Elevated amygdala response to faces and gaze aversion in autism spectrum disorder

Nim Tottenham, Margaret E. Hertzig, Kristen Gillespie-Lynch, Tara Gilhooly, Alexander J. Millner, and B.J. Casey

Autism spectrum disorders (ASD) are often associated with impairments in judgment of facial expressions. This impairment is often accompanied by diminished eye contact and atypical amygdala responses to face stimuli. The current study used a within-subjects design to examine the effects of natural viewing and an experimental eye-gaze manipulation on amygdala responses to faces. Individuals with ASD showed less gaze toward the eye region of faces relative to a control group. Among individuals with ASD, reduced eye gaze was associated with higher threat ratings of neutral faces. Amygdala signal was elevated in the ASD group relative to controls. This elevated response was further potentiated by experimentally manipulating gaze to the eye region. Potentiation by the gaze manipulation was largest for those individuals who exhibited the least amount of naturally occurring gaze toward the eye region and was associated with their subjective threat ratings. Effects were largest for neutral faces, highlighting the importance of examining neutral faces in the pathophysiology of autism and questioning their use as control stimuli with this population. Overall, our findings provide support for the notion that gaze direction modulates affective response to faces in ASD.

Keywords: fMRI; amygdala; face expressions; autism spectrum disorders; eye-tracking

At the heart of autism spectrum disorders (ASD) are deficits in social interactions (DSM-IV, 1994). Human social interactions can be ambiguous and unpredictable, but social information from facial expressions is a readily available cue that can reduce the uncertainty intrinsic to social exchanges. Unfortunately, many individuals with ASD exhibit impaired judgments of facial expressions (Bormann-Kischkel et al., 1995; Klin et al., 1999; Adolphs et al., 2001; Dawson et al., 2005; Kuusikko et al., 2009), which can significantly interfere with successful social interactions. Investigations into the neurobiology of face emotion processing of ASD have focused on the amygdala, a subcortical structure involved in detecting and learning about the motivational value of social interactions. Investigations into the neurobiology of face emotion processing in ASDs (Bookheimer, 2001) and functional abnormalities of the amygdala have been demonstrated among individuals with ASD, including both elevated (Dalton et al., 2005; Monk et al., 2010; Weng et al., 2011) and reduced amygdala reactivity (Ashwin et al., 2007; Hadijkhani et al., 2007; Bookheimer et al., 2008; Corbett et al., 2009) in response to faces. Findings such as these have led to the hypothesis that aberrations of the amygdala contribute to behavioral anomalies in face expression processing in ASDs (Baron-Cohen et al., 2000), although a full understanding of amygdala dysfunction in ASD continues to be actively pursued.

The nature of the amygdala response to faces in ASD depends in part on stimulus characteristics; for example, atypical amygdala responses among individuals with ASD may be more apparent in response to unfamiliar or dynamic faces relative to familiar or static stimuli (Pierce et al., 2004; Pelphrey et al., 2007; Pierce and Redcay, 2008). The valence of emotional expressions may also influence amygdala responses; individuals with ASD have exhibited both increased and decreased amygdala activity relative to controls in studies using positive, negative and neutral expressions (Dalton et al., 2005; Corbett et al., 2009; Monk et al., 2010; Weng et al., 2011). Decreases have been observed in response to fear (Ashwin et al., 2007; Kleinhans et al., 2011; Perlman et al., 2011). Neutral faces have resulted in both increases and decreases in the amygdala response of individuals with ASD (Hadijkhani et al., 2007; Bookheimer et al., 2008) along with diminished amygdala habituation over the scan (Kleinhans et al., 2009).

Group differences in eye contact is another influential variable in amygdala reactivity during face processing (Dalton et al., 2005). Typically developing (TD) individuals often focus most on the eye region when processing faces (Schwarzer et al., 2005), which is the most efficient region for understanding facial emotion (Baron-Cohen et al., 1997; Morris et al., 2002). However, decreased eye contact is commonly observed in individuals with ASD (Osterling and Dawson, 1994; Baron-Cohen et al., 1997; Klin et al., 2002; Dalton et al., 2005), and may contribute to the variation in functional magnetic resonance imaging (fMRI) findings across studies. It has been posited that decreased eye contact reflects social motivation impairments (Carver and Dawson, 2002; Schultz, 2005). An alternative view suggests that decreased eye contact is a means of attenuating overarousal associated with face-to-face contact. In support of this view, Dalton et al. (2005) have shown that decreased eye contact was associated with diminished amygdala response to faces in ASD. Close examination of gaze patterns has shown that ASD gaze is characterized by more eye movements away from rather than fewer eye movements toward the eyes, suggesting that decreased eye contact is an active avoidance of the eye region (Klemann et al., 2010). Additionally, Klemann et al. (2012) showed that manipulating gaze to the eye region resulted in increased amygdala response relative to gaze directed at the mouth region.
The current fMRI study used a within-subject design to manipulate gaze for the purposes of examining amygdala responses under natural viewing conditions as well as under conditions of increased eye gaze via experimental gaze manipulation. Moreover, we assessed subjective interpretations of these faces to examine whether amygdala responses were associated with affective evaluations of expressions. Given the complexities of the amygdala findings in ASD, the current report focused on amygdala response to two expressions. Because of the amygdala’s well-established role in responding to social threat (reviewed in Davis and Whalen, 2001), we chose a threatening face (angry), which has been shown to elicit a strong amygdala signal in many studies (Hariri et al., 2000; Yang et al., 2002; although see meta-analyses in Phan et al., 2002 and Fusar-Poli et al., 2009). Other expressions, such as fear, tend to show a more robust amygdala signal, but we selected angry faces in part because it is a direct signal of threat to the perceiver (Whalen, 1998 and Strauss et al., 2005). The second face type was what the scientific field typically refers to as ‘neutral’, which are created to be void of clear emotional valence.

