Sex-specific variation in facial masculinity/femininity associated with autistic traits in the general population

Abstract

Reports linking prenatal testosterone exposure to autistic traits and to a masculinised face structure have motivated research investigating whether autism is associated with facial masculinisation. This association has been reported with greater consistency for females than for males, in studies comparing groups with high and low levels of autistic traits. In the present study, we conducted two experiments to examine facial masculinity/femininity in 153 neurotypical adults selected for either low, mid-range or high levels of autistic traits. In the first experiment, their three-dimensional facial photographs were subjectively rated by 41 raters for masculinity/femininity and were objectively analysed. In the second experiment, we generated 6-face composite images which were rated by another 36 raters. Across both experiments, findings were consistent for ratings of photographs and composite images. For females, a linear relationship was observed where femininity ratings decreased as a function of higher levels of autistic traits. For males, we found a U-shape function where males with mid-range levels of traits were rated lowest on masculinity. Objective facial analyses revealed that higher levels of autistic traits were associated with less feminine facial structures in females and less masculine structures in males. These results suggest sex-specific relationships between autistic traits and facial masculinity/femininity.

Keywords: autism spectrum disorder; autistic traits; masculinisation; hypermasculinisation; facial morphology; facial features; facial composite; 3D photogrammetry

Autism spectrum disorder (ASD) is a neurodevelopmental condition characterised by difficulties in social communication as well as the presence of repetitive behaviours and restricted interests (American Psychiatric Association, 2013). ASD is approximately three times more prevalent in males than in females (Loomes, Hull, & Mandy, 2017). The male bias in the prevalence of ASD motivated the 'extreme male brain' theory which describes cognitive features of ASD as an exaggerated form of male-typical cognitive strengths (e.g., visuospatial skills) and weaknesses (e.g., emotion recognition; Baron-Cohen, 2002). For instance, compared to neurotypical individuals, both male and female autistic individuals demonstrate superior ability in processing visuospatial information (for a review see Muth, Hönekopp, & Falter, 2014) but relative difficulties in processing emotional information (Golan, Sinai-Gavrilov, & Baron-Cohen, 2015). The extreme male brain theory also posits that the shift towards hypermasculinised cognitive profiles in ASD may be influenced by exposure to increased levels of testosterone *in utero*.

Testosterone is a sex steroid with masculinising effects that is produced primarily in the male testes, and in lesser quantities by the adrenal glands in both males and females (Wood & Newman, 1995). Prior work has found that high levels of prenatal testosterone measured from amniotic fluid are associated with hypermasculinised cognitive profiles such as enhanced visuospatial skills (Auyeung et al., 2012) and reduced empathy (Knickmeyer, Baron-Cohen, Raggatt, Taylor, & Hackett, 2006). Evidence linking prenatal testosterone more specifically to autistic traits has been mixed, with some studies reporting a positive correlation (Auyeung et al., 2009; Auyeung, Taylor, Hackett, & Baron-Cohen, 2010) and others observing no relationship (Kung et al., 2016; Whitehouse, Mattes, Maybery, Dissanayake, et al., 2012). A further study reported elevated levels of several sex steroids (including testosterone) in the amniotic fluid of male fetuses who later received a diagnosis of ASD, compared to those who did not (Baron-Cohen et al., 2015). Thus, current evidence suggests that the early fetal hormone environment may play a part in the development of cognitive phenotypes associated with ASD.

During fetal development, the brain and the face evolve from the neural crest in synchrony (Le Douarin, Brito, & Creuzet, 2007). This has led to the speculation that prenatal factors that influence brain development, such as testosterone exposure, may also be linked to variations in facial structures (Marcucio, Young, Hu, & Hallgrimsson, 2011). Studies have investigated the link between early exposure to testosterone and facial masculinity using three methods for assessing prenatal testosterone exposure. First, the ratio between the length of the second and the fourth digits has been investigated, with lower ratios purportedly reflecting higher testosterone exposure (Manning, 2011). Several studies found that lower digit ratios were correlated with more masculine facial structures (Meindl, Windhager, Wallner, & Schaefer, 2012; Weinberg, Parsons, Raffensperger, & Marazita, 2015). Second, based on the proposition that girls in opposite-sex twin pairs are subjected to increased prenatal testosterone levels produced by their twin brothers relative to girls in same-sex twins (Tapp, Maybery, & Whitehouse, 2011), one study observed less feminine facial features among girls of opposite-sex twins than same-sex twins (Marečková et al., 2015). The third method involves measuring prenatal testosterone concentration directly from umbilical cord blood procured at birth. Using this method, Whitehouse et al. (2015) conducted a 20-year longitudinal study that identified a positive association between cord-derived testosterone concentrations and the degree of masculinity in facial structures in 183 typically-developing adults.

The studies reviewed so far suggest associations between prenatal testosterone exposure, the masculinised shift in ASD cognitive profiles, and facial masculinisation. Thus, it is reasonable to hypothesise that individuals with ASD may present with more masculinised facial features than neurotypical individuals, possibly through the influence of prenatal testosterone (Boutrus et al., 2017). Tan et al. (2017) investigated this assertion in prepubescent boys and girls with and without ASD, and reported that the girls and boys with ASD exhibited more masculinised facial features than their age- and sex-matched neurotypical counterparts. The aim of the current study is to investigate whether this facial hypermasculinity associated with autistic characteristics extends to the adult population more broadly.

There is accumulating evidence suggesting that autistic traits in the general population follow a smooth continuum with one end representing individuals with a clinical diagnosis of ASD (Robinson, Hallett, Happe, Plomin, & Ronald, 2011; Whitehouse, Hickey, & Ronald, 2011). This has led to several other studies investigating whether facial masculinisation observed in individuals with a clinical diagnosis of ASD extends to neurotypical individuals with relatively high levels of autistic traits. Tan et al. (2015) compared masculinity ratings of facial photographs of men with high versus low levels of autistic traits, and femininity ratings of women also selected for high versus low autistic-trait levels. Autistic traits were measured using the autism-spectrum quotient (AQ; Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001), with higher scores indicating higher levels of autistic traits. Facial masculinity ratings did not differ between the two groups of men, whereas women with high AQ scores received lower facial femininity ratings than did women with low scores on the AQ.

In another study, Scott et al. (2015) obtained facial photographs of men and women selected for high or low scores on the AQ. For each group, facial photographs were digitally manipulated to create a set of composite images. Using masculinity ratings for both male and female composite images across three experiments, raters consistently selected composites consisting of men with high AQ scores as being more masculine than images of low-AQ men. However, there was no difference in the perceived masculinity of composite images of women with high and low AQ scores. Although composite images are said to retain systematic variations related to varying autistic trait levels and minimise individual facial peculiarities that could bias ratings, ecological validity is a likely trade-off due to the limited realism of composite images (Sutherland, Rhodes, & Young, 2017). Still, the results presented in the Scott et al. study provide partial support for the hypermasculinisation account.

