Attention to social stimuli and facial identity recognition skills in autism spectrum disorder

C. E. Wilson,1,2 J. Brock1 & R. Palermo3

1 Macquarie Centre for Cognitive Science, Macquarie University, Sydney, Australia
2 Forensic and Neurodevelopmental Science, Institute of Psychiatry, London, UK
3 Department of Psychology, Australian National University, Canberra, Australia

Abstract

Background Previous research suggests that individuals with autism spectrum disorder (ASD) have a reduced preference for viewing social stimuli in the environment and impaired facial identity recognition.

Methods Here, we directly tested a link between these two phenomena in 13 ASD children and 13 age-matched typically developing (TD) controls. Eye movements were recorded while participants passively viewed visual scenes containing people and objects. Participants also completed independent matching tasks for faces and objects.

Results and Conclusions Behavioural data showed that participants with ASD were impaired on both face- and object-matching tasks relative to TD controls. Eye-tracking data revealed that both groups showed a strong bias to orient towards people. TD children spent proportionally more time looking at people than objects; however, there was no difference in viewing times between people and objects in the ASD group. In the ASD group, an individual’s preference for looking first at the people in scenes was associated with level of face recognition ability. Further research is required to determine whether a causal relationship exists between these factors.

Keywords autism, face recognition, social attention

Introduction

Individuals with autism spectrum disorder (ASD) show lifelong deficits in social interaction. Several studies have reported an impairment in facial identity recognition (e.g. Tantam et al. 1989; de Gelder et al. 1991; Boucher & Lewis 1992; Gepner et al. 1996; Hauck et al. 1998; Klin et al. 1999; Serra et al. 2003; Faja et al. 2008; Joseph et al. 2008), although many others find no evidence of impairment (e.g. Celani et al. 1999; Schultz et al. 2000; Deruelle et al. 2004; Lahaie et al. 2006). Discrepancies between previous studies are likely accounted for by the wide variation in face recognition skills that exists within the ASD population (C.E. Wilson et al. unpublished data). However, the reason why impaired face recognition emerges in the profile of some ASD individuals, but not others, is currently unclear.

Schultz (2005) hypothesised that impaired face recognition in ASD was a consequence of a lack of bias to attend to social stimuli. In line with diagnostic criteria, ASD individuals are less likely to approach other people or to express an interest in
peers, they often fail to respond to their name being called, exhibit reduced social smiling and eye contact and abnormal facial expressions (DSM IV; APA, 2000). Preferential looking studies show that toddlers with ASD fail to show the normal preference for their mother’s voice over unintelligible babble (Klin 1991, 1992) and show a reduced preference for looking at point-light animations of human actions (Klin & Jones 2008). Retrospective analyses of home videos found ASD infants spent less time looking towards people and their faces in comparison to typically developing (TD) infants during the first year of life (Osterling & Dawson 1994; Werner et al. 2000; Osterling et al. 2002). However, because ASD is rarely diagnosed before the age of 2, direct evidence of ASD infants’ viewing preferences is scarce.

C.E. Wilson et al. (unpublished data), reported indirect evidence for an association between face recognition ability and attendance to faces in infancy by looking at the advantage for own-race faces (the ‘other-race effect’) in ASD children. The other-race effect is thought to emerge as a result of experience with one’s own-race faces during a sensitive period in early childhood (cf. Nelson 2001; Kelly et al. 2007, 2009). C.E. Wilson et al. (unpublished data) found that children with ASD who performed poorly on a face-matching test also failed to show an other-race effect. This suggests that a lack of attendance to face stimuli in young ASD children had a detrimental affect on the development of specialised face processing mechanisms.

Research suggests atypical viewing preferences continue to be exhibited in the ASD population throughout later childhood and adolescence. For example, Sasson et al. (2007) found that ASD children fixated relatively less on pictures of faces and more on objects in a visual array in comparison to controls. Riby and Hancock found ASD adolescents spent less time looking at faces in scenes (Riby & Hancock 2008, 2009a), and relatively less time looking at the people compared with the background in photographs of natural scenes, compared with controls (Riby & Hancock 2009b). Fletcher-Watson et al. (2009) found that ASD adults and controls showed a similar preference for looking at the pictures containing people compared with the pictures with no people. However, ASD adults were slower to fixate on the pictures with people present relative to pictures with people absent, suggesting a reduced attentional priority for social stimuli.

