Recognition of own- and other-race faces in autism spectrum disorders

C. Ellie Wilson a b, Romina Palermo c d, A. Mike Burton e & Jon Brock a c

a Macquarie Centre for Cognitive Science, Macquarie University, Sydney, Australia
b Institute of Psychiatry, King's College, London, UK
c Australian Research Council Centre of Excellence in Cognition and its Disorders, Australia
d Department of Psychology, Australian National University Canberra, Australia
e Department of Psychology, University of Glasgow, Glasgow, UK

Available online: 14 Jul 2011

To cite this article: C. Ellie Wilson, Romina Palermo, A. Mike Burton & Jon Brock (2011): Recognition of own- and other-race faces in autism spectrum disorders, The Quarterly Journal of Experimental Psychology, 64:10, 1939-1954

To link to this article: http://dx.doi.org/10.1080/17470218.2011.603052

FULL TERMS AND CONDITIONS OF USE: http://www.tandfonline.com/page/terms-and-conditions

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or
Recognition of own- and other-race faces in autism spectrum disorders

C. Ellie Wilson1,2, Romina Palermo3,4, A. Mike Burton5, and Jon Brock1,3

1Macquarie Centre for Cognitive Science, Macquarie University, Sydney, Australia
2Institute of Psychiatry, King’s College, London, UK
3Australian Research Council Centre of Excellence in Cognition and its Disorders, Australia
4Department of Psychology, Australian National University Canberra, Australia
5Department of Psychology, University of Glasgow, Glasgow, UK

Empirical data regarding the extent of face recognition abnormalities in autism spectrum disorder (ASD) is inconsistent. Here, 27 ASD and 47 typically developing (TD) children completed an immediate two-alternative forced-choice identity matching task. We contrasted recognition of own- and other-race faces, and, counter to prediction, we found a typical advantage for recognizing own- over other-race faces in both the ASD and TD groups. In addition, ASD and TD groups responded similarly to stimulus manipulations (use of identical or different photographs for identity matching and cropping stimuli to remove hair information). However, age-standardized scores varied widely within the ASD sample, and a subgroup of ASD participants with impaired face recognition did not exhibit a significant own-race recognition advantage. An explanation regarding early experience with faces is considered, and implications for research of individual variation within ASD are discussed.

Keywords: Autism; Face recognition; Other-race effect; Individual differences.

Autism spectrum disorders (ASDs) are currently defined in terms of qualitative social and communicative impairment co-occurring with repetitive behaviours or restricted interests (American Psychiatric Association, 2000). Alongside these core deficits, a commonly reported symptom of ASD is difficulty in facial identity recognition. Numerous studies have demonstrated that individuals with ASD perform worse than matched control participants on a wide range of face-matching tasks (e.g., Boucher & Lewis, 1992; De Gelder, Vroomen, & van der Heide, 1991; Faja, Aylward, Bernier, & Dawson, 2008; Gepner, de Gelder, & de Schonen, 1996; Hauck, Fein, Maltby, Waterhouse, & Feinstein, 1988; Joseph, Ehrman, McNally, & Keehn, 2008; Klin

Correspondence should be addressed to C. Ellie Wilson, Department of Forensic and Neurodevelopmental Sciences, Institute of Psychiatry, King’s College, London, SE5 8AF, UK. E-mail: ellie.wilson@kcl.ac.uk

The authors are grateful to all the participants and staff from Southside Montessori School, Macquarie University Special Education Centre (MUSEC), and the satellite classes of Western Sydney Autism Spectrum Australia (ASPECT). We also thank David White for his assistance with the experiment stimuli. Ellie Wilson was supported by a Macquarie University Research Excellence Scholarship. Jon Brock was supported by an ARC Australian Research Fellowship (DP098466). A.M.B. is now at the University of Aberdeen.

© 2011 The Experimental Psychology Society

et al., 1999; Serra et al., 2003; Tantam, Monaghan, Nicholson, & Stirling, 1989). However, as discussed in a review (Jemel, Mottron, & Dawson, 2006), there have been a number of other studies that have failed to find evidence of significant impairment in facial identity recognition (e.g., Celani, Battacchi, & Arcidiacono, 1999; Deruelle, Rondan, Gepner, & Tardif, 2004; Lahaie et al., 2006; Schultz et al., 2000).