The Tan et al. (2015) and Scott et al. (2015) studies are similar in design whereby both studies collected perceived ratings of facial photographs taken from neurotypical university students selected for high and low levels of autistic traits using the AQ (Baron-Cohen et al., 2001). Nevertheless, the discrepancy in their findings may be attributed to two key methodological differences. First, in the Tan et al. study, raters were shown original photographs that are realistic but may lack experimental control, while Scott et al. presented composite images that are prototypes of high- and low-AQ individuals but may lack ecological validity. Second, the facial images of men and women were rated for masculinity and femininity respectively in Tan et al. whereas facial composites were rated for masculinity irrespective of gender in the Scott et al. study.

Furthermore, while several studies reported correlations between perceived facial masculinity/femininity ratings and objective measurements of sexually dimorphic facial structures, suggesting that subjective ratings could accurately reflect actual sexual dimorphism in faces (Gilani, Rooney, Shafait, Walters, & Mian, 2014; Koehler, Simmons, Rhodes, & Peters, 2004), other studies found that masculinity/femininity ratings may be influenced by facial features that were irrelevant to the sexual dimorphism (Komori, Kawamura, & Ishihara, 2011). For instance, facial ‘displacement’ (or asymmetry) is not typically associated with sex differences but greater asymmetry has been found to be related to higher facial masculinity ratings in males and higher facial femininity ratings in females. Thus, an objective examination of facial morphology using three-dimensional (3D) photogrammetry may shed some light on the inconsistencies in findings reported in the current literature.

Gilani et al. (2015) examined facial masculinity/femininity of neurotypical adults who scored high or low on the AQ using objective analyses of 3D facial images. It was reported that males with high AQ scores presented morphologically less masculine facial features than did males with low AQ scores. Females with high AQ scores presented less feminine facial structures than did females with low AQ scores. These findings align with a previous suggestion that physical characteristics associated with autistic traits may be better described as being ‘androgynous’ rather than hypermasculinised in both sexes (Bejerot et al., 2012).

Existing studies have only examined subjective facial ratings and objective morphological analyses of the neurotypical population for polarised groups of individuals with high or low scores on the AQ (i.e., the extreme ends of the distribution). It is therefore unclear whether differences in facial masculinity/femininity extend across the population or are evident only between high- and low-trait groups. The present study advances our understanding of the association between autistic traits and facial masculinity/femininity in three ways. First, we included a group of neurotypical individuals selected for mid-range levels of autistic traits in addition to the two polarised groups. Second, we conducted two experiments where we obtained ratings of facial images in their original form in Experiment 1 and ratings of composite images in Experiment 2. In both experiments, facial images of men were rated for masculinity and images of women were rated for femininity. Third, we took a further step in Experiment 1 by also objectively analysing the 3D images of the participants to characterise the morphological differences in facial masculinity and femininity.

Experiment 1

Experiment 1 provides a replication of Tan et al. (2015) in which original facial images were presented to a group of raters for masculinity/femininity ratings. In addition, 3D facial analyses were conducted to examine the association between levels of autistic traits and objectively derived morphological masculinity scores, and the association between objectively derived morphological masculinity scores and perceived masculinity/femininity ratings.

Method

Stimulus Participants. A total of 1,995 undergraduate students from the ****** completed the AQ. Students aged between 18 and 25 years and with an AQ score that fell in the bottom 15% (AQ ≤ 12; low-AQ group), middle 15% (15 ≤ AQ ≤ 18; mid-AQ group) or top 15% (AQ ≥ 22; high-AQ group) of the distribution were invited to participate further in the study. Eighty-one Caucasian females and 70 Caucasian males agreed to take part in the study (see Table 1 for descriptive statistics). Participants were identified as post-pubertal using the Puberty Developmental Scale (Petersen, Crockett, Richards, & Boxer, 1988). Informed consent was obtained from all participants, who were tested in accordance with ethical approval obtained from the Human Research Ethics Committee at the ******.

Raters. Forty-one Caucasian undergraduate students (18 males: mean age = 19.6 years, *SD* = 2.07 years; 23 females: mean age = 18.2 years, *SD* = 0.76 years) were additionally recruited to provide masculinity ratings for the male faces and femininity ratings for the female faces. To minimise the effects of facial familiarity on the ratings, the stimulus participants and raters were drawn from non-overlapping student cohorts.

Autism-spectrum Quotient (AQ; Baron-Cohen et al., 2001). The AQ is a 50-item self-report questionnaire that measures the extent of autistic traits in adults in the general population. The AQ has good test-retest reliability and validity (Austin, 2005; Baron-Cohen et al., 2001; Ruzich et al., 2015), with scores ranging from 0 to 50. The cut-offs for low and high scores employed in the current study are comparable to those used in previous studies (Jackson et al., 2013; Russell-Smith, Maybery, & Bayliss, 2011).

<insert Table 1 about here>

Facial photography. Each stimulus participant was photographed front-on using a 3dMDface system (3dMD, Atlanta, GA, USA) in a room where lighting was kept constant. The 3dMDface system generates three-dimensional (3D) images by combining multiple two-dimensional (2D) images captured using colour cameras placed at two stereo camera viewpoints. In addition, the 3dMDface system incorporates the use of infrared cameras that standardise the distance between the system and the participant such that each facial image is true to its size. All participants were free of facial hair and cosmetics and removed accessories such as ear-studs prior to photography. Participants presented a neutral facial expression with their mouths closed.

Using the3dMD face system (3dMD, Atlanta, GA, USA) and Photoshop CS 5, the 3D facial images were converted to front-facing 2D images in which the pupils were aligned horizontally and spaced 80 pixels apart. An ellipse was placed over each face so that the ears, hair and neck were covered with a grey frame. Image sizes were standardised to 320 × 420 pixels and presented in black-and-white.

Masculinity/Femininity rating. The rating procedure outlined in Tan et al. (2015) was used in the current study. For each trial, raters were presented with a facial image for 5s, followed by a 10-point slider scale anchored by ‘not at all masculine’ (rating of 0) and ‘extremely masculine’ (rating of 10) for male faces, and ‘not at all feminine’ (rating of 0) and ‘extremely feminine’ (rating of 10) for female faces. Then raters were asked if they were familiar with the facial image presented to them (none reported in the affirmative for any image). The facial images were rated in separate blocks for each sex, with order of blocks counterbalanced across raters. The order of the faces within each block was randomised for each rater.

Morphological analyses. Each facial image was pre-processed in preparation for analysis using Meshlab (Cignoni et al., 2008) which was used to remove extraneous portions (e.g., background and hair). Then each image was annotated with 42 landmarks (see Figure 1) in MorphAnalyser 2.4 (Tiddeman, Duffy, & Rabey, 2000) which were aligned in orientation, rotation, and scale using Procrustes superimposition. MorphAnalyser also resampled the surface map of each image based on a standard model annotated with the same set of landmarks. This procedure aligned the surface map of each image to those of the standard model, thus achieving homology for the entire facial surface of each image and allowing the analyses and visualisation of whole surfaces (Holzleitner et al., 2014; Tan et al., 2019).