**METHOD**

**Participants**

We recruited 94 participants (TD = 60; ASD = 34) (Table 1). Eighty-one participants (TD = 51; ASD = 30) provided complete or partial behavioral data (degrees of freedom provided for each analysis). We were able to obtain usable eye-movement data from 65 participants (39 TD; 26 ASD), and 76 participants (45 TD; 31 ASD) felt comfortable enough to participate in the MRI scanning session. Six participants (3 TD, 3 ASD) were excluded from the fMRI analyses owing to excessive head motion (>2.5 mm or 2.5° of rotation), leaving a total of 70 participants with usable fMRI data (42 TD, 28 ASD). Participants were recruited through clinical referral or advertisements. TD participants were free of psychiatric/neurological impairment as per telephone screening, where participants (or their parents) were asked to indicate whether the participant had been previously diagnosed with any psychiatric/neurological illness or behavioral/learning difficulties, whether the participant had ever taken any psychotropic medications, whether there was any first-degree relative family history of mental illness, as well as standard MRI contraindications. Participants in the ASD group had previously received a clinical diagnosis of an ASD from clinicians independent of this study, where the majority of the diagnoses was Asperger Syndrome (63%), and the remaining diagnoses were pervasive development disorder-not otherwise specified (PDD-NOS) (26%) and Autism (11%), which was confirmed with the Autism Diagnostic Observational Schedule, Generic (Lord et al., 2000) when possible (n = 14) by Dr Hertzig (co-author on this manuscript), a Child and Adolescent Psychiatrist at Weill Cornell Medical College/New York-Presbyterian Hospital with clinical and research expertise in the development of ASDs. Although we attempted to confirm this diagnosis within our own laboratory with the Autism Diagnostic Observational Schedule, Generic (Lord et al., 2000), this was not always possible (e.g. due to scheduling challenges). To further quantify autistic traits, participants or their parents completed the Autism Spectrum Quotient (AQ) (Baron-Cohen et al., 2001), which assesses social skills, attention switching, attention to detail, communication and imagination. Although the AQ score (ASD mean = 34, TD mean = 17; P < 10⁻⁷) is not diagnostic, this measure is useful support for diagnosis because it has been validated in a clinical sample (Woodbury-Smith et al., 2005), showing a cutoff point of 26 for high-functioning autism.

Thus, our approach to using the AQ to increase the confidence of the original diagnosis of participants with ASD follows that used in several other empirical studies (Welchow et al., 2005; Golan et al., 2006; Lombardo et al., 2007; Gomot et al., 2008; Ashwin et al., 2009; Minio-Paluello et al., 2009; Morita et al., 2012; Samson et al., 2012; Mathers et al., 2013). Average full-scale intelligence quotient (IQ) (Wechsler Abbreviated Scale of Intelligence; Wechsler, 1999) for the TD group (mean = 111) was not significantly different from the ASD group (mean = 103; P = 0.23). Examination of the subscales showed no group differences in block design T-scores [TD mean(s.d.) = 55(10), ASD mean(s.d.) = 51(15); P = 0.25], matrix reasoning [TD mean(s.d.) = 53(11), ASD mean(s.d.) = 52(12); P = 0.94] or similarities [TD mean(s.d.) = 55(11), ASD mean(s.d.) = 51(9); P = 0.25], but group means differed for vocabulary [TD mean(s.d.) = 57(12), ASD mean(s.d.) = 45(13); P < 0.001]. These low-vocabulary scores were consistent with scores obtained on the Peabody Picture Vocabulary Test-III (Dunn and Dunn, 1997), which trended toward being lower for the ASD group (mean = 96) relative to the TD group (mean = 111; P = 0.055). We used the Spielberger State-Trait Anxiety Inventory (STAI) (Spielberger et al., 1983) for participants ≥18 years old and the Screen for Child Anxiety Related Emotional Disorders (SCARED, parent report) (Birmaher, 1997) for participants <18 years old to assess trait anxiety. TD participants showed lower levels of trait anxiety as measured by the STAI [mean(s.d.) = 33(5), range: 24–47] compared with those in the ASD group [mean(s.d.) = 51(18), range: 20–74; F = 12.14, P < 0.005] and as measured by the SCARED [TD mean(s.d.) = 10(6), range: 1–24 and ASD mean(s.d.) = 18(11), range: 3–41; F = 8.94, P < 0.005]. All participants or their parents provided written informed consent approved by the local review board.