In the current study, we directly tested how attendance to people vs. objects in visual scenes was associated with ability to recognise images of faces and objects. We recorded eye movements while TD and ASD children passively viewed scenes containing social and non-social elements. They also completed independent tests of face and object matching. We predicted that performance on the face-matching task would be associated with bias to attend to social information. Specifically, participants with good face recognition skills would be show a greater preference for fixating on people before objects, and would also spend relatively more time looking at people in comparison to objects in the scenes.

Method

Participants

Thirteen participants (nine males) were recruited through Autism Spectrum Australia and Macquarie University Special Education Centre. All had received a formal diagnosis of autism (n = 8) or Asperger syndrome (n = 5) using the Autism Diagnostic Interview Revised (Lord et al. 1994), the Autism Diagnostic Observation Schedule (Lord et al. 2001) or the Childhood Autism Rating Scale. In addition, each child achieved scores indicative of ASD on the Social Communication Questionnaire lifetime (SCQ) (Rutter et al. 2003). Control participants were 13 TD children (six males).

Sentence comprehension was measured using the Test for Reception Of Grammar-2 (TROG, Bishop 2003), and nonverbal fluid reasoning was measured using the matrices subscale of the Weschler Abbreviated Scale of Intelligence (WASI matrices; Weschler, 1999). Neither measure of general ability was related to performance on either task, in ASD or TD children. Performance on the face- and object-matching tasks was related to chronological age, therefore we recruited a further 19 TD participants (eight males, mean age = 9.03, SD = 2.01) to complete the behavioural component so that age-standardised scores could be calculated.

Characteristics of participants that completed the eye-tracking task are provided in Table 1.
Behavioural tasks: matching faces and objects

Stimuli

Twenty-five pairs of Caucasian young male identities were originally taken from a face research database of the Glasgow Face Recognition Group (http://www.psy.gla.ac.uk/~mike/facerec.html) and had previously been used in a study by Wilson and colleagues (C.E. Wilson et al. unpublished data). For each identity there were two photographs, taken with a different camera under different lighting conditions, and transformed to greyscale. Each identity appeared twice, once as a distracter and once as a target. For the objects task, stimuli were photographs of 20 pairs of different running shoes. As with the facial identity task, the photos were taken with two different cameras under different lighting conditions, and transformed to greyscale.

Procedure

Both tasks were 2-alternative forced choice matching tests, completed on a 32 ¥ 28 cm touch screen monitor. Participants were told to touch the picture that showed the same face/object as the first picture. Each test included eight practice trials, followed by 50 test trials, with a break mid-way through. To start each trial the participant had to touch a cross that was presented in the centre of the screen, and upon release, the test item (face/object) appeared for 1000 ms. Following this, the target (which was a different picture of the same face/object) and one distracter item appeared, and remained on screen until the participant responded by touching their selected item (e.g. Fig. 1). No feedback was given.

Data analysis

Data from the 13 TD control participants and the additional 19 TD controls that completed the behavioural component clearly indicated that performance on the face- and object-matching tasks was strongly related to age (faces: $r (31) = 0.73$, $P < 0.001$; objects: $r (31) = 0.60$, $P < 0.001$). The consistency and reliability of the tests was also good, with Cronbach’s alpha values $\alpha = 0.73$ for face matching and $\alpha = 0.74$ for object matching. Using the regression equation from the TD children, we generated age-standardised scores for the face- and object-matching tasks (see Brock et al. 2007 for details of implementation).

Eye-tracking task: passive viewing of scenes

Stimuli

Twenty photographs of natural scenes were taken from the first author’s personal collection and

<table>
<thead>
<tr>
<th>Variable</th>
<th>ASD (n = 13)</th>
<th>TD (n = 14)</th>
<th>t value</th>
<th>P (df = 25)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>10.13 (1.89)</td>
<td>10.65 (2.07)</td>
<td>0.69</td>
<td>(P = 0.50)</td>
</tr>
<tr>
<td>Sentence comprehension (TROG) 100</td>
<td>81.38 (19.81)</td>
<td>106.36 (11.63)</td>
<td>4.03</td>
<td>(P &lt; 0.01)</td>
</tr>
<tr>
<td>Non-verbal Ability (WASI matrices)</td>
<td>45.00 (13.74)</td>
<td>56.07 (8.62)</td>
<td>2.53</td>
<td>(P = 0.02)</td>
</tr>
<tr>
<td>SCQ (lifetime)</td>
<td>22.85 (5.84)</td>
<td>3.00 (1.57)</td>
<td>$-12.26$</td>
<td>(P &lt; 0.001)</td>
</tr>
</tbody>
</table>

ASD, autism spectrum disorder; SCQ, Social Communication Questionnaire; TD, typically developing; TROG, Test for Reception Of Grammar-2; WASI, Weschler Abbreviated Scale of Intelligence.
edited in Adobe Photoshop so that each image included people (number of people ranged from one to a small crowd) and objects (e.g. ice-cream van, pelican, distinctive building – see Fig. 2). The background in the scenes was made as simple as possible by editing out any other items or distinctive features. The people and object regions did not overlap, and each occupied approximately one-quarter of the image. Emphasis was placed on making the scenes appear natural. Areas of interest were coded for each of the people, their faces and the objects.