In the current study, we investigated two unresolved issues in relation to face recognition in ASD. First, we considered the possibility that the use of compensatory strategies in some tasks may mask real-life face recognition difficulties, perhaps explaining the inconsistency of results across previous studies. In particular, many studies of ASD involve matching of identical images of faces, allowing participants to complete the task by focusing on idiosyncratic image characteristics rather than the identity of the faces (cf. Burton, Miller, Bruce, Hancock, & Henderson, 2001; Burton, Wilson, Cowan, & Bruce, 1999). If individuals with ASD adopt such a strategy, they should have greater difficulty when required to match different images of the same person. Also, anecdotal reports (e.g., Grandin, 1996) suggest that individuals with ASD may rely excessively on the external features of the face such as the hair. If this is the case, they should have greatest difficulty in experimental tasks where facial stimuli have been cropped to remove external features. The potential for such compensatory strategies was demonstrated by Duchaine and Weidenfield (2003), who found that typical adults could perform within the normal range on standardized tests of face recognition even when stimuli were edited to remove internal feature information.

The second unresolved issue relates to the developmental origins of the face-processing impairment. One prominent hypothesis is that face recognition skills are slow to develop in ASD due to a lack of interest and attention to faces (Chawarska & Volkmar, 2007; Dawson et al., 2002). Studies suggest that ASD infants and young children orient towards faces less than their typically developing peers (Kikuchi, Senju, Tojo, Osanai, & Hasegawa, 2009; Osterling, Dawson, & Munson, 2002), and they look at faces for less time than typically developing children do (Osterling & Dawson, 1994; Swettenham et al., 1998). Our own research has shown that a tendency to look at faces before looking at objects in visual scenes is associated with level of face recognition skill (Wilson, Brock, & Palermo, 2010). There is considerable evidence from typical development that experience influences the acquisition of face recognition skills (McKone, Crookes, & Kanwisher, 2009). For example, six-month-old infants are able to reliably discriminate between monkey faces but older children lose this ability as their experience with human faces increases (Pascalis, de Haan, & Nelson, 2002). Likewise, infant monkeys are able to discriminate between individual monkey faces as well as between individual human faces, but after one month of exposure to either human or monkey faces they lose the ability to discriminate between individuals of the other species whom they did not experience during the month (Sugita, 2008). Chawarska and Volkmar (2007) hypothesized that this “other-species” effect should be reduced in infants with ASD due to a lack of attention to faces. Unfortunately, the infants in their study failed to demonstrate discrimination for either species, thus results could neither support nor contradict their hypothesis.

A related and potentially more sensitive phenomenon is the “other race” effect, whereby children and adults tend to be better at differentiating between unfamiliar faces that are members of their own racial group than they are with faces of another racial group (Ellis, Deregowski, & Shepherd, 1975; Meissner & Brigham, 2001; Walker & Tanaka, 2003). Importantly, the magnitude of this effect has been shown to vary in relation to an individual’s degree of social contact with the other race (Walker, Silvert, Hewstone, & Nobre, 2008). Studies with children indicate that the effect size increases with age between 3 and 5 years (Sangrigoli & de Schonen, 2004), but there are no further changes in effect size between 6 and 14 years (de Heering, de Liedekerke, Deboni, & Rossion, 2010; see Crookes & McKone, 2009). Given the above, we hypothesized that the magnitude of the other-race effect would be
reduced in ASD due to a lack of attention to own-race faces. To our knowledge, this is the first investigation of the other-race effect in ASD individuals.

In the current study, therefore, we tested typically developing and ASD children on a simple sequential face-matching task while making three orthogonal manipulations to the stimuli. First, stimuli were either own- or other-race faces. Second, the target images were to be matched either to identical or to different images of the same person. Third, stimuli were either whole faces or were cropped to remove external features including hair.

Our primary interest was in children’s relative performance across the different conditions. We therefore tested a relatively large sample of typically developing children and then selected from this group children who were individually matched to the ASD children in terms of their overall matching performance. This ensured that any group differences in the effects of race, image type, or cropping could not be explained in terms of differences in overall performance (cf. Strauss, 2001).

In addition to looking at effects at the group level, we also considered individual variation within ASD (Brock, 2011). Specifically, we utilized the full set of data from typically developing children to create age-standardized scores with which to index the severity of the face-matching impairment demonstrated by each child with ASD. We then considered various potential correlates of face recognition performance and the extent to which the effects of our stimulus manipulations might differ as a function of the severity of the child’s face recognition impairment.