**Procedure**

Data collection occurred over 2 separate days. On the first day, participants completed behavioral and eye-tracking measures and were acclimated to a mock scanner to determine whether participants felt comfortable for the MRI scanning session, which took place on a second day within the following 2 weeks. Although data collection measures included angry, neutral and happy face stimuli (described below), our analyses focus on angry and neutral faces for the purpose of this report. Results from the happy condition are reported in the Supplementary Figure 1.

**Expression processing**

**Subjective threat ratings**

Participants were shown 18 gray scale images of facial expressions (Tottenham et al., 2009b) presented singly on flashcards. Participants...
Amygdala and autism

provided ratings on a scale from '1' (not threatening) to '10' (extremely threatening) to assess how threatening they perceived each expression. Participants proceeded at their own pace. Responses were recorded and averaged to obtain one threat score for each expression.

**Labeling**

Participants were shown 18 gray scale images of facial expressions (Tottenham et al., 2009b) presented singly on flashcards. Participants were asked to indicate whether these faces 'felt' angry/sad/neutral/surprised/happy/afraid/disgusted/none of these. Participants proceeded at their own pace. Responses were recorded and scored for accuracy. The order of the labeling task and the subjective threat ratings task was counterbalanced across participants.

**fMRI task**

Face stimuli (angry, neutral, happy; Tottenham et al., 2009b) were presented in two counterbalanced runs (1: natural viewing, 2: experimental gaze-manipulation), with 36 stimuli per run, with a random fixed order within each run. That is, we used a randomization procedure to order the stimuli, and then used this order for all of the participants. In the natural viewing run, participants passively viewed each face and, to increase task engagement, were instructed to press a button each time a face stimulus appeared (Figure 3A). Participants were instructed to alternate pressing the button with their index and middle fingers for each trial [two fingers were used to match the behavior in the experimental gaze-manipulation run (described next)]. Thus, participants were instructed to use their index finger for the first stimulus, middle finger for the second stimulus, and so on, continuing to alternate fingers throughout the task. In the experimental gaze-manipulation run (Figure 4A), participants viewed these same faces, but with a visually degraded geometric shape placed either in the right or the left eye of the face stimulus. The task was to locate the shape (right or left) with either the index (if on left) or the middle finger (if on right). This condition was designed to increase eye movements toward the eye region, which was confirmed with out-of-scanner eye tracking (see below). Each trial lasted 2500 ms, which was composed of 300 ms presentations of face stimuli presented at a vertical visual angle of 15° arranged such that the tip of the nose (rather than the eyes) was aligned with the intertrial interval central fixation cross] and 2200 ms of fixation, during which responses were collected. In addition, an average 5 s jitter was included between each trial. Each of the two runs lasted 5 min, 35 s. Advancement of each trial was independent of participants’ responses. There were no group differences in accuracy [TD mean (s.d.) = 84% (12%); ASD mean (s.d.) = 81% (15%); P = 0.41]. Although the average accuracy scores were fairly high, all participants included in this manuscript had accuracy scores ≥50%. We chose this liberal threshold because this study included young children and the behavioral task was used primarily to ensure task engagement during an essentially passive viewing task.

Eye-movement measures were obtained during an identical task out of the scanner during the first visit using table-mounted eye-movement equipment (ISCAN, Inc.). Eye movements were recorded at a rate of 60 data points/s (60 Hz), averaged over both eyes. Face stimuli (300 ms) were arranged such that the tip of the nose was aligned with the intertrial interval central fixation cross (1000 ms). Eye-movement measures captured the initial saccade made following stimulus onset. The variable of interest was eye movements in the upward direction toward the eyes (Kliemann et al., 2010). Eye-movement coordinates were output in ASCII format. From these coordinates, a change score was computed in vertical gaze coordinates from central fixation to initial saccade using the output gaze coordinates; trials with positive change scores were scored as a 1, and those with no change or a negative change we marked with a 0. Thus, we calculated the proportion of trials with upward direction (toward the eye-region) for each participant.

**Image acquisition**

Subjects were scanned with a General Electric Signa 3.0-T fMRI scanner (Milwaukee, Wisconsin) with a quadrature head coil. A high-resolution T1-weighted anatomical scan [3D magnetization prepared rapid acquisition gradient echo (MPRAGE) 256 × 256 in-plane resolution, 240 mm field of view (FOV); 124 sagittal slices of 1.5 mm] was acquired for transformation and localization of functional data into Talairach space (Talairach and Tournoux, 1988). A spiral in-and-out sequence (Glover and Thomason, 2004) was used to collect functional data (repetition time (TR) = 2500, echo time = 30, FOV = 200 mm, Flip angle = 90, 64 × 64 matrix). We obtained 34 coronal slices of 4 mm thickness (skip 0) with a resolution of 3.125 × 3.125 mm.