**Procedure**

Participants completed four blocks. Each block comprised a calibration procedure, followed by passive viewing of five scenes. Trials were presented in a fixed randomised order. Between each block participants completed a delayed memory test for faces. These data are reported elsewhere.

A remote Eyelink 1000 (SR-Research) eye-tracking system was used to record the point of gaze of the right eye at a sampling rate of 500 Hz. A 9-point calibration method was used to calibrate

![Figure 1 Matching task procedure; (a) face matching; (b) shoe matching.](image-url)
and validate eye tracking for each participant. This was repeated as many times as necessary at the start of each block until adequate accuracy was achieved. To ensure eye-tracking data were accurate throughout the experiment, a drift correction was conducted before each trial. The fixation point for the drift-correct was positioned such that, when the picture was presented, participants were not fixating on either the people or the objects. Each scene appeared for 10 s, and participants were instructed to simply look at the screen.

Data analysis

Due to technical errors, the data from the final block were lost for three ASD participants. These participants completed 15 trials, and their scores on the eye-tracking variables are calculated accordingly.

If a participant did not fixate on an interest area at all, the maximum time until first fixation (10 000 ms), and the minimum percent of viewing time (0%) was entered for that trial. For each par-
We calculated the mean time to first fixation on the people and objects, and the mean percentage of time that people and objects were viewed. The proportion of times that a participant fixated on people before objects was calculated. We also calculated the ratio of time spent viewing people vs. objects. By including the objects in the scenes and analysing proportional viewing times to people vs. objects, the overall looking times to the images was controlled for. For the faces, median scores were calculated for each participant. Note that ANOVA’s were inappropriate because the interest areas were not independent variables (cf. Klin et al. 2002) therefore we used t-tests to compare the measures of viewing behaviour between the two groups for each interest area.

### Results

#### Behavioural performance

In the combined TD sample (n = 31) there was no effect of sex on the face- and object-matching tasks ($P$'s $> 0.4$).

Three ASD participants performed at chance level on the face-matching task; however, these same participants all performed well above chance level on the equivalent object-matching task (61%, 71% and 72%) therefore we can be confident that they understood instructions and their poor performance was as a result of difficulties matching faces specifically, and have therefore included them in the following analysis. Table 2 provides mean percent accuracy on the tests of face and object matching for the TD and ASD groups, and shows that the TD group outperformed the ASD group on both tasks. When face and object-matching scores were standardised according to age, the ASD participants’ scores were well below the TD children on the face-matching task, $t (24) = 3.91$, $P = 0.001$, and slightly below the TD children on the object-matching task, $t (24) = 2.34$, $P = 0.03$. In both tasks there was a wide range of standardised scores in the ASD group (Fig. 3).

#### Eye movement on the scenes

**People vs. objects**

The mean time to first fixation on the people and the objects are presented in Fig. 4a, and the mean percentage of viewing time on the people and objects are presented in Fig. 4b. As is clear from Fig. 4a, both groups showed a bias for fixating on the people before the objects, although in the TD group this was a highly consistent preference with almost all children fixating very rapidly on the people. Overall, the TD children were far quicker to fixate on the people, $t (24) = 2.96$, $P < 0.01$, and there were no group differences in time taken to fixate on the objects, $P > 0.9$. To control for abso-
lute time taken to fixate on an interest area, we calculated the mean proportion of trials on which people were fixated before the objects. This was 70.67% (SD = 15.57) for the ASD group, and 76.54% (SD = 7.47) for the TD group. These were not significantly different, $t(24) = 1.23$, $P = 0.23$.

To compare the time spent viewing the people and objects between the two groups, we calculated ratios for the mean percentage of viewing time on each region $[(\text{people} - \text{objects})/(\text{people} + \text{objects})]$. Figure 4b shows that the ASD group spent a similar amount of time viewing each region, the ratio being 0.03 (SD = 0.28). By contrast, the TD group spent far more time viewing the people, the ratio being 0.24 (SD = 0.17). The group difference between the ratios was significant, $t(24) = 2.24$, $P = 0.04$. 