All 3D images were subjected to Principal Component Analysis (PCA) using MorphAnalyser. The PCA identified 24 factors which accounted for 86.3% of the overall variations in facial shapes. The resulting factor loadings were used to compute a 'morphological masculinity score' (MMS) for each participant. First, a weight W for each factor was calculated as the ratio of the between-sex difference in factor loadings L for the factor, to the square root of the sum of squared sex differences across all factors, that is,

$$W_i = \frac{\overline{L}_{M_i} - \overline{L}_{F_i}}{\sqrt{\sum_{j=1}^{24} (\overline{L}_{M_j} - \overline{L}_{F_j})^2}}, \; i = 1, 2, \ldots, 24 \qquad (1),$$

where $\overline{L}_M$ is the mean factor loadings of males and $\overline{L}_F$ is the mean factor loadings of females. The MMS for each participant n was then computed as the dot product of the individual factor loadings and the weights, that is,

$$MMS_n = L_n \cdot W \qquad (2),$$

where MMS_n is the unstandardised morphological masculinity score calculated for each participant. A standardised MMS_n = 0 indicates facial structures that are neither masculine nor feminine, a MMS_n > 0 indicates masculinity, and a MMS_n < 0 indicates femininity.

< Insert Figure 1 about here >

Statistical analyses. Masculinity/Femininity rating data were analysed using linear mixed effects modelling conducted in R (R Core Team, 2017) and models were estimated using the lmer() function of lme4 (Bates, Mächler, Bolker, & Walker, 2015). AQ scores were entered as fixed effects and raters were included in the model as random effects. As male and female stimulus participants were rated on separate scales, that is, males were rated for masculinity while females were rated for femininity, analysis was conducted separately for each sex. Linear regression analyses were used to examine the relationship between AQ scores and the morphological masculinity scores, and the relationship between masculinity/femininity ratings and morphological masculinity scores.

Results

AQ scores and masculinity/Femininity ratings. Item-rest correlation analyses were used to identify any raters whose ratings correlated poorly with the rest of the raters. All correlations were above the recommended cut-off of $r = .20$ (Zijlmans, Tijmstra, van der Ark, & Sijtsma, 2018) and none was excluded from subsequent analyses. Intraclass correlation coefficient analysis revealed excellent inter-rater agreement in the femininity ratings of the female faces ($r = .97$, $p < .001$) and high internal consistency (Cronbach's α= .98). Similarly, the masculinity ratings of male faces showed excellent inter-rater agreement ($r = .95$, $p < .001$) and high internal consistency (Cronbach's α = .96).

Linear mixed model analysis indicates that females with higher AQ scores received lower femininity ratings (β = – 0.029, SE = 0.0043, $t = 6.66$, $p < .001$; see Figure 2). For males, the association between AQ scores and masculinity ratings was not statistically significant ($p = .22$). However, group comparison using repeated measures ANOVA suggests that the association between AQ scores and masculinity ratings may be non-linear (see Figure S1 in the Supplementary Material). Therefore, a quadratic term (squared-AQ scores) was added to the model. Indeed, squared-AQ scores significantly predicted facial masculinity ratings across a U-shape function, where lower and higher AQ scores were associated with higher masculinity ratings while mid-range levels of AQ were related to lower masculinity ratings (β = 0.0037, SE = 0.00054, $t = 6.91$, $p < .001$; see Figure 2).

<insert Figure 2 about here>

AQ scores and morphological masculinity scores. Linear regression analyses revealed that AQ score was a significant predictor of morphological masculinity score for females, $F(1, 79) = 4.98$, $p = .03$, $R^2 = .06$, where women with higher AQ scores had higher morphological masculinity scores (i.e., less feminine facial structures, β = 0.019, SE = 0.0084, $t = 2.23$, $p = 0.03$). For males, AQ scores also significantly predicted morphological masculinity scores, $F(1, 68) = 4.23$, $p = .04$, $R^2 = .06$, where men with higher AQ scores presented *lower* morphological masculinity scores (β = – 0.022, SE = 0.011, *t* = 2.06, *p* = .04; Figure 3).

<insert Figure 3 about here>

Masculinity/femininity ratings and morphological masculinity scores. Morphological masculinity scores significantly predicted femininity ratings in females, *F*(1, 79) = 10.8, *p* = .002, R^2 = 0.12, where women with higher morphological masculinity scores received lower femininity ratings, β = –0.77, SE = 0.24, *t* = 3.29, *p* = .002. However, morphological masculinity scores did not predict masculinity ratings in males (*p* = .39).

Discussion

The results of the current experiment provide a clear replication of the effects reported by Tan et al. (2015), which are displayed in Figure S2. In particular, as in Tan et al., AQ scores were negatively correlated with femininity ratings along a linear pattern where females with higher AQ scores received lower facial femininity ratings. For males, while outcomes were consistent with Tan et al.'s findings in that facial masculinity ratings did not differ between high and low AQ scorers, the inclusion of men with mid-range levels of autistic traits showed a U-shape quadratic trend where mid-AQ men were rated lowest on facial masculinity.

The morphological masculinity scores generated from the analyses of the 3D images indicated that females with higher AQ scores had structurally less feminine facial features, whereas males with higher AQ scores presented structurally less masculine facial features. These findings neatly align with those of Gilani et al. (2015), who identified the same pattern of associations. We also examined whether morphological masculinity scores predicted subjective masculinity/femininity ratings. For females, higher morphological masculinity scores predicted lower femininity ratings, suggesting that structural aspects of female faces did influence perceived femininity. However, for males, morphological masculinity scores did not predict perceived masculinity, which suggest that masculinity ratings may be driven by facial structures that are unrelated to sexual dimorphism. Given the possibility that individual facial idiosyncrasies may have influenced the masculinity/femininity ratings of original facial images, we created composite images for each AQ group to directly test this assertion in Experiment 2.

Experiment 2

As previously discussed, the different findings reported by Tan et al. (2015) and Scott et al. (2015) might be attributed to the use of original versus composite images (i.e., Tan et al. used original images and observed an effect in women but not in men; Scott et al. used composite images and observed an effect in men but not in women). With this second experiment, we used the original facial photographs detailed in Experiment 1 to create 6-face composite images for each AQ group. If the outcomes of Experiment 2 are similar to those reported in Experiment 1 and in Tan et al., this would imply that the facial ratings were unlikely to have been influenced by unsystematic individual-level variations captured in the original images (i.e., photographs). In contrast, if the results of Experiment 2 replicate findings reported in Scott et al., this would suggest methodology is critical in the use of original versus composite images in face rating studies.

Method

Composite images. Following the method described in Scott et al. (2015), 18 photographs were randomly chosen from each AQ group for each sex. The 18 facial images were then randomly divided into three sets of six images that were used to generate three 6-face composite images using Psychomorph (Tiddeman, Burt, & Perrett, 2001). In summary, three composite images were generated for each AQ group for each sex, resulting in a total of 18 composite images used for the rating task. Demographic information of the stimulus participants included in Experiment 2 is presented in Table 2.

Raters. An additional group of 36 Caucasian raters was recruited for Experiment 2 (11 males, mean age = 21.5 yrs, *SD* = 1.81; 25 females, mean age = 20.5 yrs, *SD* = 1.30).

Raters provided masculinity ratings for male composite faces and femininity ratings for female composite faces using the same procedure described in Experiment 1.

Results

Item-rest correlation of 11 raters who provided masculinity ratings and 10 raters who provided femininity ratings were below the cut-off of $r = .20$ (Zijlmans et al., 2018), hence their ratings were excluded from the analyses. Intraclass correlation coefficient analysis revealed good inter-rater agreement ($r = .84$, $p < .001$) and high internal consistency (Cronbach's $\alpha = .93$) of the femininity ratings of the female composite faces. Similarly, the masculinity ratings of male composite faces showed good inter-rater agreement ($r = .83$, $p < .001$) and high reliability (Cronbach's $\alpha = .94$).