**Imaging data analysis**

Functional imaging data were preprocessed and analyzed with the Analysis of Functional NeuroImages (AFNI) software package (Cox, 1996). After slice time correction, images were registered to the first volume and smoothed with an isotropic 6 mm Gaussian kernel. Time series were normalized to percent signal change by dividing signal intensity at each time point by the mean intensity for that voxel and multiplying the result by 100. An individual model was fit for each subject, which included regressors for each stimulus type, accuracy and 6 motion parameters by convolving the stimulus timing files with a gamma-variate hemodynamic response function. Standard general linear modeling without auto-regression correction was performed to fit the time courses to each regressor. Linear and quadratic trends were modeled in each voxel timecourse to control for correlated drift, and data were transformed into the standard Talairach coordinate space and resampled resolution of 3 mm³. Group-level linear mixed effects (LME) models were conducted with the 3dLME program within AFNI, which uses functions from the R software package (http://www.R-project.org). Three separate voxel-wise LME models were computed: Group (TD, ASD) by Emotion (angry, neutral) under natural viewing conditions; Group (TD, ASD) by Emotion (angry, neutral) under the experimental gaze-manipulation condition; and Emotion (angry, neutral) by Face Viewing Condition (natural viewing, gaze manipulation) within the ASD group to examine within-group differences. All LMEs were performed with age as a covariate. Each trial lasted 2500 ms. Correction for multiple comparisons was applied at the cluster level following Monte Carlo simulations conducted in the AlphaSim program within AFNI (for alpha <0.05, small-volume correction: FWHM = 6; # simulations = 10 000; individual voxel threshold = 0.02; the minimum number of voxels necessary to achieve P < 0.05 = 8 × 3 × 3 voxels). Clusterwise false positive rates of P < 0.05 for small volume correction were applied (Kim et al., 2004). Beta (β) coefficients were extracted from significant regions of the right and left amygdala, which served as our parameters of interest and analyzed with in SPSS.

**RESULTS**

**Eye-movements**

**Natural viewing**

A 2 × 2 (Group, Emotion) repeated measures analysis of covariance (ANCOVA) was performed on the proportion of trials directed toward the eye region during natural viewing as the dependent measure, with age and sex entered as covariates. There was a trend-level main effect of group [F(1,61) = 3.46, P = 0.07, η² = 0.05], where the ASD group made fewer eye movements toward the eye region. Importantly,
there was a Group × Emotion interaction \(F(1,61) = 5.14, P<0.05, \eta^2 = 0.08\). Post hoc tests showed that eye movements were most different for neutral faces \(F(1,61) = 7.20, P<0.01, \eta^2 = 0.11\). As Figure 1A shows, participants in the ASD group were significantly less likely to direct gaze toward the eyes for neutral faces, whereas the two groups were similar in eye movement for angry faces (\(P = 0.55\)).

**Confirmation of gaze manipulation**

To confirm that we successfully increased gaze directed toward the eye region, we examined the eye-movement data in the two face viewing conditions, natural looking and experimental manipulation. A 2 × 2 (Group, Face Viewing Condition) repeated measures ANCOVA confirmed that there was a significant increase in proportion of trials directed upward toward the eye region in the experimental condition across participants \(F(1,61) = 8.21, P<0.01, \eta^2 = 0.12\;\text{Figure 4B}\).

**Threat ratings**

A 2 × 2 (Group, Expression) repeated measures ANCOVA was performed on average threat ratings, with age and sex entered as covariates. There was a main effect of emotion \(F(1,64) = 9.08, P<0.005, \eta^2 = 0.12\), such that angry faces were rated as more threatening than neutral faces by all participants. There were no other main effects or interactions.

To examine whether the eye-movement data were associated with threat ratings within the ASD group, we performed a linear regression on the dependent measure of eye movements toward the eye region during natural viewing for neutral faces with the independent variable of threat appraisals for neutral faces, controlling for age and sex. As Figure 1B shows, there was an inverse association such that those participants who gave high threat ratings for neutral faces were less likely to produce eye movements toward the eyes of neutral faces (\(\beta = -0.46, P<0.05\)). There was no significant association between threat ratings for angry faces and eye movements toward the eyes of angry faces (\(\beta = -0.28, P=0.25\)), nor was there any association between eye movements toward the eye region and threat ratings for the TD group (angry \(\beta = -0.03, P=0.91\); neutral \(\beta = 0.00, P=0.98\)).