Figure 3 ASD and TD children’s age-standardised scores on (a) face-matching task, (b) object-matching task. Bars show means and standard error means. ASD, autism spectrum disorder; TD, typically developing.

Figure 4 (a) Mean time to first fixation on people and objects, (b) Mean percentage of viewing times on the people and objects. Bars show group means and standard error means. ASD, autism spectrum disorder; TD, typically developing.
Neither measure (proportion of times people were fixated before objects; ratio of viewing times on people and objects) were associated with standardised TROG or WASI matrices scores in either the TD or ASD group (all $P’s > 0.3$).

Faces

For the faces of people in scenes, median scores for each participant were calculated because there were several trials in which a participant failed to fixate on a face at all. On average, the time taken for the ASD group to fixate on a face was 3246 ms ($SD = 3319$), whereas the TD group took only 1087 ms ($SD = 609$). This difference was significant, $t_{(25)} = 2.40, P = 0.04$ (Fig. 5a). The average percentage of time viewing the faces was 6.97% ($SD = 5.20$) in the ASD group, and 13.01% ($SD = 6.42$) in the TD group. Again, this difference was significant, $t_{(25)} = 2.67, P = 0.02$ (Fig. 5b).

Associations between viewing behaviour and face and object recognition ability

Our primary aim was to determine associations between visual scanning behaviour and face- and object-recognition ability. For the TD group, correlation analyses revealed no association between standardised performance on the face- and object-matching tasks, and either proportion of times people were looked at before objects, or ratio of looking times to people and objects (all $P’s > 0.3$; Table 3). In contrast, for the ASD group, age-standardised face-matching scores were associated with proportion of trials that people were fixated before objects, $r_{(13)} = 0.66, P = 0.02$ (Fig. 6a), and the association approached significance with the ratio of viewing times on people and objects, $r_{(13)} = 0.45, P = 0.13$ (Table 3). The correlations remained significant after controlling for standardised WASI matrices scores, $r_{(10)} = 0.61, P = 0.04$, the standardised TROG scores, $r_{(10)} = 0.64, P = 0.03$, and SCQ, $r_{(10)} = 0.69, P = 0.01$. Importantly, the associations with age-standardised object recognition ability and either measure of visual scanning were non-significant (Fig. 6b; Table 3).

Percent viewing times on the faces and time to first fixation on the faces, were also correlated with age-standardised matching scores, however all associations were non-significant (all $P’s > 0.3$).
Discussion

The current study examined the hypothetical link between face-recognition impairment and reduced attention to social stimuli in ASD (Schultz 2005). Behavioural tests of face and object matching were conducted, as well as an eye-tracking task including passive viewing of scene photographs. In the eye-tracking component, TD children showed a strong preference for fixating on the people before the objects in scenes, which replicates previous studies documenting an inherent bias to attend to social elements of the surroundings, particularly faces (Palermo & Rhodes 2007; Ro et al. 2007; Langton et al. 2008). The ASD children showed a similar tendency to fixate on people before objects; however, they were slower to fixate on the people than the controls. Over the whole 10-s viewing period, the TD children showed a strong preference for looking at the people rather than the objects. By contrast, the ASD children looked at the people for a similar amount of time as they looked at the objects. In addition, the ASD children spent less time looking at the faces of the people than the TD children.

Table 3 Correlations (r-values) between age-standardised face- and object-matching scores, and viewing behaviour. Time to 1st fixation: proportion $P < O$, proportion of time the people were fixated on before the objects. Viewing time: ratio $P : O$, (people – objects)/(people + objects)

<table>
<thead>
<tr>
<th></th>
<th>ASD ($n = 13$)</th>
<th>TD ($n = 13$)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Time to 1st fixation (proportion $P &lt; O$)</td>
<td>Viewing time (ratio $P : O$)</td>
</tr>
<tr>
<td>Face matching (age-standardised)</td>
<td>0.66</td>
<td>0.45</td>
</tr>
<tr>
<td>$P = 0.02$</td>
<td>$P = 0.13$</td>
<td>$P = 0.75$</td>
</tr>
<tr>
<td>Object matching (age-standardised)</td>
<td>0.26</td>
<td>0.12</td>
</tr>
<tr>
<td>$P = 0.40$</td>
<td>$P = 0.70$</td>
<td>$P = 0.53$</td>
</tr>
</tbody>
</table>

ASD, autism spectrum disorder; TD, typically developing.