Method

Participants

Thirty-one children were recruited from Autism Spectrum Australia (ASPECT) and Macquarie University Special Education Centre (MUSEC). All met criteria for ASD (25 autism, 6 Asperger’s) according to the DSM–IV–TR (Diagnostic and Statistical Manual of Mental Disorders–Fourth Edition, Text Revision; American Psychiatric Association, 2000) and had previously been given a formal diagnosis by a qualified clinician, using tools including the Autism Diagnostic Interview Revised (ADI–R; Lord, Rutter, & Le Couteur, 1994), the Autism Diagnostic Observation Schedule (ADOS; Lord, Rutter, DiLavore, & Risi, 2001) or the Childhood Autism Rating Scale (CARS; Schopler, Reichler, DeVellis, & Daly, 1980). In addition, all scored above the cut-off for ASD on the Social Communication Questionnaire (lifetime; Rutter, Bailey, & Lord, 2003), a parental questionnaire based on the Autism Diagnostic Interview, with which it has good agreement (e.g., Berument, Rutter, Lord, Pickles, & Bailey, 1999; Bishop & Norbury, 2002). Four participants were excluded from the analysis because they were unable to follow instruction and complete the face recognition task, leaving a final sample size of 27 (21 males). A total of 20 ASD subjects were tested at school, and the remaining 7 were tested in their homes. Four ASD participants were of South-East Asian origin (the Philippines and China) but had lived in Australia since birth. The remaining 23 were Caucasian.

Forty-seven typically developing (TD) children (22 males) were recruited as a comparison group to cover the range of age and face-recognition performance of the ASD group. Three were tested at Macquarie University, four were tested at home, and the remaining 40 were tested at school. Five TD participants were of South-East Asian origins (Korea, Japan) but had lived in Australia since birth, and the remaining 42 were Caucasian. Twenty-seven of these children were selected to form a matched control group on the basis of their overall performance on the sequential face-matching task described below.

Group characteristics are shown in Table 1. Receptive language skills were assessed using the Test for Reception of Grammar–2 (TROG; Bishop, 2003). Nonverbal ability was assessed via the matrices subscale of the Wechsler Abbreviated Scale of Intelligence (WASI matrices; Wechsler, 1999). We also administered the immediate recognition of faces subset of the Children’s Memory Scale (Cohen, 1997). In this test, participants are shown a set of faces sequentially for 2 seconds each and are then required to make old/new judgements...
on a larger set of images, half of which are identical images of faces already viewed. Scores of typically developing children confirmed age-appropriate face recognition skills. However, the majority of participants with ASD failed to complete the task or scored at chance level, indicating a failure to comprehend or comply with task demands.

Throughout this paper, we refer to Caucasian faces as “own-race” and Egyptian faces as “other-race”. Although some of the participants were of Asian ethnicity, all participants had spent their whole lives living in Australia, and we can be confident that their exposure to Caucasian faces was much greater than their exposure to Egyptian faces. According to the Australian Bureau of Statistics 2006 census, less than 0.2% of the Australian population claim Egyptian ancestry, compared with 93.2% who claim European ancestry. Nevertheless, all analyses were repeated without the inclusion of participants of Asian ethnicity. This had no effect on the pattern of results, so all participants are included in the reported analyses.

Table 1. Participant group characteristics

<table>
<thead>
<tr>
<th>Variable</th>
<th>ASD (n = 27)</th>
<th>TD (n = 47)</th>
<th>Ability-matched TD (n = 27)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>Range</td>
</tr>
<tr>
<td>Age (years)</td>
<td>10.07</td>
<td>2.05</td>
<td>6.83–15.58</td>
</tr>
<tr>
<td>Overall face-matching performance</td>
<td>72.28</td>
<td>6.74</td>
<td>57.08–86.67</td>
</tr>
<tr>
<td>Reception of grammar (TROG)</td>
<td>74.63</td>
<td>21.97</td>
<td>55–118</td>
</tr>
<tr>
<td>Nonverbal ability (WASI)(a)</td>
<td>42.82</td>
<td>13.13</td>
<td>20–64</td>
</tr>
<tr>
<td>SCQ (lifetime)(b)</td>
<td>23.90</td>
<td>5.44</td>
<td>15–35</td>
</tr>
</tbody>
</table>

Note: ASD = autism spectrum disorder. TD = typically developing. TROG = Test for Reception of Grammar. WASI = Wechsler Abbreviated Scale of Intelligence. SCQ = Social Communication Questionnaire.