**Labeling accuracy**

A 2 × 2 (Group, Expression) repeated measures ANCOVA was performed on the proportion correct in labeling accuracy, with age and sex entered as covariates. There was a main effect of group \(F(1,53) = 8.25, P<0.01, \eta^2 = 0.14\); the Group × Emotion interaction remained \(F(1,53) = 4.27, P<0.05, \eta^2 = 0.07\) even when accounting for verbal ability. Figure 2B shows the distribution of errors for neutral faces, where neutral faces were often mislabeled as negative facial expressions. To quantify this observation, we computed the number of times a neutral face was mislabeled as a positive expression (happy) and a negative expression (angry, fear, disgust, sad) (we divided this value by 4 to account for the greater number of negative options) and performed an additional 2 × 2 (Group, Error Valence) repeated measures ANCOVA on the dependent measure of neutral labeling errors. Confirming what is shown in Figure 1B, there was a main effect of group \(F(1,54) = 12.35, P<0.001, \eta^2 = 0.17\), which was qualified by a Group × Error Valence interaction \(F(1,54) = 11.98, P<0.001, \eta^2 = 0.18\). Post hoc tests showed that the ASD group was more likely to mislabel neutral faces as a negative emotion than the TD group \(F(1,54) = 9.64, P<0.005, \eta^2 = 0.15\). There were no other main effects or interactions.

To examine the association between threat appraisals and labeling accuracy for neutral faces, we performed a linear regression controlling for group, age and sex with the independent variable of neutral threat ratings and the dependent variable of negatively valenced labeling errors for neutral faces. There was a strong positive association such that, as Figure 2C shows, higher threat ratings for neutral faces were associated with more negatively valenced labeling errors for neutral faces (\(\beta = 0.52, P<10^{-5}\)).
Imaging results

Natural viewing of faces—group differences

A $2 \times 2$ (Group, Emotion) repeated measures ANCOVA on the dependent measure of reaction time, controlling for sex and age revealed no main effects or interactions on reaction time ($P > 0.05$). As Figure 3B (top) shows, the LME revealed a significant main effect of group in the right ($F = 4.94$, $P < 0.05$, small-volume corrected; $xyz = 25 0 18$) and left amygdala ($xyz = 25 4 17$). To explore the nature of this main effect, beta weights from these functional regions were plotted in Figure 3B (bottom). Results from additional tests are provided in the Supplementary Figure 2.

Experimental gaze-manipulation to the eye region—group differences

We examined group differences in amygdala response in the experimental condition when gaze was directed upward toward the eye region. We performed a $2 \times 2$ (Group, Emotion) repeated measures ANCOVA on the dependent measure of reaction time under the gaze-manipulation condition, controlling for age and sex. There were no main effects or interactions on reaction time ($P > 0.05$). The results of the LME in AFNI revealed a significant Group $\times$ Emotion interaction in the right ($F = 4.93$, $P < 0.05$, small volume corrected; $xyz = 23 2 21$) and left ($xyz = 23 2 14$) amygdala (Figure 4C). To examine the nature of the interactions, post hoc tests were performed on the extracted beta weights from these activated regions, controlling for age and sex. As shown in Figure 4D, the source of the interaction was the high amygdala response to neutral faces in the ASD group ($P < 0.0005$, left $P < 0.05$). That is, compared with the TD group, those with ASD showed significantly higher amygdala response when presented with neutral faces but no difference from the TD group when presented with angry faces.

Effect of gaze manipulation within the ASD group

To further explore the effect of gaze on amygdala response within the ASD group alone, another LME was conducted in AFNI that directly compared amygdala response under natural viewing conditions with amygdala response when gaze was experimentally manipulated toward the eye region. Specifically, a $2 \times 2$ (Emotion $\times$ Face Viewing Condition) LME revealed an Emotion $\times$ Face Viewing Condition interaction in the right ($F = 5.27$, $P < 0.05$ small volume corrected; $xyz = 19 6 17$) and left ($xyz = 16 5 15$) amygdala (Figure 5A).
To explore the nature of these interactions, post hoc tests were performed on the extracted beta weights from these regions, controlling for age and sex. Experimental manipulation of gaze to the eye region did not influence amygdala response to angry faces (right $P = 0.92$, left $P = 0.31$), but potentiated amygdala response for neutral faces for both the right [$F(1,24) = 4.32$, $P < 0.05$, $\eta^2 = 0.13$] and left amygdala [$F(1,24) = 6.68$, $P < 0.025$, $\eta^2 = 0.19$; Figure 5B]. This last finding suggests that individuals with ASD are not typically looking at the eye region of neutral faces, a suggestion that was supported by our earlier eye-movement findings. Taken together, these analyses show that for individuals with ASD, forcing eye contact with neutral faces potentiates amygdala activity.

**Eye gaze associated with magnitude of amygdala change within ASD group**

We compared eye-gaze measures under the natural viewing condition for neutral faces, which were taken out of the scanner, with amygdala potentiation via experimental gaze manipulation. We calculated a change score for amygdala response to neutral faces during the experimental gaze-manipulation condition minus the natural viewing condition. A linear regression controlling for age and sex showed that the proportion of naturally occurring eye movements toward the eye region was inversely associated with the amygdala difference (Figure 5C; $\beta = -0.47$, $P < 0.05$). That is, participants who made fewer eye movements toward the eye region showed the largest potentiation in amygdala response when gaze was experimentally driven upward toward the eyes.