Linear mixed model analysis showed that higher AQ scores were associated with lower femininity ratings of the female composite faces ($\beta = -0.05$, SE = 0.01, $t = 4.49$, $p < .001$; see Figure 4). However, AQ scores did not predict masculinity ratings of male composite faces in males ($p = .20$). As with males included in Experiment 1, group comparison suggests that the relationship between AQ scores and masculinity may be non-linear (see Figure S3 in the Supplementary Material). Hence, squared-AQ scores were added to the model, with the analysis supporting a U-shape function between AQ scores and masculinity ratings, where extreme AQ scores were related to higher masculinity ratings while mid-range AQ scores were associated with lower masculinity ratings ($\beta = 0.02$, SE = 0.003, $t = 6.36$, $p < .001$; see Figure 4).

<insert Figure 4 about here>

Discussion

The results of Experiment 2 with composite images largely replicate the effects reported in Experiment 1 and Tan et al. (2015) with original photographs of individuals. Specifically, there was a negative linear relationship between the facial femininity ratings of composite images and autistic traits in women, and a U-shape quadratic trend between masculinity ratings and AQ scores was observed in men. Together, the results of Experiments 1 and 2 provided compelling evidence that masculinity/femininity ratings are unlikely to be influenced by unsystematic variations in the facial images presented to the raters.

General Discussion

In the present study, we conducted two experiments that investigated: (1) subjective and objective measures of facial masculinity/femininity in neurotypical men and women with varying levels of autistic traits and (2) the subjective ratings of 6-face composite images of participants with low, mid-range, or high levels of autistic traits. Using subjective masculinity/femininity ratings of original images, we found that ratings of extreme-scoring groups (i.e., low- versus high-AQ) were consistent with those previously reported by Tan et al. (2015). The inclusion of groups of individuals with mid-range AQ scores revealed intriguing new findings. For women, femininity ratings decreased as a function of AQ score along a linear pattern. However, for men, a U-shaped trend was observed in the ratings of original and composite images, where males with mid-range levels of autistic traits received lower masculinity ratings than did males with high or low trait levels. Morphological analyses of the 3D facial images of these participants showed that females with higher AQ scores presented more masculine facial structures which were associated with lower observer femininity ratings. Conversely, males with higher AQ scores presented less masculine facial structures which did not correlate with observer masculinity ratings.

We then tested the possibility that the ratings of original facial images may be influenced by individual facial idiosyncrasies by creating composite images, each morphed from images of six male or female participants with similar AQ scores. However, the patterns of variation as a function of AQ score between masculinity/femininity ratings of original images and those of composite images were remarkably similar, where a negative linear relationship was observed in females and a U-shape quadratic function was found in males.

Overall, we observed a negative association between level of autistic traits and either perceived or morphological facial femininity in women. The current results are consistent with previous findings that have reported less feminine facial structures in prepubescent girls (Tan et al., 2017) and adult women (Bejerot et al., 2012) with an ASD diagnosis, and in neurotypical women with high AQ scores (Gilani et al., 2015; Tan et al., 2017), relative to appropriate comparison groups. Reduced facial femininity has also been linked to elevated exposure to perinatal testosterone in women from the general population (Whitehouse et al., 2015). While Scott et al. (2013) reported no statistically significant difference in the masculinity ratings for composite images of high- and low-AQ women, this outcome could be because masculinity ratings are less sensitive than femininity ratings in judging female faces (Mitteroecker, Windhager, Müller, & Schaefer, 2015; Munson, 2007). Therefore, existing evidence converges on the conclusion that reduced facial femininity in women with high levels of autistic traits may be associated with exposure to heightened levels of testosterone *in utero* and that its ‘de-feminising’ effects on facial structure remain stable throughout development.

By contrast, the association between perceived and morphological facial masculinity and levels of autistic traits in men appears to be more complex in that the variation in facial masculinity ratings follows a quadratic trend across increasing levels of autistic traits whereas morphological analyses suggested that males with higher AQ scores presented structurally less masculine faces. In light of a recent study which reported hypermasculinised facial structure in prepubescent autistic boys (Tan et al., 2017), it is possible that variations in facial masculinity as a function of autistic traits are age- and sex-specific. We speculate that the relationship between facial masculinity and autistic traits may be transient: evident in early development (Tan et al., 2017) but weaker or more complex in nature in post-pubertal men (Bejerot et al., 2012; Gilani et al., 2015; Scott et al., 2015; Tan et al., 2015).

It is unclear why morphological masculinity scores derived from morphological analyses failed to predict subjective masculinity ratings in the current study. One possible explanation is the presence of facial characteristics other than sexually dimorphic features that may have influenced raters' perceived masculinity ratings. A previous study provided evidence suggesting that in addition to sexual dimorphism in facial structures, impressions such as perceived height and weight could independently influence judgements of facial masculinity (Holzleitner et al., 2014). Furthermore, even though participants in the current study presented a neutral expression during facial photography, several studies have found variations in subtle emotional expressions in presumably neutral poses (Henderson, Holzleitner, Talamas, & Perrett, 2016; Oosterhof & Todorov, 2009). Subtle expressions of anger were previously found to increase the perception of masculinity while expressions of happiness were linked to increase femininity (Becker, Kenrick, Neuberg, Blackwell, & Smith, 2007).

To our knowledge, this is the first study to examine facial masculinity/femininity of neurotypical adults with low, mid-range or high AQ scores. The main strength of the current study is the inclusion of individuals with mid-range AQ scores that has enabled the examination of variations in facial masculinity/femininity across the broader continuum of autistic traits. Another key strength of the current study was the use of morphological analyses of 3D facial images which have shown that higher levels of autistic traits were associated with less feminine facial structures in females and less masculine facial structures in males. These findings are in line with a proposed hypothesis that physical characteristics associated with autism are related to androgyny in each sex rather than hypermasculinisation across both sexes. The sex-specific relationships reported here converge with evidence from other research literature, for instance, in neuroscience (Lai et al., 2013; Nordahl et al., 2011) and language development (Hollier et al., 2013; Whitehouse, Mattes, Maybery, Sawyer, et al., 2012) to support an emerging view that autism-related characteristics develop and manifest differently in males and females (Halladay et al., 2015; Lai, Lombardo, Auyeung, Chakrabarti, & Baron-Cohen, 2015).

There are several limitations of the current study. First, we focused on a general population sample and did not include any individuals with a clinical diagnosis of ASD. Given that the only study reporting data collected from autistic adults used a gender coherence scale (Bejerot et al. 2012) that is difficult to interpret (Tan et al., 2015), further investigation is required to examine masculinity and femininity ratings and morphological analyses of the faces of adults with ASD. Second, the present study necessarily included only Caucasian participants to limit possible extraneous influence of ethnicity on the facial ratings. Consequently, it is unclear whether our current observations extend to other ethnic populations.