Figure 6 Association between proportion of times people were fixated before objects, and (a) age-standardised face-matching score, (regression line shows relationship in ASD group; $r^2 = 0.43$); (b) age-standardised shoe matching score, (all correlations non-significant). ASD, autism spectrum disorder; TD, typically developing.
In sum, we found that the ASD group showed an initial preference for attending to people that was comparable to typical controls, but their overall looking times at people was reduced. These findings are consistent with some previous studies (e.g., Sasson et al., 2007; Riby & Hancock 2008, 2009a,b) but contrast with others. Fletcher-Watson et al. (2009) found that adults with ASD were slower to fixate on ‘person-present’ scenes relative to ‘person-absent’ scenes than typical adults, although throughout a 3-s viewing period, the groups showed a similar allocation of attention, with a bias towards the people-present scene. The different samples might account for some discrepancy – their study included ASD adults with normal intelligence, whereas the current study included ASD children. The different presentation times could also contribute to the contradictory results, because the lack of interest in social stimuli might only have emerged over the longer time periods that our study presented.

The current study also documented significant impairments in face recognition, and moderate impairments in object recognition in the ASD children relative to chronological age. However, our primary aim was to determine whether associations existed between viewing behaviour and recognition ability for faces and objects. Despite finding group differences in overall looking times at people and objects, there were no associations with face or object recognition ability on this measure. However, we found evidence that the initial preference for fixating on people relative to objects was associated with face recognition ability in ASD children. This association could not be accounted for by cognitive ability, as measured by the TROG or WASI matrices, or by the level of autistic symptoms present (according to the SCQ). Moreover, the association was specific to the ability to recognise faces, and did not generalise to objects, suggesting the result was relevant to a component of visual processing that is specialised for face stimuli.

This intriguing result implies that an important factor associated with face-recognition ability in ASD is the way that the individual instinctively allocates their attention, rather than the way attention is sustained over time. The clear implication here is that research with ASD children requires direct associations between behavioural measures to be drawn at the individual level, without assuming that group level deficits in different domains are related. An invalid assumption about an individual, based solely on the average statistics of the group they belong to has been termed the ‘ecological fallacy’ (Robinson 1950). Before considering possible explanations for this result, we note that the lack of a significant association in the TD individuals may well be because of the small variation on both measures in this group. Therefore, we do not suggest that the relationship is necessarily specific to the ASD population, but more likely that the greater variation in performance, where some ASD participants scored well below, but others scored in the same range as the TD controls, led to the significant association emerging in this group. The idea that in the ASD population skill level forms a continuum that continues into the TD population is consistent with previous work researching traits common to ASD (Baron-Cohen et al. 2001).

The correlation between initial viewing preference and face-recognition ability in the ASD group could be interpreted in two ways. First, the instinctive allocation of attention in childhood could be indicative of the way that individuals conducted visual scanning of their environment during infancy. Nelson (2001) proposed that the development of specialised face-processing skills relies crucially on an infant gaining experience with face stimuli during a sensitive period, at around 6–9 months of age (Pascalis et al. 2005). It is thought that a reduced attendance to faces during this time would limit the experience gained, and thus inhibit acquisition of expertise for processing face information. This is in line with our previous study indicating that a reduced advantage for own-race faces is associated with impaired face recognition overall (C.E. Wilson et al. unpublished data). However, this interpretation is tentative since it relies on the assumption that social attention in childhood is highly correlated with social attention in infancy.

A second possibility is that an individual’s proficiency in recognizing faces could dictate the way that they attend to the environment. For example, if a child is poor at processing social information, they may develop a bias for attending to non-social information. If a causal relationship exists between
preferential viewing of social information and face-recognition ability, one approach to ascertaining the direction of causality might be to conduct training studies with children that show deficits. If training a child to attend more to social elements of a scene leads to an improvement in their face-recognition skills, or vice versa, this would imply that one characteristic has a causal affect on the other. Understanding this relationship could also provide insight into disorders such as developmental prosopagnosia, whereby impaired facial recognition is observed in the absence of any apparent deficits in social functioning (Duchaine et al. 2010; Wilson et al. 2010).

Conclusion

To summarise, the current study replicated previous findings suggesting ASD children are impaired at facial identity recognition, and to a lesser extent object recognition, although wide variation existed within the sample. We also confirmed that ASD children exhibit abnormal viewing preferences to components of visual scenes. Moreover, we provided evidence for a relationship between a particular component of viewing behaviour, specifically bias to attend to people before objects, and the ability to recognise facial identity. This work provides the basis for further investigations into how natural viewing preferences may be related to specific abilities to process social and perceptual information in ASD.

Acknowledgements

The authors are extremely grateful to all the participants and their families that took part in this research. We would also like to thank Helen Dodd for her helpful comments on an earlier draft.

References


Accepted 16 September 2010