial expression, autistic children show less activation of the fusiform gyrus, compared with controls (Wang, Dapretto, Hariri, Sigman, & Bookheimer, 2005). A similar pattern of results was found by Piggot et al. (2004). Bolte et al. (2006) recently showed that training in facial emotion recognition did not boost the response of the fusiform gyrus to facial expressions in autistic individuals.

However, a significant number of studies have failed to find deficits in emotion recognition in individuals with autism (Ozonoff, Pennington, & Rogers, 1990; Buitelaar, van der Wees, Swaab-Barneveld, & van der Gaag, 1999; Castelli, 2005), or have found that deficits are restricted to particular emotions, such as surprise (Baron-Cohen, Spitz, & Cross, 1993) or fear (Howard et al., 2000). Inconsistencies in previous reports may be due to two factors. First, a wide range of testing paradigms and stimuli have been used across studies. Some studies have focussed on emotion recognition from a single cue, such as facial expression (e.g. Bormann-Kischkel et al., 1995; Castelli, 2005). Others have used a range of tests relying on different cues (Lindner & Rosen, 2006) or cross-modal paradigms (Hobson, 1986a, 1986b). Second, there is considerable variability in the composition of autistic groups studied (children or adults, high functioning or low functioning), and the matching of autistic and control groups (e.g. by verbal IQ or overall IQ). Ozonoff et al. (1990) observe that matching by verbal mental age (MA) is likely to be a more conservative approach than by overall MA or non-verbal MA since autistic individuals typically have a higher non-verbal than verbal MA. Congruent with this, all three of the studies mentioned above that matched groups on non-verbal measures of MA (Hobson, 1986a, 1986b; Tantam et al., 1989) found a deficit in the autistic group, whilst of the four studies that matched according to verbal ability, two found a deficit in the autistic group (Hobson et al., 1988; Lindner & Rosen, 2006) and two did not (Ozonoff et al., 1990; Castelli, 2005).

To complicate matters further, emotion recognition is not a unitary ability that may either be impaired or unimpaired. Many different personal cues can be used to recognise emotion, such as information from different parts of the face, tone of voice, gestures, and an individual’s behaviour and interaction with others. It is unclear whether, if autistic individuals are impaired in emotion recognition, their deficit affects the use of all cues, or is specific to certain cues such as facial expression. In this study, we tested the ability of a single group of participants to use two different cues to recognise emotion. We employed a set of standardised pictures of facial emotions, and also developed a novel test of emotion recognition, using abstract animations.

The animations were used to test a subject’s ability to use motion cues, such as speed and trajectory of movement, and movement in relation to others, to infer emotion. Similar abstract animations have been used for many decades to test mental state attribution (Heider & Simmel, 1944). Viewers tend to attribute mental states to geometrical shapes if they move in a ‘social’ way, i.e. if one shape moves in a self-propelled manner and is seen to influence the movements of another in a manner more complex than simple physical causality (Bloom & Veres, 1999; Heider & Simmel, 1944). Such animated stimuli activate brain areas that subserve a social cognitive function, including the superior temporal sulcus (STS), temporal poles, and medial prefrontal cortex (MPFC; Castelli, Happe, Frith, & Frith, 2000; Schultz et al., 2003).

In contrast to control subjects, people with autism do not spontaneously attribute mental states to such animations. Abell, Happe, and Frith (2000) devised a set of four animations involving two triangles, the movement of which was designed to evoke mental state attributions. In contrast, in the four control animations, the triangles appeared to be taking simple goal-directed actions. The subject’s task was to give a running commentary on the film, describing the actions of the triangles. Typically-developing children tended to attribute more complex mental states to the first set of animations than to the control animations. In the commentaries of children with autism, mental state descriptions were more likely to be inappropriate to the animation in question, compared with the control group. Castelli, Frith, Happe, and Frith (2002) found a similar response pattern in adults with high-functioning autism or Asperger syndrome.

Another way of isolating motion cues is to use point-light displays, in which a moving person is filmed in a dark room, with light sources attached to the joints of the body. Even with this limited amount of visual information, it is possible to recognise the particular emotion felt by the person represented by the point-lights (Heberlein, Adolphs, Tranel, & Damasio, 2004). Although the earliest studies of the recognition of biological motion from point-light displays did not indicate that people with autism were impaired in their interpretation (Moore and Hobson, 1997), subsequent studies have found some evidence of impairment (Blake, Turner, Smoski, Pozdol, & Stone, 2003), though these did not focus specifically on the recognition of emotion.