Conclusion

In conclusion, the current study replicated previous findings and provided converging evidence using two methodological approaches for eliciting masculinity/femininity ratings of facial images of individuals with low, medium, and high AQ scores, as well as 3D morphological analyses. These results provide partial support for the androgyny hypothesis where females with higher levels of autistic traits presented facial features that were perceptually and structurally less feminine. More critically, males with higher levels of autistic traits presented structurally less masculine facial features, though their associations of AQ scores with perceived masculinity ratings were unclear. It is recommended that future research using the general population considers group differences across three or more bands of AQ scores rather than just two polarised samples. Additionally, it is also important to consider sex-specific variations to further our understanding of how autism-related characteristics may manifest in males and females.

Acknowledgements

****** was supported by an International Postgraduate Research Scholarship in Australia. ****** is supported by a National Health and Medical Research Council Senior Research Fellowship (Grant No.: ******). The authors would like to thank ****** for their research assistance.

Conflict of Interest

All authors declare that they have no conflict of interest.

Ethical approval

The studies involved human participants who provided informed written consent and were tested in accordance with the ethical standards of the Human Research Ethics Committee at the ****** and with Australia's 2007 National Statement on Ethical Conduct in Human Research and its later amendments.

References

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). Washington, DC: Author.

Austin, E. J. (2005). Personality correlates of the broader autism phenotype as assessed by the Autism Spectrum Quotient (AQ). *Personality and Individual Differences*, *38*(2), 451–460. https://doi.org/10.1016/j.paid.2004.04.022

Auyeung, B., Baron-Cohen, S., Ashwin, E., Knickmeyer, R., Taylor, K., & Hackett, G. (2009). Fetal testosterone and autistic traits. *British Journal of Psychology*, *100*(1), 1–22. https://doi.org/10.1348/000712608X311731

Auyeung, B., Knickmeyer, R., Ashwin, E., Taylor, K., Hackett, G., & Baron-Cohen, S. (2012). Effects of Fetal Testosterone on Visuospatial Ability. *Archives of Sexual Behavior*, *41*(3), 571–581. https://doi.org/10.1007/s10508-011-9864-8

Auyeung, B., Taylor, K., Hackett, G., & Baron-Cohen, S. (2010). Foetal testosterone and autistic traits in 18 to 24-month-old children. *Molecular Autism*, *1*(1), 11. https://doi.org/10.1186/2040-2392-1-11

Baron-Cohen, S. (2002). The extreme male brain theory of autism. *Trends in Cognitive Sciences*, *6*(6), 248–254.

Baron-Cohen, S., Auyeung, B., Nørgaard-Pedersen, B., Hougaard, D. M., Abdallah, M. W., Melgaard, L., … Lombardo, M. V. (2015). Elevated fetal steroidogenic activity in autism. *Molecular Psychiatry*, *20*(3), 369–376. https://doi.org/10.1038/mp.2014.48

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). *The Autism-Spectrum Quotient (AQ): Evidence from Asperger Syndrome/High-Functioning Autism, Males and Females, Scientists and Mathematicians*. *31*(1).

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1). https://doi.org/10.18637/jss.v067.i01

Becker, D. V., Kenrick, D. T., Neuberg, S. L., Blackwell, K. C., & Smith, D. M. (2007). The confounded nature of angry men and happy women. *Journal of Personality and Social Psychology*, *92*(2), 179–190. https://doi.org/10.1037/0022-3514.92.2.179

Bejerot, S., Eriksson, J. M., Bonde, S., Carlström, K., Humble, M. B., & Eriksson, E. (2012). The extreme male brain revisited: Gender coherence in adults with autism spectrum disorder. *British Journal of Psychiatry*, *201*(2), 116–123. https://doi.org/10.1192/bjp.bp.111.097899

Boutrus, M., Maybery, M. T., Alvares, G. A., Tan, D. W., Varcin, K. J., & Whitehouse, A. J. O. (2017). Investigating facial phenotype in autism spectrum conditions: The importance of a hypothesis driven approach. *Autism Research*, *10*(12), 1910–1918. https://doi.org/10.1002/aur.1824

Cignoni, P., Callieri, M., Corsini, M., Dellepiane, M., Ganovelli, F., & Ranzuglia, G. (2008). *MeshLab: An Open-Source Mesh Processing Tool*. 129–136.

Gilani, S. Z., Rooney, K., Shafait, F., Walters, M., & Mian, A. (2014). Geometric Facial Gender Scoring: Objectivity of Perception. *PLoS ONE*, *9*(6), e99483. https://doi.org/10.1371/journal.pone.0099483

Gilani, S. Z., Tan, D. W., Russell-Smith, S. N., Maybery, M. T., Mian, A., Eastwood, P. R., … Whitehouse, A. J. (2015). Sexually dimorphic facial features vary according to level of autistic-like traits in the general population. *Journal of Neurodevelopmental Disorders*, *7*(1). https://doi.org/10.1186/s11689-015-9109-6

Golan, O., Sinai-Gavrilov, Y., & Baron-Cohen, S. (2015). The Cambridge Mindreading Face-Voice Battery for Children (CAM-C): Complex emotion recognition in children with and without autism spectrum conditions. *Molecular Autism*, *6*(1), 22. https://doi.org/10.1186/s13229-015-0018-z

Halladay, A. K., Bishop, S., Constantino, J. N., Daniels, A. M., Koenig, K., Palmer, K., … Szatmari, P. (2015). Sex and gender differences in autism spectrum disorder: Summarizing evidence gaps and identifying emerging areas of priority. *Molecular Autism*, *6*(1). https://doi.org/10.1186/s13229-015-0019-y

Henderson, A. J., Holzleitner, I. J., Talamas, S. N., & Perrett, D. I. (2016). Perception of health from facial cues. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *371*(1693), 20150380. https://doi.org/10.1098/rstb.2015.0380

Hollier, L. P., Mattes, E., Maybery, M. T., Keelan, J. A., Hickey, M., & Whitehouse, A. J. O. (2013). The association between perinatal testosterone concentration and early vocabulary development: A prospective cohort study. *Biological Psychology*, *92*(2), 212–215. https://doi.org/10.1016/j.biopsycho.2012.10.016

Holzleitner, I. J., Hunter, D. W., Tiddeman, B. P., Seck, A., Re, D. E., & Perrett, D. I. (2014). Men's Facial Masculinity: When (Body) Size Matters. *Perception*, *43*(11), 1191–1202. https://doi.org/10.1068/p7673

Jackson, B. L., Blackwood, E. M., Blum, J., Carruthers, S. P., Nemorin, S., Pryor, B. A., … Crewther, D. P. (2013). Magno- and Parvocellular Contrast Responses in Varying Degrees of Autistic Trait. *PLoS ONE*, *8*(6), e66797. https://doi.org/10.1371/journal.pone.0066797

Knickmeyer, R., Baron-Cohen, S., Raggatt, P., Taylor, K., & Hackett, G. (2006). Fetal testosterone and empathy. *Hormones and Behavior*, *49*(3), 282–292. https://doi.org/10.1016/j.yhbeh.2005.08.010

Koehler, N., Simmons, L. W., Rhodes, G., & Peters, M. (2004). The relationship between sexual dimorphism in human faces and fluctuating asymmetry. *Proceedings of the Royal Society B: Biological Sciences*, *271*(Suppl 4), S233–S236.