\(a\)Seven TD participants were too young to complete the WASI, therefore TD n = 38. \(b\)Due to questionnaires not being returned to the experimenter, the total number of SCQ (lifetime) was 29 for the TD participants’ (18 in ability-matched TD).

Stimuli
One hundred and twenty young male adults (60 Caucasian, 60 Egyptian) were selected from a database of unfamiliar faces compiled by the Glasgow Face Recognition Group (www.psy.gla.ac.uk/~mike/facerec.html; see Burton, White, & McNeill, 2010). All had short, neat hairstyles and no piercings, facial hair, or spectacles. Faces displayed neutral expressions and were facing the camera. All images were converted to greyscale. Two photographs of each individual were used, which had been taken with a different camera under different lighting conditions, thus creating two sets of images of the Caucasian individuals and two sets of the Egyptian individuals. Original photographs included the whole head. Cropped stimuli were created by putting an oval shape around each image, thereby removing most hair information. Identities were organized into pairs based on similar characteristics such as eyebrow shape and size of eyes and mouth. Members of a pair were of the same race.

A previous study tested Caucasian and Egyptian adults on an identity-matching task, using faces from this stimulus set, and found no main effect of race of stimuli, demonstrating that the Caucasian and Egyptian faces were of comparable levels of difficulty (Megreya, White, & Burton, 2011). Furthermore, Megreya et al. found a crossover interaction pattern, in which the two participant groups were superior at recognizing faces of their own race and significantly worse at recognizing other-race faces.

Design
A sequential two-alternative forced-choice (2-AFC) matching task was used. There were three
orthogonal conditions, each with two levels. The first condition was race; faces were either “own race” (e.g., Figures 1a and 1b) or “other race” (e.g., Figure 1c). The second condition was image type. In the “identical” condition, the target was matched to an identical image (e.g., Figure 1a). In the “different” condition, the target was matched to a different image of the same person (e.g., Figures 1b and 1c). The third condition was cropping, whereby all three images seen in a trial showed either the whole face (e.g., Figures 1a and 1b) or the internal face region only (e.g., Figure 1c). The same face pairs were used in every condition, but the face used as the target differed and was counterbalanced across conditions.

Protocol
The task was presented on a 32 × 28-cm touch screen monitor using SuperLab Pro, Version 4 (Abboud & Sugar, 1997). To start each trial, a 2 × 2-cm shape (blue circle, green square, red triangle, or yellow star) appeared in the centre of the screen (see Figure 2). The participant touched the shape, and, upon release, a test face appeared in the centre of the screen. This ensured that the participant was attending to the screen when the test face was presented. The test face was shown for 500 ms. Pilot testing revealed that a 500-ms presentation resulted in performance that was neither at floor nor at ceiling level in children with special educational needs (non-ASD students at MUSEC). The target and distractor items appeared immediately after the test face, and the participant was instructed to “touch the same person” as the test face. The target and distractor remained on the screen until a response was made, and no time limit was enforced. No feedback was given to indicate whether they were correct, but participants were verbally encouraged and praised by the experimenter throughout the task. As soon as a response was made, the shape to initiate the next trial appeared on screen.

In order to prevent participants developing a routine of repeatedly touching the same place on the screen to respond, the pairs of faces were presented 100 pixels above or below the test face. The location and identity of the target face were fully counterbalanced.

Practice session
Prior to testing, participants completed practice trials with pictures of animals (two trials) and cartoon faces (two trials) before moving on to complete practice with photographed faces (eight trials). Feedback was provided if the participant responded incorrectly, and the trial was run again. If the participant did not understand the aim of the task, or appeared to be guessing, they were
shown the test, target, and distractor faces at the same time on a piece of paper and were asked to select the “same face”.

Test session
Each participant completed two testing sessions on consecutive days. Participants were tested with whole faces in the first session and cropped faces in the second. Within each session, participants completed one block of testing in the identical faces condition followed by a block in the different faces condition. Each block took approximately 12 min to complete and included 30 trials with own-race faces, followed by 30 trials with other-race faces. Within each race condition, trials were presented in the same pseudorandom order for each participant. The TROG was completed midway through Session 1 and the WASI matrices midway through Session 2.