**Amygdala response is associated with threat appraisals**

We examined whether there was an association between amygdala response under the experimental gaze-manipulation condition and threat ratings for neutral faces within the ASD group. To test this association, we performed a linear regression comparing the amygdala response with neutral faces when gaze was manipulated toward the eye region with threat appraisals for neutral faces. As Figure 6 shows, when controlling for group, age and sex, right amygdala activity under conditions of experimental gaze directed at the eye region of neutral faces was positively associated with threat appraisals for neutral faces ($\beta = 0.42$, $P < 0.01$).

**Hierarchical regression: amygdala as mediator between group and labeling errors**

To test whether amygdala response statistically mediated the association between group and likelihood of mistaking a neutral face for a
negatively valenced face, we used hierarchical regression as specified by Baron and Kenny (1986). This analysis showed that group (when controlling for age and sex) was a significant predictor of labeling errors ($\beta = 0.47$, $P < 0.05$) (Table 2). In the second step of the model, group was simultaneously regressed on labeling errors along with amygdala signal to neutral faces under the experimental gaze manipulation condition (beta weights extracted from Emotion x Face Viewing interaction performed in the ASD group) as the mediator variable. The association between group and amygdala signal was significant as was the association between amygdala signal and labeling errors. Moreover, the association between group and labeling errors was mediated by amygdala signal ($\beta = 0.41$, $P < 0.05$), which when included in the analysis, explained the majority of the variance attributed to group, and the coefficient between group and labeling errors became non-significant when amygdala signal was included ($\beta = 0.27$, ns). The Sobel test (1982) examining the indirect effects of amygdala signal was significant ($Z = 2.02$, $P < 0.05$). This analysis shows that group differences in labeling errors for neutral faces were statistically mediated by amygdala signal to neutral faces.

DISCUSSION

Building on the growing literature exploring the association between face expression processing and amygdala response in ASD, we used a combination of eye-tracking, behavioral and fMRI methods to assess responses to images of angry and neutral faces. Collectively, our findings support the hypothesis that face emotion processing is altered in ASD, and amygdala response to faces are atypical, showing heightened reactivity to faces, consistent with the findings of several other fMRI studies (Dalton et al., 2005; Monk et al., 2010; Weng et al., 2011). Importantly, the largest group effects were evident for neutral faces, not angry. Individuals with ASD made more errors in labeling neutral faces, mostly confusing them for negatively valenced expressions, and these errors were positively associated with threat ratings for neutral faces. The eye-tracking data showed that during natural viewing, individuals with ASD showed diminished gaze toward the eyes of neutral faces, and eye contact was negatively associated with threat ratings, such that individuals in the ASD group who rated neutral faces as most threatening were least likely to direct gaze toward the eye region of those faces.

Gaze direction when viewing neutral faces was an important variable for examining amygdala response. The fMRI results showed that in the natural viewing condition, individuals with ASD showed diminished gaze toward the eyes of neutral faces, and eye contact was negatively associated with threat ratings, such that individuals in the ASD group who rated neutral faces as most threatening were least likely to direct gaze toward the eye region of those faces.

Gaze manipulation increased looks toward the eye region, a Group x Emotion interaction emerged (C), where individuals in the ASD group showed elevated amygdala response to neutral faces, averaged over both hemispheres for this figure (D). TD, typically developing; ASD, autism spectrum disorder; *between-group difference, $P < 0.05$; **within-group difference $P < 0.01$.
amygdala response. The out-of-scanner eye-tracking data suggested that naturally occurring gaze toward the eye region was diminished for neutral faces, but then increased via experimental manipulation, and amygdala potentiation was thus observed. The amount of naturally occurring gaze toward the eye region was associated with the magnitude of amygdala potentiation during the experimental condition relative to the natural viewing condition. That is, those individuals in the ASD group who gazed toward the eye region least showed the largest increase in amygdala signal during the experimental gaze-manipulation condition. Moreover, amygdala response under the experimental gaze-manipulation condition correlated with subjective threat ratings of neutral faces and mediated group differences in labeling errors for neutral faces.

Collectively, the behavioral, eye-tracking and neuroimaging results of this study suggest that neutral faces are important to consider, as individuals with ASD may not always perceive them as ‘neutral’, which calls into question the appropriateness of using neutral faces as a baseline condition in fMRI studies. Our sample size was larger than most fMRI studies of ASD, which may have allowed for the observation of neutral faces’ effects. The current study is not the first to observe aberrant processing of neutral faces in ASD, as previous work has already shown that the greatest amount of eye-contact avoidance in ASD occurs with neutral faces (Kliemann et al., 2012) and others have shown associations between hyperactive amygdala signal and neutral faces in ASD (Kleinhans et al., 2009; Dichter et al., 2012). It may be that neutral faces are less familiar or more affectively ambiguous, stimulus characteristics that have been shown to increase amygdala response in TD children (Thomas et al., 2001). One interpretation of the findings in typical children is that the heightened amygdala response to neutral faces may reflect an increased sensitivity to the affective ambiguity of neutral faces in children, perhaps as a result of developmental differences in experience (discussed in Tottenham et al., 2009a). It may be that in ASD, neutral faces similarly have great affective ambiguity (perhaps due to decreased visual experience with faces in general) and therefore elicit a strong amygdala response. In general, higher amygdala response to facial stimuli may be an index of