The most common stimuli used to test emotion recognition are pictures of facial expressions. It has been suggested that some autistic individuals might be able to use abnormal, compensatory strategies with this type of stimulus, which might therefore mask an underlying emotion recognition deficit (Teunisse & de Gelder, 2001). This is one reason behind the use of novel computerised animations in this study. These animations are a highly abstract type of stimulus, with which our participants would not have had prior experience. The stimuli therefore enabled us to test emotion recognition in the absence of compensatory strategies. The animations were designed to portray basic emotions (happy, angry, scared and sad). As well as being highly abstract, the task had the advantage of not being verbally demanding, as it was rated by responding with a key-press rather than a narrative description. It is therefore suitable for testing people with poor verbal skills.
Experiment 1 of this study employed this animation-based test of emotion recognition in an autistic and comparison population. In Experiment 2, we tested emotion recognition from a set of standardised and widely-used still photographs of facial expressions. It has been proposed that social information processing, including the interpretation of facial expressions and body movements, is linked to social interaction deficits in autism (Joseph & Tager-Flusberg, 2004). We tested the hypothesis that the interpretation of emotionally salient movement patterns as measured in our task would be related to an autistic participant’s social reciprocity as assessed by the reciprocal social interaction (RSI) subscale of the Autism diagnostic observation schedule (ADOS; Lord et al., 2000).

2. Experiment 1: Emotion recognition from animations

2.1. Methods

2.1.1. Participants

Two groups of participants were recruited by advertisement in literature of national autism groups and societies: 11 individuals with autism (9 males), and 11 typically developing controls (9 males). The groups were matched for age (see Table 1). All participants in the autistic group had a diagnosis of autism, Asperger syndrome or autistic spectrum disorder from a GP or psychiatrist. The diagnosis was confirmed by administering the Autism Diagnostic Observation Schedule (ADOS) (Lord et al., 2000). All participants gave informed consent, and the local ethics committee approved the study.

The groups were matched for verbal and performance IQ, as measured by administering all four subtests of the Wechsler Abbreviated Scale of Intelligence (WASI). Full details of the two groups can be found in Table 1.

Table 1
Details of the autistic and control groups

<table>
<thead>
<tr>
<th>Group</th>
<th>Subject</th>
<th>Age (years)</th>
<th>Verbal IQ</th>
<th>Performance IQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Autism</td>
<td>1</td>
<td>60.2</td>
<td>130</td>
<td>133</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>59.4</td>
<td>126</td>
<td>103</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>51.5</td>
<td>119</td>
<td>121</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>33.2</td>
<td>112</td>
<td>103</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>24.5</td>
<td>111</td>
<td>103</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>24.4</td>
<td>109</td>
<td>121</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>37.1</td>
<td>108</td>
<td>127</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>37.1</td>
<td>132</td>
<td>135</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>24.8</td>
<td>109</td>
<td>96</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>32.4</td>
<td>121</td>
<td>137</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>19.6</td>
<td>118</td>
<td>107</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>36.7</td>
<td>118</td>
<td>117</td>
</tr>
<tr>
<td>NC</td>
<td>1</td>
<td>57.6</td>
<td>135</td>
<td>127</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>59.8</td>
<td>127</td>
<td>124</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>53.2</td>
<td>99</td>
<td>108</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>31.4</td>
<td>89</td>
<td>85</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>24.9</td>
<td>98</td>
<td>111</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>24.4</td>
<td>115</td>
<td>116</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>33.8</td>
<td>99</td>
<td>117</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>36.4</td>
<td>109</td>
<td>111</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>20.7</td>
<td>96</td>
<td>119</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>32.8</td>
<td>119</td>
<td>127</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>20.7</td>
<td>107</td>
<td>105</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>33.8</td>
<td>108</td>
<td>114</td>
</tr>
<tr>
<td>S.D.</td>
<td></td>
<td>13.2</td>
<td>14.4</td>
<td>12.0</td>
</tr>
<tr>
<td>Group comparison</td>
<td>( t_{(20)} = 0.128, ) NS</td>
<td>( t_{(20)} = 1.863, ) NS</td>
<td>( t_{(20)} = 0.565, ) NS</td>
<td></td>
</tr>
</tbody>
</table>

Fig. 1. Frame from one of the 12 animations, showing the circle and triangle, and the following question. Eight animations were designed to evoke the attribution of a particular emotion (angry, happy, sad or scared), whereas in the other four the shapes were designed to appear inanimate. Animations were followed by a question in which the participant selected their response from a numerical rating scale.