Results
As mentioned in the introduction, the scores for each participant were collapsed across tasks, and each individual ASD participant was matched to a TD participant on overall face-matching performance. Each ability-matched pair had scores that were within two percentage points of each other. The ability-matched TD subgroup was used for the first set of analyses (see Table 1 for the two groups). Thus, differential performance across the conditions could be examined, without being affected by differences in overall performance level. The groups were mismatched on age, receptive grammar, and nonverbal reasoning. The relationship between these measures and individual differences in face recognition ability are addressed in later analyses.

Effects of race, image type, and cropping on face recognition ability
The first analysis compared performance of the children with ASD with their ability-matched control group. Data were entered into an analysis of variance (ANOVA) with three within-participant factors—race (own/other), image type (identical/different), and cropping (whole/internal)—and group (ASD/TD) as the between-subjects factor (Figures 3a–3c). There was a highly significant main effect of race, with own-race faces being matched more successfully than other-race faces,
F(1, 52) = 42.24, p < .001; Caucasian participants only, F(1, 44) = 42.05, p < .001. The absence of a race by group interaction suggests similar effects of race in both groups, F(1, 52) = 1.12, p = .30.

A significant main effect was also found for image type, F(1, 52) = 238.21, p < .001. Again, there was no interaction with group, (F < 1), indicating that both groups found the identical images a great deal easier to match than the different images. Surprisingly, a significant main effect in the cropping condition revealed that participants performed significantly better on the internal images than on the whole images, F(1, 52) = 5.23, p = .03. Again, we found no interaction with group (F < 1).

The only significant interaction was between image type and cropping, F(1, 52) = 12.24, p = .001, driven by the fact that the internal faces were matched significantly better than the whole faces in the identical condition, t(53) = -3.46, p = .001, but there was no effect of cropping in the different condition, t(53) = 0.78, p = .44. Importantly, however, there was no three-way interaction with group.

Effect of age on the race effect
It was important to ascertain whether the magnitude of the other-race effect varied with age in TD children in this study, as in order to match groups on overall performance, they were necessarily mismatched on age. To determine the effect of age in our TD sample, we conducted a repeated measures ANOVA on data from the whole sample of TD children (n = 47), with race as the within-subjects factor and age as a covariate. Results showed main effects of race, F(1, 45) = 6.15, p = .02, and age, F(1, 45) = 83.15, p < .001; Caucasian participants only, F(1, 38) = 5.95, p = .03. Importantly, the interaction between race and age was nonsignificant, F(1, 45) = 1.56, p = .22, indicating that performance improved with age overall, but that the advantage for own-over other-race faces was fairly constant between the ages of 4 and 16 years (in line with the review by McKone et al., 2009).

Individual differences in overall performance
We also aimed to examine variation in face recognition ability at the individual level. Figure 4a shows performance (mean score collapsed across conditions) of each participant plotted as a function of chronological age. A regression analysis showed that age accounted for a large proportion (62%) of variance in performance of the TD group. Using this regression model, we calculated

\[ F(1, 52) = 42.24, \ p < .001; \ \text{Caucasian participants only, } F(1, 44) = 42.05, \ p < .001. \]

\[ \text{The absence of a race by group interaction suggests similar effects of race in both groups, } F(1, 52) = 1.12, \ p = .30. \]

\[ \text{A significant main effect was also found for image type, } F(1, 52) = 238.21, \ p < .001. \]

\[ \text{Again, there was no interaction with group, } (F < 1), \text{ indicating that both groups found the identical images a great deal easier to match than the different images.} \]

\[ \text{Surprisingly, a significant main effect in the cropping condition revealed that participants performed significantly better on the internal images than on the whole images, } F(1, 52) = 5.23, \ p = .03. \]

\[ \text{Again, we found no interaction with group } (F < 1). \]

\[ \text{The only significant interaction was between image type and cropping, } F(1, 52) = 12.24, \ p = .001, \text{ driven by the fact that the internal faces were matched significantly better than the whole faces in the identical condition, } t(53) = -3.46, \ p = .001, \text{ but there was no effect of cropping in the different condition, } t(53) = 0.78, \ p = .44. \]

\[ \text{Importantly, however, there was no three-way interaction with group.} \]

\[ \text{Effect of age on the race effect} \]