Fig. 5 Effect of experimental gaze manipulation with ASD group. When gaze was directed to the eye region, amygdala response was potentiated for neutral faces in the ASD group. Panel (A) shows effect of Emotion × Face viewing condition interaction in the right and left amygdala, and panel (B) explains the source of the interaction, where the experimental gaze manipulation condition resulted in a potentiated amygdala response to neutral faces, averaged across both hemispheres for this figure. Panel (C) shows an inverse association between gaze directed toward the eye region of neutral faces under natural viewing conditions and amygdala potentiation by experimental eye gaze (experimental eye-gaze condition minus natural viewing); residuals controlling for age and sex are plotted. ASD, autism spectrum disorder; ns, not significant; *within-group difference, P < 0.025.
affective immaturity, as has been found in typical samples of children (Gee et al., 2013) and in the examination of age group effects in the current study (presented in the Supplementary Section). These age trends did not interact with group; that is, children in both the TD and ASD groups showed increased amygdala signal to faces relative to older ages, although individuals in the ASD group had higher amygdala signal over all. Ambiguous (blended) facial expressions have been shown to be difficult for individuals with ASD to interpret and resulted in a negativity bias in their responses (Kuusikko et al., 2009). Ambiguity and uncertainty are features that have attenuated visual attention in ASD in other domains (unpredictable toys Ferrara and Hill, 1980), as they are attributes that are at odds with the inflexible adherence to routine and predictability that characterize ASD. These findings raise the question of whether social ambiguity and uncertainty itself is an aversive aspect of facial stimuli for individuals with ASD and may be associated with the ‘resistance to change’ characteristic common in ASD (Gomot et al., 2008; Lionello-DeNolf et al., 2010; Qian and Lipkin, 2011; Duerden et al., 2012). If true, then face-training interventions should result in reduced amygdala response over time to ambiguous facial expressions like neutral.

These findings are consistent with the hypothesis that individuals with ASD who avoid eye contact may do so to reduce the emotional overarousal that accompanies direct eye contact (Dalton et al., 2005; Kliemann et al., 2010). Naturally occurring eye contact was diminished for neutral faces in the ASD group, and diminished eye contact was associated with higher threat ratings provided by the participants.

**Table 2** Hierarchical regression: amygdala activity to neutral faces mediates association between group and rate of mistaking neutral faces with a negatively valenced expression

<table>
<thead>
<tr>
<th>Variables in model</th>
<th>B</th>
<th>SE B</th>
<th>β</th>
<th>ΔR²</th>
<th>Sobel test</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Step 1</strong> Group</td>
<td>0.28</td>
<td>0.11</td>
<td>0.47*</td>
<td>0.25*</td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td>0.06</td>
<td>0.12</td>
<td>0.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age group</td>
<td>-0.10</td>
<td>0.06</td>
<td>-0.27</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Step 2</strong> Group</td>
<td>0.16</td>
<td>0.12</td>
<td>0.27</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex</td>
<td>0.06</td>
<td>0.11</td>
<td>0.10</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age group</td>
<td>-0.08</td>
<td>0.06</td>
<td>-0.22</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amygdala response</td>
<td>0.29</td>
<td>0.12</td>
<td>0.41*</td>
<td>0.13*</td>
<td>2.02*</td>
</tr>
</tbody>
</table>

*P < 0.05. B, Unstandardized Coefficient; SE B, Standard Error.
Amygdala response to neutral faces was potentiated when gaze was experimentally directed toward the eye region for individuals in the ASD group, and this potentiation was greatest for those individuals with the least amount of naturally occurring eye contact who provided the highest threat ratings for neutral faces. This within-subject neuroimaging finding suggests that under natural viewing conditions, individuals with ASD may modulate amygdala-mediated arousal by averting their gaze from direct eye contact.

In contrast to neutral expressions, angry faces are affectively anchored expressions, conveying clearer meaning than neutral faces. It is perhaps for this reason that angry face stimuli produced no behavioral group differences. That is, there were no group differences in labeling accuracy, eye contact or threat ratings for angry faces. Amygdala response to angry faces was higher in the ASD group under natural viewing conditions, but amygdala response to angry faces was not modulated between viewing conditions in the ASD group, suggesting that individuals with ASD were already more likely to be looking at the eye region of angry faces under natural viewing conditions—an assumption confirmed by the eye-tracking measure taken out of the scanner. The lack of effects for angry faces may in part be due to the fact that the angry faces used in the study were highly caricatured faces, and exaggerated faces have been shown to improve face processing in ASD (Rutherford and McIntosh, 2007). Moreover, this sample had IQ scores in the normal range, and thus, the caricatured angry faces may have presented little challenge. Additionally, although speculative, many face-training interventions emphasize canonical emotional expressions such as angry, fear, sad and happy, but may not include training on neutral faces (Silver and Oakes, 2001; Solomon et al., 2004; Golan et al., 2010; Lopata et al., 2010; Hopkins et al., 2011; Tanaka et al., 2012). Therefore, individuals with ASD may have a disproportionate amount of experience identifying angry faces relative to neutral. This experience may have attenuated group differences for angry faces. Future studies that use a face-training component can address this possibility.