All participants were screened for exclusion criteria by self-report prior to taking part in the study. Those with dyslexia, epilepsy, or other neurological or psychiatric conditions were excluded. The screening process also included a short questionnaire, in which participants were required to give unambiguous definitions (as judged by the experimenter) of the emotions under investigation: ‘angry’, ‘happy’, ‘sad’ and ‘scared’. All participants were able to do this task without difficulty.

2.1.2. Animations

Twelve silent animations were used, each featuring a black outline triangle and a black outline circle moving on a white background (Fig. 1). In eight of the animations, the triangle moved in a self-propelled, non-linear manner, designed to evoke the impression that it was ‘living’ or animate (Scholl & Tremoulet, 2000; Blakemore et al., 2003). Each of these eight animations was designed to evoke the attribution of a particular emotion: angry, happy, sad or scared, with two examples of each. For example, in a ‘scared’ animation, the triangle appeared...
to ‘run away’ from the circle, whereas in a ‘happy’ animation it approached the circle in an affectionate manner. In an ‘angry’ animation, the triangle jabbed repeatedly at the circle, and in a ‘sad’ animation it ‘pushed away’ the circle when it was approached. The animations were of similar duration for all emotions (angry = 5.5 s, happy = 5.6 s, sad = 5.5 s, scared = 5 s).

In the other four animations, the triangle moved in a manner designed to make it appear inanimate (‘non-living’). For example, it might appear to be moving as if falling under gravity. The trajectory of movement of the circle was identical in all 12 animations. Animations were designed using Flash MX 2004, and presented to the subjects using Matlab v6.5, and Cogent Graphics, on a Dell Latitude 100L laptop computer with a 15” LCD display screen.

2.1.3. Design and procedure
Participants viewed each living animation three times, and each non-living animation once, making a total of 28 presentations, in a pseudorandom order. After each presentation of an animation, a question appeared on the screen. There were two types of question: ‘emotion’ questions, and ‘living’ questions.

‘Emotion’ questions were of the format: ‘was the triangle ANGRY?’ For each living animation, an emotion question followed one of the two presentations. One of these referred to the actual emotion intended to be perceived in the animation. The other referred to an alternative emotion that was not intended to be perceived. Emotion questions were not asked for the non-living animations. Participants answered the emotion question using a rating scale from 0 (“not at all…” ) to 5 (extremely… ”).

‘Living’ questions were of the format ‘was the triangle LIVING?’ A living question followed each of the four non-living animations, and one presentation of each of the eight living animations. Each living animation was therefore shown twice with an ‘emotion’ question, and once with a ‘living’ question. Again for the ‘living’ question, the participant answered using a rating scale from 0 (“definitely non-living”) to 5 (“definitely living”). The rating scale was explained to each participant before the experiment, and participants completed four practice trials.

2.1.4. Control task
Before the main part of the task, participants completed a control task, with four animations—a subset of the animations used in the main task. Participants were required to indicate, using the same 0–5 rating scale, the vertical position of the triangle on the screen at the end of animations, with 5 representing the top of the screen, and 0 representing the bottom. This task was designed to ensure that the participant was able to perceive and attend to the animations, read the questions, and use the rating scale. Participants who did not complete this task correctly were excluded from the study. All participants completed the task correctly and so no participants were excluded.

2.1.5. Data analysis
For each living animation, participants gave a rating for the actual emotion present, and an alternative emotion. First, we calculated the mean ratings for the actual and alternative emotions for both subject groups, for each of the four emotions. Accurate recognition of the emotion would be indicated by a high rating for that emotion, and a low rating for any of the three possible alternative emotions. We therefore calculated a score for each subject by subtracting the alternative emotion rating from the actual emotion rating. This gave a score ranging from −5 to +5, with a higher score indicating a greater accuracy in correctly identifying the actual emotion. This meant that a participant could not get a high score by indiscriminately giving high ratings in response to all questions. The scores were averaged for each of the two presentations of each emotion for each subject, and the averages were used in the statistical analysis.