\[ \text{It was important to ascertain whether the magnitude of the other-race effect varied with age in TD children in this study, as in order to match groups on overall performance, they were necessarily mismatched on age. To determine the effect of age in our TD sample, we conducted a repeated measures ANOVA on data from the whole sample of TD children } (n = 47), \text{ with race as the within-subjects factor and age as a covariate. Results showed main effects of race, } F(1, 45) = 6.15, \ p = .02, \text{ and age, } F(1, 45) = 83.15, \ p < .001; \text{ Caucasian participants only, } F(1, 38) = 5.95, \ p = .03. \text{ Importantly, the interaction between race and age was nonsignificant, } F(1, 45) = 1.56, \ p = .22, \text{ indicating that performance improved with age overall, but that the advantage for own-over other-race faces was fairly constant between the ages of 4 and 16 years (in line with the review by McKone et al., 2009).} \]

\[ \text{Individual differences in overall performance} \]

\[ \text{We also aimed to examine variation in face recognition ability at the individual level. Figure 4a shows performance (mean score collapsed across conditions) of each participant plotted as a function of chronological age. A regression analysis showed that age accounted for a large proportion (62%) of variance in performance of the TD group. Using this regression model, we calculated} \]
age-standardized scores for each participant. This was achieved by using each subject’s age to calculate their predicted score, then subtracting the predicted score from their actual score and dividing the difference by the standard error of the regression model (see Brock, Jarrold, Farran, Laws, & Riby, 2007, for details of implementation). The standardized scores are shown in Figure 4b. Scores above 0 indicate that the participant performed better than predicted for their age, and scores below 0 indicate their performance was below the predicted level. A score of less than −1.64 is considered to represent significant impairment (cf. Crawford & Garthwaite, 2002). The mean age-standardized score of the ASD group was −1.72 (SD = 1.50) and the mean score of the ability matched TD group was −0.32 (SD = 0.18). A one-sample t test showed that, on average, the ASD group performed significantly below age-appropriate levels, t(26) = −5.99, p < .001. However, there was a great deal of variation within the ASD group. While some participants were severely impaired, others performed well within the range of their TD peers.

To explore this individual variation further, we examined the associations between age-standardized scores on the face recognition test and age-standardized measures of verbal and nonverbal ability based on the published norms for the TROG and WASI matrices, respectively. Within the TD group, there was no association between face recognition and either verbal, r(45) = −.07, p = .66, or nonverbal ability, r(38) = .05, p = .78. Within the ASD group, there was no association between face recognition ability and verbal abilities, r(27) = .31, p = .51, but a correlation that approached significance with nonverbal ability, r(27) = .34, p = .08 (Figure 5). Nevertheless, as Figure 5 shows, there were several participants with good nonverbal ability, but impaired face recognition ability. The reverse was also true in two cases.
Finally, a correlation analysis between Social Communication Questionnaire (SCQ, lifetime) score and age-standardized face recognition scores was also nonsignificant in the ASD group, \( r(25) = -0.07, p = .76 \).

**The race effect in individuals with age-appropriate and impaired face recognition ability**

Contrary to prediction, the preceding analyses showed that ASD children, on average, demonstrate an apparently normal other-race effect. However, they also reveal considerable heterogeneity within the ASD population in terms of their overall face-matching skills. Some children have age-appropriate face recognition abilities, but others demonstrate severe impairments. We therefore reanalysed our data for these two subgroups separately, comparing each with their own ability-matched control groups.

Figure 6a shows the race effect in children with ASD who have age-appropriate levels of overall performance, together with data from their ability-matched typically developing controls. An ANOVA with race as a within-subjects factor and group as a between-subjects factor revealed a main effect of race in the expected direction, \( F(1, 22) = 38.59, p < .01 \). The effect of group, \( F(1, 22) = 0.22, p = .64 \), and the interaction between race and group, \( F(1, 22) = 1.87, p = .19 \), were nonsignificant.

By contrast, Figure 6b suggests that the race effect may in fact be reduced in the subset of children with ASD who have impaired face recognition. An effect of race persisted, \( F(1, 28) = 18.10, p < .01 \), but there was an interaction between group and race that narrowly failed to reach significance, \( F(1, 28) = 3.97, p = .06 \). Specifically, while the control group showed a significant advantage for own-race faces, \( t(14) = 4.92, p < .001 \), this effect was absent in the ASD subgroup, \( t(14) = 1.46, p = .17 \).