There are limitations to this study. We were not able to collect in-scanner eye-tracking measures, and therefore cannot say with confidence where individuals with ASD looked in the natural viewing condition. However, the out-of-scanner eye-tracking measures and the experimentally induced change in amygdala signal to neutral faces increase our confidence that individuals with ASD were not looking at the eye region for neutral faces under natural viewing. The stimuli used in the behavioral and fMRI sessions were the same, which was done by design to compare behavioral and neural responses within the same individual. It is possible that experience with face stimuli in the behavioral session could have influenced neural responses collected at the scanning session. While the amygdala in typical adults has shown habitation effects within a single scan session (Breiter et al., 1996), individuals with ASD showed reduced habitation (Kleinhaus et al., 2009; Swartz et al., 2013). Sessions that are separated by several weeks have shown high reliability in amygdala signal in typical adults (Johnstone et al., 2005). This reliability was higher for faces like fear, and showed more variability for neutral (not necessarily a uniform increase or decrease across individuals). It is unknown whether group differences in habitation would be observed across multiple sessions as was used in the current study. Another matter concerns the reliability of the amygdala signal across testing sessions at the individual subject level. Previous work suggests that the amygdala’s response in typical populations may fluctuate at the individual level (although not the group level) in response to emotional faces (Plichta et al., 2012; van den Bulk et al., 2013). Although the current study did not acquire multiple scans, the reliability tests performed in these previous studies may suggest that amygdala signal is variable and perhaps subject to state effects. Importantly, there were significant associations between amygdala signal and behavioral measures in the current study, and behavior has been shown to be more stable within the individual (van den Bulk et al., 2013). Nonetheless, it will be important for future work to examine test-retest reliability at the individual level within atypical populations. We were unable to administer diagnostic interviews to all participants in the ASD group [e.g. Autism diagnostic observation schedule (ADOS)] to confirm diagnosis, owing to the difficulties of scheduling. Obtaining this confirmation is ideal. The ADOS interviews that were obtained confirmed the presence of an ASD in all cases, and the high AQ scores of the ASD group, although not diagnostic in and of themselves, provided confidence that the previous ASD diagnoses participants had were accurate. Another limitation pertains to the generalizability of these findings to all individuals with ASD. As is common of most study participants who can tolerate fMRI, the individuals included in this study had high IQ scores and verbal ability. These individuals may not be representative of all individuals with ASD and therefore, the results from the current study may not generalize to individuals who are more functionally impaired. We also used a wide age range in this study, a practice used in other studies examining this special and difficult-to-test population (Dalton et al., 2005; Palmen et al., 2006), although our groups were balanced with regard to age. Additionally, we could not obtain a balanced sample of male and female participants. We had to use this wide-ranging sample because of the difficulty obtaining a large enough sample for fMRI methods. We chose to statistically control for age and sex in all of our analyses, although we included supplemental findings to show trends. The developmental relationship between atypical neural and behavioral responses to faces in ASD remains an important question for future research.

Taken together, the findings are consistent with the hypothesis that the amygdala is hyperresponsive to facial expressions in ASD and this response is associated with increased threat ratings and negative interpretations. The current findings show that as a group, individuals with ASD are particularly prone to interpreting neutral faces as negative and exhibiting elevated amygdala response to these faces. For individuals with ASD, decreasing eye contact seems to be a means of modulating this response. We draw this conclusion based on the associations between eye tracking and threat ratings and potentiated amygdala response resulting from increasing eye contact. However, these data should not discourage the use of interventions that increase eye contact. On the contrary, face expression processing is a learned skill (Adolphs et al., 1995; Tottenham et al., 2009a) and interventions for face processing deficits need to include visual experience with faces. Indeed, individuals with ASD who make more eye contact show enhanced emotion recognition skills (Kirchner et al., 2011). Moreover, the heightened amygdala response observed in this study may be requisite for learning about faces. Previous work has shown that initial learning of any affective association necessitates amygdala signal increases (LaBar et al., 1998; Holland and Gallagher, 2006; Sarinopoulos et al., 2010). Therefore, we believe that increased arousal caused by eye contact may be a necessary and unavoidable aspect of face expression training programs. These data do not discourage the use of eye contact in face training programs, and provide insight into the neural mechanisms involved in eye contact in ASD.

**SUPPLEMENTARY DATA**

Supplementary Data are available at SCAN Online.

**REFERENCES**