As ratings of the four emotions were considered to be separate dependent variables, and emotion type was not considered to comprise different levels of the same continuous factor, data were analysed using a multiple analysis of variance (MANOVA), with two groups (autism and control) and four emotions (angry, happy, sad and scared). Cohen’s d (Cohen, 1988) was also calculated as a measure of effect size for each of the four emotions, independent of sample size. Cohen’s d is obtained by dividing the difference between the two group means by the pooled standard deviation of the groups.

For the ‘living’ task, the mean ratings of the two groups were compared for each of the emotional animations, and for the animations that were designed to appear inanimate, using a MANOVA, with two groups and five conditions.

2.1.6. Reliability of test
As this test is a novel measure, its reliability was assessed using a test-retest paradigm. Twenty normal adults (10 males; mean age = 26.2 ± 4.0 years) were tested twice, using the procedure above, with an interval of between 5 and 14 days. On each testing occasion, a single score was calculated for each subject for emotion recognition, averaged across all four emotions. These scores were compared for time 1 and time 2 as a measure of reliability, and used to calculate an Intraclass Correlation Coefficient (Shrout & Fleiss, 1979), as well as a Technical Error of Measurement (Mueller & Martorell, 1988).

2.2. Results

2.2.1. Animations: Emotion recognition task
Fig. 2a shows how the autistic and control groups rated each animation type for the actual and alternative emotions. As described earlier, these ratings were subtracted to give a score for each participant for each emotion. This was done for two reasons. First, a difference score more accurately indicates how well each participant could identify the correct emotion for each animation. Second, these difference scores were more normally distributed than the raw ratings, and thus more appropriate for the application of parametric statistical tests.

Fig. 2b shows these calculated scores for the two participant groups. Control participants scored higher than autistic
Table 2
Scores for all participants

<table>
<thead>
<tr>
<th>Group</th>
<th>Subject</th>
<th>Emotion recognition - animations</th>
<th>Living judgement - animations</th>
<th>Emotion recognition - faces</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Angry</td>
<td>Happy</td>
<td>Sad</td>
</tr>
<tr>
<td>Autism 1</td>
<td>3.0</td>
<td>0.5</td>
<td>3.5</td>
<td>−0.5</td>
</tr>
<tr>
<td>2</td>
<td>3.0</td>
<td>1.5</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>3</td>
<td>5.0</td>
<td>2.5</td>
<td>0.5</td>
<td>4.5</td>
</tr>
<tr>
<td>4</td>
<td>0.0</td>
<td>2.5</td>
<td>0.0</td>
<td>3.0</td>
</tr>
<tr>
<td>5</td>
<td>1.0</td>
<td>2.0</td>
<td>1.5</td>
<td>2.0</td>
</tr>
<tr>
<td>6</td>
<td>4.0</td>
<td>5.0</td>
<td>3.0</td>
<td>5.0</td>
</tr>
<tr>
<td>7</td>
<td>3.5</td>
<td>2.0</td>
<td>1.5</td>
<td>5.0</td>
</tr>
<tr>
<td>8</td>
<td>1.5</td>
<td>5.0</td>
<td>0.5</td>
<td>1.0</td>
</tr>
<tr>
<td>9</td>
<td>4.5</td>
<td>4.0</td>
<td>1.5</td>
<td>3.0</td>
</tr>
<tr>
<td>10</td>
<td>4.0</td>
<td>−1.0</td>
<td>2.5</td>
<td>1.5</td>
</tr>
<tr>
<td>11</td>
<td>1.0</td>
<td>0.0</td>
<td>0.5</td>
<td>3.0</td>
</tr>
<tr>
<td>Mean</td>
<td>2.77</td>
<td>2.18</td>
<td>1.45</td>
<td>2.59</td>
</tr>
<tr>
<td>S.E.</td>
<td>0.50</td>
<td>0.58</td>
<td>0.34</td>
<td>0.54</td>
</tr>
<tr>
<td>NC 1</td>
<td>2.5</td>
<td>2.0</td>
<td>2.0</td>
<td>2.5</td>
</tr>
<tr>
<td>2</td>
<td>5.0</td>
<td>2.5</td>
<td>2.5</td>
<td>3.5</td>
</tr>
<tr>
<td>3</td>
<td>5.0</td>
<td>3.0</td>
<td>3.0</td>
<td>1.5</td>
</tr>
<tr>
<td>4</td>
<td>5.0</td>
<td>4.0</td>
<td>3.5</td>
<td>1.0</td>
</tr>
<tr>
<td>5</td>
<td>1.5</td>
<td>4.5</td>
<td>2.0</td>
<td>1.0</td>
</tr>
<tr>
<td>6</td>
<td>5.0</td>
<td>5.0</td>
<td>3.5</td>
<td>3.0</td>
</tr>
<tr>
<td>7</td>
<td>2.5</td>
<td>3.5</td>
<td>2.5</td>
<td>1.5</td>
</tr>
<tr>
<td>8</td>
<td>−1.0</td>
<td>0.5</td>
<td>1.5</td>
<td>1.0</td>
</tr>
<tr>
<td>9</td>
<td>1.5</td>
<td>2.5</td>
<td>1.5</td>
<td>2.0</td>
</tr>
<tr>
<td>10</td>
<td>4.0</td>
<td>0.5</td>
<td>5.0</td>
<td>5.0</td>
</tr>
<tr>
<td>11</td>
<td>5.0</td>
<td>2.5</td>
<td>3.5</td>
<td>5.0</td>
</tr>
<tr>
<td>Mean</td>
<td>3.27</td>
<td>2.77</td>
<td>2.77</td>
<td>2.45</td>
</tr>
<tr>
<td>S.E.</td>
<td>0.61</td>
<td>0.44</td>
<td>0.31</td>
<td>0.46</td>
</tr>
</tbody>
</table>