Similar analyses considering the effect of image type and cropping showed no group by condition interactions for either subgroup.
Discussion

Aberrant social interaction is a defining feature of ASD and is commonly manifested in young children as a reduced interest in social stimuli such as faces (Osterling & Dawson, 1994; Osterling et al., 2002; Swettenham et al., 1998). It has been proposed that reduced attention to faces early in life could impair the development of face recognition skills in ASD (Dawson et al., 2002). In the current study, we investigated the other-race effect in ASD, with the view that reduced attention to faces during infancy could inhibit the development of a face-processing system that is specialized to process the type of face that is most salient in the environment—that is, own-race faces. However, results of previous studies investigating own-race face recognition in ASD have been extremely inconsistent, with some studies reporting significant impairment among individuals with ASD but others finding no group differences. Therefore, we also took the opportunity to investigate the effect of various stimulus manipulations on face-matching abilities of children with ASD, and individual differences in skill level, which might explain the contradictory results of previous studies.

We suspected that discrepant findings on own-race face recognition in previous research might have been a function of different stimuli used in different studies. We hypothesized that individuals with ASD may be able to successfully complete some face-matching tasks by relying on superficial image characteristics or by using information from the hair rather than the internal features. However, we found no evidence for such compensatory strategies. Relative to control children matched on overall performance, ASD children were no better at matching identical images than they were at matching different images of the same face, nor were they adversely affected by cropping of the stimuli to remove the hair. We did find an unexpected interaction between these
two factors, such that the effect of cropping was greater for identical than for different images. The origin of the interaction is unclear, although all participants completed the conditions in the same fixed order, so it could simply reflect practice effects. Crucially, there was no three-way interaction with group, indicating that whatever caused the interaction between cropping and image type, it did not affect the ASD group any more or less than the comparison group.

Given that none of the within-subject factors interacted with group membership, we were able to collapse scores across conditions and consider individual variation in overall performance. Within the typically developing group, age was a strong predictor of performance, so we used the scores of the typically developing children to derive age-standardized scores for all participants. Overall, the ASD group performed significantly below the level predicted for their age. However, the variation in performance level within the ASD group was far greater than that found in the typically developing group, ranging from well within the normal range to severely impaired. This variation in face recognition skills is often overlooked in ASD research (although see Barton et al., 2004; and Rutherford, Clements, & Sekuler, 2007) but could potentially account for much of the inconsistency of findings. Specifically, if researchers only take note of group mean performance, the outcome of a study will be highly sensitive to participant sampling (i.e., whether the majority of ASD participants happen to fall within or below the normal ability range for face recognition).

Given that ASD is a heterogeneous disorder with wide variation in terms of the severity and type of symptoms and cognitive impairments, it is no surprise that there should be heterogeneity of face recognition skills. However, there was no association between standardized face-matching scores and severity of symptoms in our ASD group, as indexed by scores on the SCQ. We also found no significant association between standardized face-matching scores and either reception of grammar or nonverbal ability in either group. While there was a nonsignificant positive association between face recognition and nonverbal ability in the ASD group, there were numerous individuals who exhibited either normal nonverbal ability and impaired face recognition or the opposite pattern of results. The relative independence of face recognition from measures of general cognitive ability in the current study is consistent with previous studies of typically developing children and adults (Bowles et al., 2009; Jeffery & Anderson, 2004; McKone & Palermo, 2010; Wilmer et al., 2010) and suggests that variation in intellectual ability does not provide a complete explanation for heterogeneity of face recognition in ASD.

Instead, the clearest difference between ASD individuals with and without face-matching difficulties was their susceptibility to the other-race effect. When we looked at the ASD group as a whole, we found no significant difference in the magnitude of the race effect relative to ability-matched controls. However, when we divided the ASD group based on their age-standardized face-matching scores, we found that children with age-appropriate performance showed a normal advantage for own-race faces, whereas those with impaired face matching showed no significant race effect. Caution is warranted here because the impaired subgroup did show a trend in the direction of an own-race advantage, and, when these individuals were compared to their individually matched control participants, the race by group interaction narrowly failed to reach significance. Thus, before considering the implications of these results, we stress that conclusions drawn here are necessarily tentative.