Komori, M., Kawamura, S., & Ishihara, S. (2011). Multiple mechanisms in the perception of face gender: Effect of sex-irrelevant features. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(3), 626–633. https://doi.org/10.1037/a0020369

Kung, K., Spencer, D., Pasterski, V., Neufeld, S., Glover, V., O'Connor, T., … Hines, M. (2016). No relationship between prenatal androgen exposure and autistic traits: Convergent evidence from studies of children with congenital adrenal hyperplasia and of amniotic testosterone concentrations in typically developing children. *Journal of Child Psychology and Psychiatry*. https://doi.org/10.1111/jcpp.12602

Lai, M.-C., Lombardo, M. V., Auyeung, B., Chakrabarti, B., & Baron-Cohen, S. (2015). Sex/Gender Differences and Autism: Setting the Scene for Future Research. *Journal of the American Academy of Child & Adolescent Psychiatry*, *54*(1), 11–24. https://doi.org/10.1016/j.jaac.2014.10.003

Lai, M.-C., Lombardo, M. V., Suckling, J., Ruigrok, A. N. V., Chakrabarti, B., Ecker, C., … Baron-Cohen, S. (2013). Biological sex affects the neurobiology of autism. *Brain*, *136*(9), 2799–2815. https://doi.org/10.1093/brain/awt216

Le Douarin, N. M., Brito, J. M., & Creuzet, S. (2007). Role of the neural crest in face and brain development. *Brain Research Reviews*, *55*(2), 237–247. https://doi.org/10.1016/j.brainresrev.2007.06.023

Loomes, R., Hull, L., & Mandy, W. P. L. (2017). What Is the Male-to-Female Ratio in Autism Spectrum Disorder? A Systematic Review and Meta-Analysis. *Journal of the American Academy of Child & Adolescent Psychiatry*, *56*(6), 466–474. https://doi.org/10.1016/j.jaac.2017.03.013

Manning, J. T. (2011). Resolving the role of prenatal sex steroids in the development of digit ratio. *Proceedings of the National Academy of Sciences*, *108*(39), 16143–16144. https://doi.org/10.1073/pnas.1113312108

Marcucio, R. S., Young, N. M., Hu, D., & Hallgrimsson, B. (2011). Mechanisms that underlie co-variation of the brain and face. *Genesis*, *49*(4), 177–189. https://doi.org/10.1002/dvg.20710

Marečková, K., Chakravarty, M. M., Lawrence, C., Leonard, G., Perusse, D., Perron, M., … Paus, T. (2015). Identifying craniofacial features associated with prenatal exposure to androgens and testing their relationship with brain development. *Brain Structure and Function*, *220*(6), 3233–3244. https://doi.org/10.1007/s00429-014-0852-3

Meindl, K., Windhager, S., Wallner, B., & Schaefer, K. (2012). Second-to-fourth digit ratio and facial shape in boys: The lower the digit ratio, the more robust the face.

Proceedings of the Royal Society B: Biological Sciences, *279*(1737), 2457–2463. https://doi.org/10.1098/rspb.2011.2351

Mitteroecker, P., Windhager, S., Müller, G. B., & Schaefer, K. (2015). The Morphometrics of "Masculinity" in Human Faces. *PLOS ONE*, *10*(2), e0118374. https://doi.org/10.1371/journal.pone.0118374

Munson, B. (2007). The Acoustic Correlates of Perceived Masculinity, Perceived Femininity, and Perceived Sexual Orientation. *Language and Speech*, *50*(1), 125–142. https://doi.org/10.1177/00238309070500010601

Muth, A., Hönekopp, J., & Falter, C. M. (2014). Visuo-Spatial Performance in Autism: A Meta-analysis. *Journal of Autism and Developmental Disorders*, *44*(12), 3245–3263. https://doi.org/10.1007/s10803-014-2188-5

Nordahl, C. W., Lange, N., Li, D. D., Barnett, L. A., Lee, A., Buonocore, M. H., … Amaral, D. G. (2011). Brain enlargement is associated with regression in preschool-age boys with autism spectrum disorders. *Proceedings of the National Academy of Sciences*, *108*(50), 20195–20200. https://doi.org/10.1073/pnas.1107560108

Oosterhof, N. N., & Todorov, A. (2009). Shared perceptual basis of emotional expressions and trustworthiness impressions from faces. *Emotion*, *9*(1), 128–133. https://doi.org/10.1037/a0014520

Petersen, A., Crockett, L., Richards, M., & Boxer, A. (1988). A self-report measure of pubertal status: Reliability, validity, and initial norms. *Journal of Youth and Adolescence*, *17*(2), 117–133. https://doi.org/10.1007/BF01537962

R Core Team. (2017). *R: A language and environment for statistical computing.* Retrieved from https://www.R-project.org/

Robinson, E. B., Hallett, V., Happe, F., Plomin, R., & Ronald, A. (2011). Evidence That Autistic Traits Show the Same Etiology in the General Population and at the Quantitative Extremes (5%, 2.5%, and 1%). *ARCH GEN PSYCHIATRY*, *68*(11), 9.

Russell-Smith, S. N., Maybery, M. T., & Bayliss, D. M. (2011). Relationships between autistic-like and schizotypy traits: An analysis using the Autism Spectrum Quotient and Oxford-Liverpool Inventory of Feelings and Experiences. *Personality and Individual Differences*, *51*(2), 128–132. https://doi.org/10.1016/j.paid.2011.03.027

Ruzich, E., Allison, C., Smith, P., Watson, P., Auyeung, B., Ring, H., & Baron-Cohen, S. (2015). Measuring autistic traits in the general population: A systematic review of the Autism-Spectrum Quotient (AQ) in a nonclinical population sample of 6,900 typical adult males and females. *Molecular Autism*, *6*(2), 1–12. https://doi.org/10.1186/2040-2392-6-2

Scott, N. J., Jones, A. L., Kramer, R. S. S., & Ward, R. (2015). Facial Dimorphism in Autistic Quotient Scores. *Clinical Psychological Science*, *3*(2), 230–241. https://doi.org/10.1177/2167702614534238

Sutherland, C. A. M., Rhodes, G., & Young, A. W. (2017). Facial Image Manipulation: A Tool for Investigating Social Perception. *Social Psychological and Personality Science*, *8*(5), 538–551. https://doi.org/10.1177/1948550617697176

Tan, D. W., Foo, Y. Z., Downs, J., Finlay-Jones, A., Leonard, H., Licari, M., … Alvares, G. (2019). *A preliminary investigation of the effects of prenatal alcohol exposure on facial morphology in children with Autism Spectrum Disorder* [Preprint]. https://doi.org/10.31219/osf.io/xy7m6

Tan, D. W., Gilani, S. Z., Maybery, M. T., Mian, A., Hunt, A., Walters, M., & Whitehouse, A. J. O. (2017). Hypermasculinised facial morphology in boys and girls with Autism Spectrum Disorder and its association with symptomatology. *Scientific Reports*, *7*(1). https://doi.org/10.1038/s41598-017-09939-y

Tan, D. W., Russell-Smith, S. N., Simons, J. M., Maybery, M. T., Leung, D., Ng, H. L. H., & Whitehouse, A. J. O. (2015). Perceived Gender Ratings for High and Low Scorers on the Autism-Spectrum Quotient Consistent with the Extreme Male Brain Account of Autism. *PLOS ONE*, *10*(7), e0131780. https://doi.org/10.1371/journal.pone.0131780

Tapp, A. L., Maybery, M. T., & Whitehouse, A. J. O. (2011). Evaluating the twin testosterone transfer hypothesis: A review of the empirical evidence. *Hormones and Behavior*, *60*(5), 713–722. https://doi.org/10.1016/j.yhbeh.2011.08.011

Tiddeman, B, Burt, D. M., & Perrett, D. (2001). *Computer Graphics in Facial Perception Research*. *21*(5), 42–50.