Emotion recognition scores are generated by subtracting the ratings for alternative emotions from those for actual emotions, and range from −5 to +5. A positive score indicates some ability to discriminate correctly between actual and alternative emotion labels. Living ratings of the animations range from 0 to 5. Emotion recognition scores from faces range from 0 to 10.

participants for three of the four emotions. Scores for all conditions were greater than zero, indicating that both participant groups had some ability to distinguish the actual emotion label from the alternative one. Scores for individual participants are shown in Table 2.

The MANOVA revealed no significant difference between the groups for angry ($F_{(1, 20)} = 0.40; p = 0.53$), happy ($F_{(1, 20)} = 0.65; p = 0.43$) or scared ($F_{(1, 20)} = 0.04; p = 0.85$) animations, but a significant difference for sad animations ($F_{(1, 20)} = 7.99; p = 0.01$). These data are shown in Fig. 2b. Table 2 shows that all of the control participants had higher sadness recognition scores than the mean score of the autistic group. The calculated effect sizes revealed a large effect of group only for sadness ($d = 1.26$), a medium effect size for happiness ($d = 0.36$) and anger ($d = 0.28$), and almost no effect for fear ($d = 0.09$).

2.2.2. Animations: Animacy recognition task

The MANOVA revealed no significant difference between the autistic and control groups for the ‘living’ rating of any emotion (angry: $F_{(1, 20)} = 0.06; p = 0.82$; happy: $F_{(1, 20)} = 1.73; p = 0.20$; sad: $F_{(1, 20)} = 1.37; p = 0.82$; scared: $F_{(1, 20)} = 0.77; p = 0.39$), or for the ‘living’ rating of the animations that were designed to appear inanimate ($F_{(1, 20)} = 0.43; p = 0.52$). However, sad animations were rated as less ‘living’ by both autistic and control participants (see Fig. 3). The ratings of all participants are shown separately for the four emotions in Table 2.

2.2.3. Task scores compared to degree of social impairment

Module 4 of the ADOS (Lord et al., 2000) was used to confirm the diagnosis of each autistic participant, by assessing his or her behaviour in social interaction with the examiner. In this module of the ADOS, ratings of impairment are derived...
in terms of subscales measuring reciprocal social interaction, social communication skills, imagination/creativity and repetitive behaviour patterns. A higher score on a subscale indicates a greater degree of impairment.

For the participants with autism, we correlated emotion recognition performance with their social interaction skills as assessed by the RSI measure of the ADOS. A Spearman’s non-parametric test was used to assess the correlation between each participant’s score on the RSI scale, and his or her emotion recognition score from the animation task. As a comparison, we also assessed the correlation between the RSI scores and the mean ‘living’ rating of the sad animations. There was a significant negative correlation between magnitude of impairment on the RSI subscale and the sadness recognition score ($r^2 = 0.64$; $p = 0.002$, 1-tailed). Greater social impairment correlated with poorer recognition of sadness. There was no significant correlation between the RSI score and the ‘living’ rating of the sad animations ($r^2 = 0.004$; $p = 0.42$, 1-tailed). These results can be seen in Fig. 4. Correlations between RSI scores and emotion recognition scores for the other emotions were not significant (angry: $p = 0.16$, happy: $p = 0.30$, scared: $p = 0.47$, all 1-tailed).