One obvious concern is that the absence of a race effect in the “impaired” subgroup is merely a reflection of floor effects. However, the presence of an own-race advantage in the control group negates this possibility. Another potential concern is that face-processing impairments may simply reflect a lack of attention to the stimuli during the task. Again, this seems unlikely. In a recent eye-tracking study using similar stimuli (Wilson, Brock, & Palermo, 2011), we found no difference in gaze directed at faces between individuals with and without ASD and no association between total
looking time and age-standardized performance. While we did not track eye movements during the current study, the task was designed such that participants had to attend to the region of the display that would show the face in order to initiate each trial. Moreover, an attention deficit would predict an overall decrement in performance but not a qualitative difference as reflected in the changing sensitivity to the other-race effect.

Instead, we consider the role of early developmental changes to face-processing systems. Face perception mechanisms are thought to be “experience expectant”, requiring appropriate visual input early in life to develop to maturity (Morton & Johnson, 1991). A study by Anzures, Quinn, Pascalis, Slater, and Lee (2010) demonstrated a shift in face-processing skills between 6 to 9 months. Both 6- and 9-month-olds could differentiate between faces of their own race but not faces of another race. However, only the 9-month-olds could categorize faces according to race, suggesting that the way the facial information is processed changes between these ages. This is consistent with Nelson’s (2001) proposal that an infant’s face-processing mechanisms undergo a “sensitive period”, thought to be between 6 and 9 months (Kelly et al., 2007). Experience of face stimuli in this period leads to an increasingly refined face prototype through a process of “perceptual narrowing”, whereby face-processing mechanisms are refined to optimally process faces that are similar to those they have frequently experienced. In addition, the capacity to discriminate between faces that are atypical of their experience is reduced (Kelly, Liu, Lee, Quinn, Pascalis, & Slater, 2009; Kelly et al., 2007; Pascalis et al., 2005). Whilst our results demand replication, they are certainly consistent with the hypothesis that reduced attention to faces in infancy can disrupt the perceptual narrowing of face-processing mechanisms, leading to abnormal and impaired face recognition skills later in life.

Assuming these results do hold up to replication, an intriguing question for future research is whether individual differences in perceptual narrowing in face processing are related to perceptual narrowing in other domains. In particular, perceptual narrowing has been well documented in the typical development of language. All infants are born with the ability to discriminate the phonetic components of any language, but they go on to develop the skills necessary to discriminate and reproduce only the language to which they are frequently exposed (Kuhl, Tsao, & Liu, 2003). In a recent study, Constantino et al. (2007) investigated this effect in ASD. Although these authors failed to demonstrate the predicted autistic advantage for discriminating non-native phonemes, their participants were all high-functioning individuals with a mean full scale IQ of 110, suggesting that few, if any, had clinically significant language difficulties.

By analogy with the current face-processing data, we would predict that reduced perceptual narrowing for language stimuli would most likely be observed amongst autistic individuals demonstrating severe language impairments. Such an effect has been reported in a study of infants born prematurely who, unlike control infants, retained the ability to discriminate non-native phonemes beyond 6–12 months of age, (Jansson-Verkasalo et al., 2010). At a two-year follow-up, the premature infant group exhibited significant language delay.

Conclusions

Here, we present evidence suggesting that some ASD individuals have age-appropriate face-matching skills and evidence a normal other-race effect. However, some ASD children have severely impaired facial identity recognition skills overall and do not appear to exhibit the typical advantage for recognizing own- over other-race faces. ASD is a pervasive developmental disorder, and cognitive and perceptual systems that undergo significant changes involving brain-environment interactions could become derailed by atypical social experience, with lifelong consequences. Specifically, we propose that a reduction in perceptual narrowing during the development of the face-processing system could underlie impaired face recognition.

While our conclusions are tentative, and our results demand replication, they clearly highlight the importance of considering the heterogeneity that exists within ASD, not only in terms of face
processing, but in other domains as well (Ingram, Takahashi, & Miles, 2007; Prior et al., 1998; Rapin, Dunn, Allen, Stevens, & Fein, 2009). While researchers readily acknowledge the fact that ASD is a heterogeneous disorder, most published reports only consider group averages (Brock, 2011). By taking an individual differences approach, it may be possible to reconcile contradictory findings, to identify meaningful subgroups within ASD, and to begin to link together findings at different levels of explanation in order to generate causal developmental models of the behavioural symptoms associated with ASD.

Original manuscript received 13 January 2011
Accepted revision received 9 June 2011
First published online 7 September 2011

REFERENCES


THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY, 2011, 64 (10)