Tiddeman, B, Duffy, N., & Rabey, G. (2000). Construction and visualisation of three-dimensional facial statistics. *Computer Methods and Programs in Biomedicine*, *63*(1), 9–20. https://doi.org/10.1016/S0169-2607(00)00072-9

Weinberg, S. M., Parsons, T. E., Raffensperger, Z. D., & Marazita, M. L. (2015). Prenatal sex hormones, digit ratio, and face shape in adult males. *Orthodontics & Craniofacial Research*, *18*(1), 21–26. https://doi.org/10.1111/ocr.12055

Whitehouse, A. J. O., Gilani, S. Z., Shafait, F., Mian, A., Tan, D. W., Maybery, M. T., … Eastwood, P. (2015). Prenatal testosterone exposure is related to sexually dimorphic facial morphology in adulthood. *Proceedings of the Royal Society B: Biological Sciences*, *282*(1816), 20151351. https://doi.org/10.1098/rspb.2015.1351

Whitehouse, A. J. O., Hickey, M., & Ronald, A. (2011). Are Autistic Traits in the General Population Stable across Development? *PLoS ONE*, *6*(8), e23029–e23029. https://doi.org/10.1371/journal.pone.0023029

Whitehouse, A. J. O., Mattes, E., Maybery, M. T., Dissanayake, C., Sawyer, M., Jones, R. M., … Hickey, M. (2012). Perinatal testosterone exposure and autistic-like traits in the general population: A longitudinal pregnancy-cohort study. *Journal of Neurodevelopmental Disorders*, *4*(1), 1–1. https://doi.org/10.1186/1866-1955-4-25

Whitehouse, A. J. O., Mattes, E., Maybery, M. T., Sawyer, M. G., Jacoby, P., Keelan, J. A., & Hickey, M. (2012). Sex-specific associations between umbilical cord blood testosterone levels and language delay in early childhood: Perinatal testosterone and language delay. *Journal of Child Psychology and Psychiatry*, *53*(7), 726–734. https://doi.org/10.1111/j.1469-7610.2011.02523.x

Wood, R., & Newman, S. (1995). Hormonal influence on neurons of the mating behaviour pathway in male hamsters. In *Neurobiological effects of sex steroid hormones.* Cambridge, UK: Cambridge University Press.

Zijlmans, E. A. O., Tijmstra, J., van der Ark, L. A., & Sijtsma, K. (2018). Item-Score Reliability in Empirical-Data Sets and Its Relationship With Other Item Indices. *Educational and Psychological Measurement*, *78*(6), 998–1020. https://doi.org/10.1177/0013164417728358

Tables and Figures

Table 1

Original photographs: Descriptive statistics for female and male stimulus participants' age, AQ scores, masculinity/femininity ratings, and morphological masculinity scores.

		Females			Males		
Variables	Statistics	Low-AQ	Mid-AQ	High-AQ	Low-AQ	Mid-AQ	High-AQ
	n	27	27	27	23	23	24
Age in yrs	Mean	18.6	18.9	19.1	19.3	19.3	20.1
	SD	1.31	1.74	1.34	1.85	1.80	2.09
	Range	18–24	18–25	18–22	18–25	18–25	17–25
AQ	Mean	7.85	17.1	26.2	9.25	16.5	26.3
	SD	2.11	0.85	4.47	1.80	0.78	3.87
	Range	3–10	16–18	22–43	5–12	15–18	23–36
Ratings†	Mean	5.35	5.06	4.88	5.88	5.58	5.83
	SD	0.51	0.56	0.62	0.66	0.69	0.68
	Range	4.26–7.05	3.53–6.66	3.43–6.58	4.87–8.00	4.17–7.91	4.65–8.04
Morphological masculinity scores	Mean	−0.86	−0.74	−0.53	1.00	0.77	0.70
	SD	0.73	0.53	0.56	0.71	0.62	0.68
	Range	−2.49–0.59	−1.61–0.17	−1.59–0.37	−0.01–2.46	−0.55–1.74	−0.48–1.82

Note. †Femininity ratings for female faces and masculinity ratings for male faces.

Table 2

Composite images: Descriptive statistics for female and male stimulus participants' age, AQ scores and masculinity/femininity ratings.

		Females			Males		
Variables	Statistics	Low-AQ	Mid-AQ	High-AQ	Low-AQ	Mid-AQ	High-AQ
	n	18	18	18	18	18	18
Age in yrs	Mean	18.6	18.6	19.2	19.0	19.6	19.9
	SD	1.14	0.78	1.40	1.37	1.89	2.19
	Range	18–22	18–20	18–22	18–22	18–24	17–25
AQ	Mean	8.28	16.6	26.9	8.89	16.1	24.6
	SD	2.32	1.20	4.84	1.81	0.83	2.57
	Range	3–11	15–18	22–43	5–11	15–18	22–32
Ratings†	Mean	7.42	6.87	6.77	4.74	3.83	4.30
	SD	1.32	1.27	1.23	1.49	1.65	1.70
	Range	4.20–9.67	3.93–9.37	3.65–9.03	2.47–7.90	1.43–7.33	1.73–8.40

Note. †Femininity ratings for female faces and masculinity ratings for male faces.

Figure 1

A three-dimensional ‘average’ face with 42 landmarks.