2.2.4. Reliability of test

The test-retest data were from 20 normal adults, each tested on two occasions, between 5 and 14 days apart. The scores for each subject from time 1 and time 2 were subtracted to give a difference score. The mean difference score was $0.03 \pm 0.96$, with only four of the 20 participants obtaining a difference score of greater than one point. A Bland–Altman plot (Bland & Altman, 1986) of these differences (Fig. 5) shows that the difference score does not depend on the subject’s overall score, i.e. the reliability of the test is not affected by the level of performance. Scores from the two testing sessions gave an intraclass correlation coefficient of 0.69, which falls within the ‘good’ range (Cicchetti, 1994), and a technical error of measurement of 0.66, indicating reasonable agreement between testing sessions.

3. Experiment 2: Emotional faces task

3.1. Methods

3.1.1. Participants

All 11 of the autistic participants from Experiment 1, and nine of the control participants (seven males), took part in this study (mean age of controls: $34 \pm 15.0$ years, mean verbal IQ: $107 \pm 15.6$, mean performance IQ: $112 \pm 12.5$).

3.1.2. Stimuli

This task used 60 black and white photographs from a standard set of pictures of facial affect (Ekman & Friesen, 1976). These comprised photos of 10 different individuals, each showing facial expressions of anger, happiness, sadness, fear, surprise and disgust.

3.1.3. Design and procedure

Participants viewed each photograph, presented on a computer screen, for as long as they liked, and selected the appropriate emotion from a list of six words adjacent to the photograph. The task was preceded by a short practice task containing six images, one of each of the six possible emotions. The individual in the practice images was different from those in the main task.

Fig. 5. Bland–Altman plot of participants’ scores at time 1 and time 2 in the test-retest paradigm. The x-axis shows the mean score of each subject (maximum score = 5; a positive score indicates some ability to discriminate correctly between actual and alternative emotion labels). The y-axis shows the difference between that subject’s scores for time 1 and time 2.
3.1.4. Data analysis

As in the animations task, we considered the score for each emotion to be a separate dependent variable, and emotion type not to comprise different levels of the same factor. Therefore, data were analysed using a MANOVA, with two groups (autism and control) and six emotions (angry, happy, sad, scared, disgusted and surprised). As with the animations task, Cohen’s d was calculated from autistic and control data for each of the six emotions, as a measure of effect size.

We did not test reliability of the task on this occasion, but the reliability of a computerised presentation of the Ekman–Friesen faces has been evaluated as satisfactory in a separate study (Skuse, Lawrence, & Tang, 2005).

3.2. Results

The autistic group scored slightly lower than the control group in the recognition of all emotions apart from disgust, in which scores for the two groups were equal. The MANOVA revealed a significant difference between the autistic and control groups for the sad faces ($F_{(1, 20)} = 8.11; p = 0.01$), but no difference for the other emotions (angry: $F_{(1, 20)} = 1.53, p = 0.23$; happy: $F_{(1, 20)} = 0.82, p = 0.38$; scared: $F_{(1, 20)} = 1.26, p = 0.28$; surprised: $F_{(1, 20)} = 0.171, p = 0.684$; disgusted: $F_{(1, 20)} = 0, p = 1$). Cohen’s d revealed a large effect size of group for sadness recognition (d = 1.34), medium effect sizes for some other emotions (angry: d = 0.58, happy: d = 0.42, scared: d = 0.53), and little or no effect for surprise (d = 0.19) and disgust (d = 0). Fig. 6 shows the mean scores for the two groups for all six emotions.

As in the animations task, a Spearman’s non-parametric test was used to assess the correlation between an autistic individual’s score on the RSI scale, and their emotion recognition scores, for the four emotions common to both the animations and faces task. The correlations were not significant for any emotion (angry: $p = 0.47$, happy: $p = 0.11$, sad: $p = 0.26$, scared: $p = 0.58$, all 1-tailed).

A Spearman’s test was also used to assess the correlation between an individual’s sadness recognition score from the animated task and from the face task. This was conducted separately for the two participant groups, and was not significant for either group (autism: $p = 0.31$, control: $p = 0.50$, both 1-tailed).

The lack of significance in tests of correlation using the face task scores may have been due to the relatively small range of possible scores, and the fact that scores are integers.