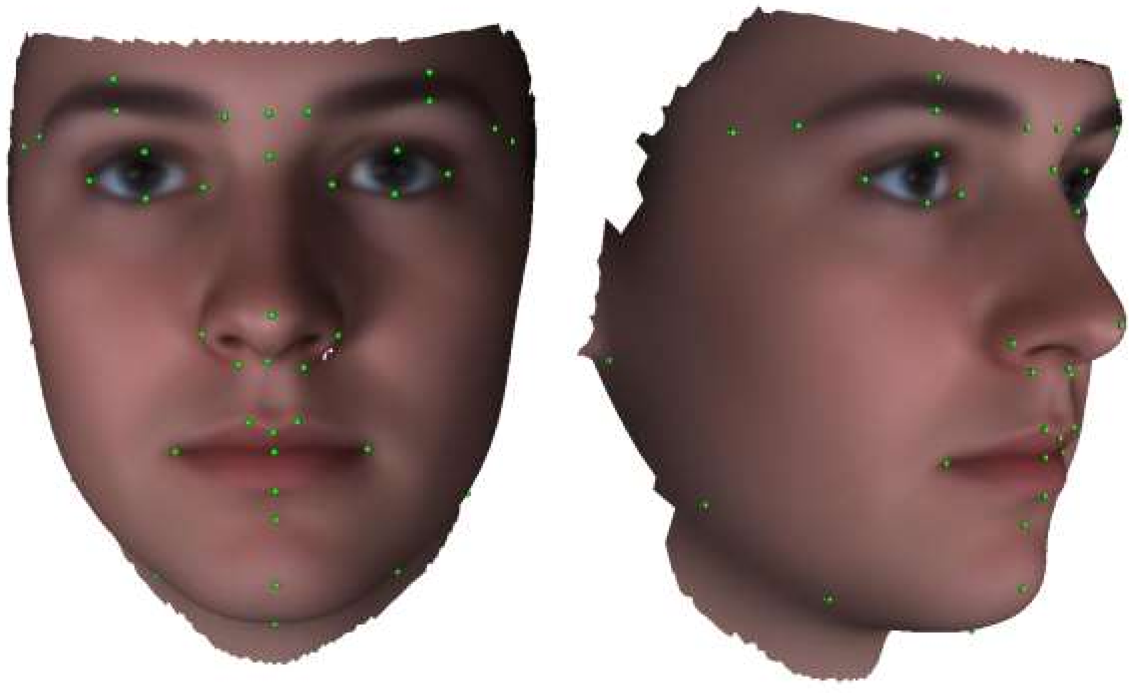

Figure 2

The variation in femininity ratings (left panel) and masculinity ratings (right panel) across AQ scores. The black dot points indicate individual ratings of the original images and the blue dot points represent the overall mean rating for each AQ score. The red line is the linear model fit and the red shaded area is the 95% CI.

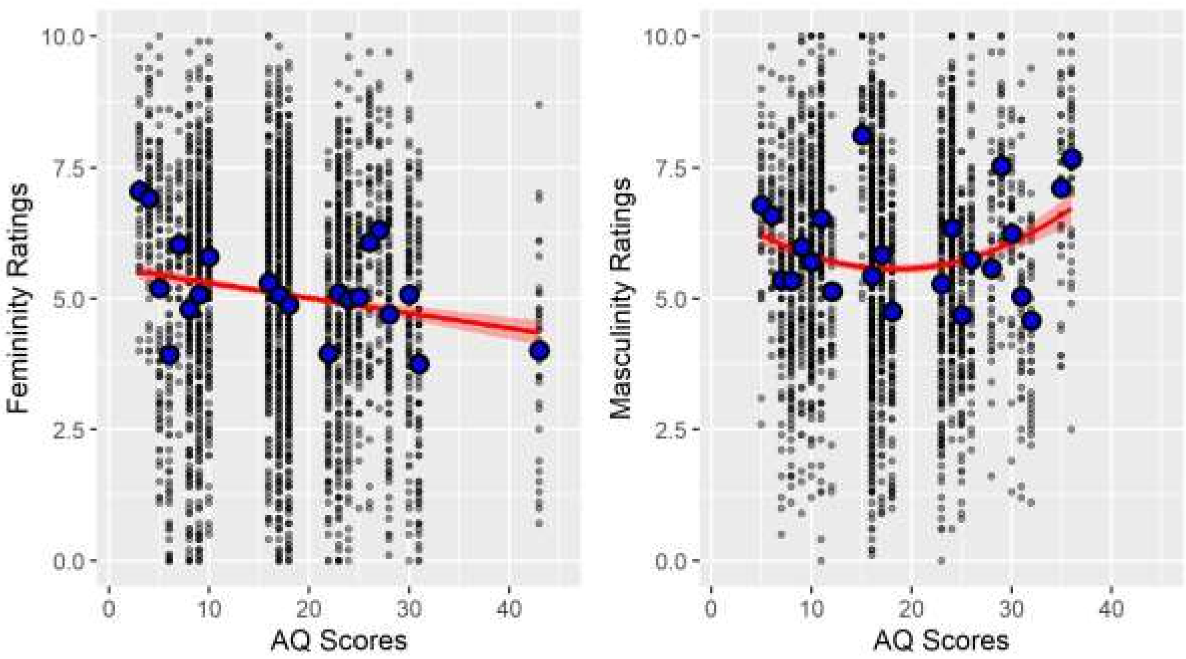

Figure 3

The variation as a function of AQ score in the morphological masculinity scores of 3D facial images for females (left panel) and males (right panel). Each black dot point indicates the facial masculinity score of each participant. The red line is the linear model fit and the red shared area is the 95% CI.

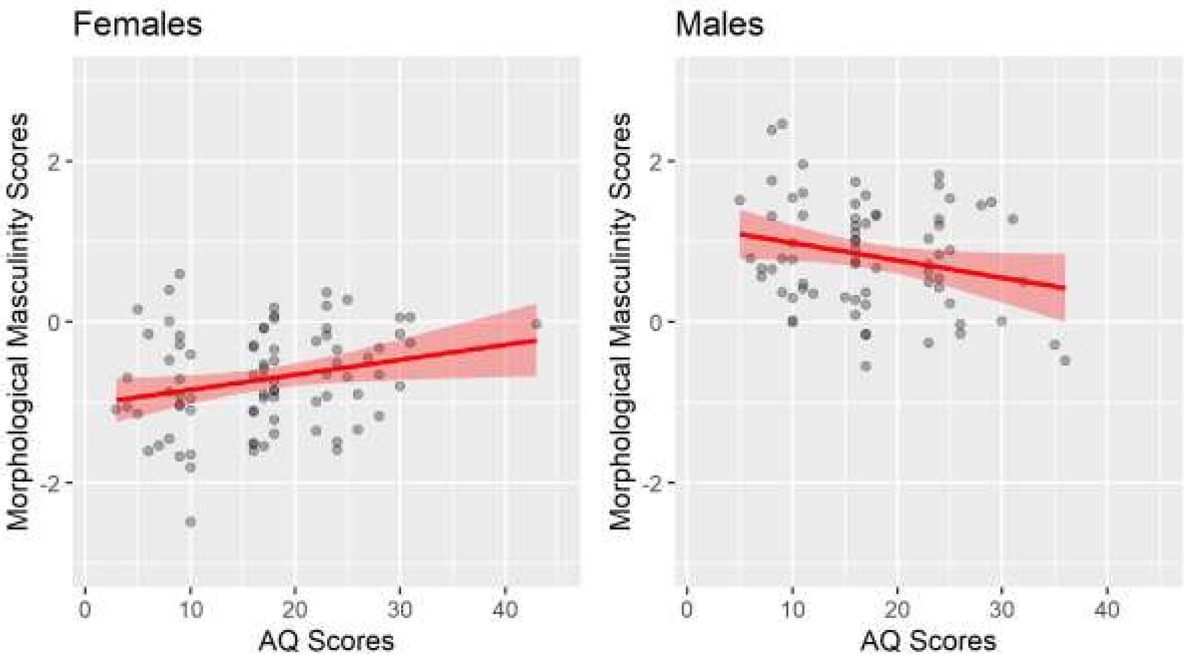

Figure 4

The variation in femininity ratings (left panel) and masculinity ratings (right panel) across the mean AQ scores of participants included in the composites. The black dot points indicate individual ratings of the composite images and the blue dot points represent the overall mean rating for each AQ score. The red line is the linear model fit, and the red shaded area is the 95% CI.