4. Discussion

The main finding of this study was a deficit in sadness recognition from both animations and photographs of faces, in the autistic group. This effect was not simply due to the sadness task being more difficult: the scores of control participants do not suggest this was the case in either experiment (see Figs. 2b and 6). Indeed other studies have found sadness to be one of the emotions most easily recognised by normal individuals (Ekman & Friesen, 1978). That the sadness recognition deficit in autism was found in both experiments suggests that it is not specific either to facial expressions or motion cues. These results suggest that neural systems that underlie emotion recognition, in particular the recognition of sadness, might be abnormal in autism.

First of all, however, it is necessary to consider alternative explanations for the apparent sadness recognition deficit indicated by the data, with particular regard to the animations task, as this is a novel measure. Poor performance on the task could occur for a number of reasons. Firstly, the task requires the processing of visual motion, which has been shown to be abnormal in autistic individuals (Milne et al., 2002; Spencer et al., 2000). Second, the task relies on holistic viewing of the animations to identify the emotions successfully, and there is evidence in autism of impaired global processing (Frith, 1989) or a bias towards local processing (“weak central coherence theory”; Happe, 1996).

Third, a failure to recognise the animated shapes as living might impair the attribution of the correct emotion to the triangle. The scores of the autism group on the ‘living’ task from Experiment 1 suggest these alternative explanations cannot explain the emotion recognition data. There was no significant difference between the two groups’ perception of animacy in the animations in this task. The living task has the same demands as the emotion recognition task in terms of motion processing, and in the holistic viewing of the animations. The fact that the autistic participants were able to perform this living task as well as controls suggests that poor performance on the sadness recognition task was not due to a general visual motion processing deficit, impaired global processing or a failure to recognise the shapes as living. This suggests that low scores on the animated emotion recognition task in the autistic group reflect a deficit in sadness recognition. In addition, none of these alternative explanations can be used to explain the impairment in sadness recognition from static faces that we found in Experiment 2.

In addition, an interesting finding from the ‘living’ task is that both the autistic and control groups were less inclined to rate the sad animations as living compared with the other emotions. The nature of the triangle’s movement in this cartoon may have contributed to the relatively low living ratings for control and autistic participants.

4.1. Tasks using animations

The animations task used in this study is less verbally demanding than existing animated tests using abstract shape
interactions (e.g. Klin, 2000; Castelli et al., 2000), making it more suitable for testing child clinical populations with language delay. Using abstract, unfamiliar stimuli avoids the potential confound of the use of compensatory strategies by high-functioning autistic individuals, which might mask an underlying deficit for emotion recognition if familiar stimuli such as facial expressions are used (Teunisse & de Gelder, 2001).

The test-retest results presented here demonstrate a reasonable level of reliability within the task. Furthermore, tests with an epidemiological sample of 2000 adolescents (Boraston et al., unpublished data) revealed performance that was significantly above chance in the recognition of all four basic emotions, with no ceiling effect. The animations used in the task reliably evoke the perception of specific emotions, but the precise neural impairment that causes poor performance in autism is not yet known. As work in this field progresses, it will be useful to combine animated stimuli with eye-tracking technology, to examine patterns of eye movement when performing this task.

4.2. Sadness recognition deficit in autism

There have been no previous reports of specific deficits in sadness recognition in the autistic population, though a frequent comment from parents of autistic children is that they fail to recognise when their parents are upset. The mother of one autistic girl commented that she (the mother), “broke down in tears in front of her and she didn’t notice” (Skuse, personal communication). Deficits in recognising sad facial expressions tend to be associated with psychopathic traits (Blair, 1995; Stevens, Charman, & Blair, 2001; Blair & Coles, 2000). Blair and Coles (2000) found in a study of normal adolescent children that problems recognising sadness and fear were associated with a higher level of affective-interpersonal disturbance in the form of callous or unemotional (CU) traits. Rogers, Viding, Blair, Frith, and Happe (2006) found that the recognition of sadness from faces was particularly impaired in boys with autism who also had high CU tendencies compared with boys with autism with low CU tendencies. We did not specifically measure psychopathic traits in the samples under investigation in the current study.

There has been relatively little work on the neural correlates of sadness recognition. Blair, Morris, Frith, Perrett, and Dolan (1999) showed that viewing sad facial expressions activated the left amygdala and right temporal pole. Goldin et al. (2005) found that viewing sad films resulted in activations in a network of regions including the medial prefrontal cortex, superior temporal gyrus, precuneus, lingual gyrus and the amygdala. A deficit in sadness recognition could therefore be explained by disrupted amygdala-cortical connectivity (Schultz, 2005). Wang, McCarthy, Song, and Labar (2005) also found that the amygdala was activated by sad images, though a study by Killgore and Yurgelun-Todd (2004) failed to find amygdala activation.

4.3. Emotions other than sadness

In contrast to previous studies (Howard et al., 2000; Pelphrey et al., 2002), we did not find any significant deficit in fear recognition from facial expression in the autistic group. On inspection of the data, this appears to be because although the difference between the group means is comparable to that for sadness, the variance in the autistic group is much larger for fear than for sadness (Fig. 6). This heterogeneity could explain why some studies have found a fear recognition deficit whilst others have not (e.g. Buitelaar et al., 1999; Castelli, 2005).

We also found no deficit in fear recognition from abstract animations (Fig. 2). In this task, scores were almost identical for the autism and control groups. This suggests that the fear recognition deficits found in previous studies might be specific to the perceptual processing of faces, and therefore not extend to abstract stimuli. Results from eye-tracking studies with amygdala lesion patients (Adolphs et al., 2005) and adults with autism (Spezio, Adolphs, Hurley, & Piven, 2006, 2007; Dalton et al., 2005) suggest that specific deficits in fear recognition from facial expression could be caused by a failure to look at the eyes. This would not affect the perception of fear from animations (in which eyes are absent).

Our study showed no significant difference between the autistic and control groups in the recognition of emotions other than sadness in either experiment. However, it is notable that the autistic individuals scored slightly lower than controls on the recognition of all emotions in the faces task (Fig. 6), and all but one emotion in the animations task (Fig. 2b), implying that this deficit may not be specific to sadness. In addition to large effect sizes for sadness, the values of Cohen’s d obtained indicate moderate effect sizes for anger and happiness in both experiments, so testing with a larger sample might reveal significant differences for these other emotions. It should be noted that in the faces task, a ceiling-effect for happiness makes it difficult to determine whether the autistic group was impaired in the recognition of this emotion.

4.4. Cross-modal nature of deficit

We found parallel deficits in both experiments, despite the difference in nature of the two tasks. This could be linked to the fact that both tasks involved ‘social’ stimuli. There is evidence that visual information signalling the activities and intentions of other humans is processed in and near the STS (Allison, Puce, & McCarthy, 2000). The STS is activated by a number of social cognitive tasks, including emotion recognition from facial expressions (Narumoto, Okada, Sadato, Fukui, & Yonekura, 2001) and also the attribution of mental states to abstract animations (Castelli et al., 2002).

The STS has strong functional connections with the amygdala, which is involved in the perception of social cues, such as the recognition of facial expressions of emotion (Fitzgerald, Angstadt, Jelsone, Nathan, & Phan, 2006). The impairment of autistic individuals in perceiving and interpreting social cues from others has features in common with cases of amygdala damage (Adolphs, Sears, & Piven, 2001), and there is evidence that circuits involving the amygdala are dysfunctional in autistic individuals (Bachevalier & Loveland, 2006). It has been suggested that impaired amygdala-cortical connectivity could underlie some of the deficits seen in autism (Schultz, 2005).
4.5. Reciprocal social interaction impairment and emotion recognition

As well as sadness recognition being significantly impaired for autistic individuals, there was a significant correlation between sadness recognition and reciprocal social interaction skills as assessed by the ADOS (Lord et al., 2000). This is the first study to demonstrate a correlation between a measure of emotion recognition and the severity of autistic symptoms. This fits with evidence that the level of social impairment in autism correlates with behavioural variables such as eye-gaze patterns (Klin, Jones, Schulz, Volkmann, & Cohen, 2002), and brain responses to socially salient stimuli (Dapretto et al., 2006).

5. Conclusion

We found evidence for impaired sadness recognition in autism, using tasks that relied on two quite different cues—facial expressions and movement patterns. Given the evidence that similar brain areas, in particular the STS, are activated by both images of faces and abstract animated stimuli, abnormal functioning of this brain region could lead to a deficit in emotion recognition of this type. Further experiments are needed to investigate this.

Acknowledgements

This study was supported by the Wellcome Trust and the Royal Society, UK. Z.B. is supported by a Wellcome Trust four-year PhD Programme in Neuroscience at University College London.

References


