Empirical Article

Clinical Psychological Science
2015, Vol. 3(2) 230–241

DOI: 10.1177/2167702614534238
cpx.sagepub.com

Facial Dimorphism in Autistic Quotient Scores

Naomi Jane Scott[1], Alex Lee Jones[1], Robin Stewart Samuel Kramer[2], and Robert Ward[1]
[1]School of Psychology, Bangor University, and [2]School of Psychology, University of Aberdeen

Abstract
Baron-Cohen's extreme male brain theory proposes that autism results from elevated prenatal testosterone levels. In the present study, we assessed possible correlated effects of androgen exposure on adult morphology and, in particular, the development of facial features associated with masculinity. We created composite images capturing statistical regularities in facial appearance associated with high and low Autism-Spectrum Quotient (AQ) scores. In three experiments, we assessed correlations between perceived facial masculinity and AQ scores. In Experiment 1, observers selected the high-AQ males as more masculine. We replicated this result in Experiment 2, using different photographs, composite-image methods, and observers. There was no association of masculinity and AQ scores for female faces in either study. In Experiment 3, we created high- and low-AQ male composites from the five AQ subscales. High-AQ images were rated more masculine on each of the subscales. We discuss these findings with respect to the organizational-activational hypothesis of testosterone activity during development.

Keywords
autism, face perception, facial features, gender bias

Received 9/1/13; Revision accepted 4/7/14

The extreme male brain theory of autistic spectrum disorder (ASD; Baron-Cohen, 2002) assumes a continuum of individual differences in social and cognitive style that ranges from an empathizing "female" brain to a systematizing "male" brain. The empathizing end of the spectrum is represented by social cognitive skills, such as identifying the emotions and thoughts of others. The systematizing end of the scale is represented by skills in analysis of the underlying constructs in mechanical (i.e., nonsocial) systems. Baron-Cohen (2002) identified many behavioral traits that support the gender assignments of this continuum; for example, females are more likely to exhibit empathic behaviors, such as responding to the distress of others, taking turns, and showing sensitivity to the facial expressions of others. In turn, males are more likely to have good mathematical and engineering skills, pay more attention to relevant details of mechanical systems, and show a greater aptitude for construction. Therefore, within the context of a broad range of individual differences, Baron-Cohen argued that the average woman would lie more toward the empathizing end of the scale and the average man more toward the systematizing end.

Individuals with ASD have been shown to exhibit impaired empathizing abilities and high degrees of systematic behavior in various ways. Lawson, Baron-Cohen, and Wheelwright (2004) found that, as expected, females outperformed males on empathizing tasks; however, males with Asperger disorder scored even lower than did male control participants. Similarly, compared with control participants, both males and females with ASD have been shown to have lower scores on the Empathizing Quotient (Stauder, Cornet, & Ponds, 2011). The Empathizing Quotient asks participants to respond to items such as "I really enjoy caring for other people"; responses are made on a 4-point scale ranging from *strongly agree* to *strongly disagree*. Participants are given

Corresponding Author:
Naomi Jane Scott, School of Psychology, Bangor University, Adeilad Brigantia, Gwynedd LL57 2AS, Wales
E-mail: pspc5c@bangor.ac.uk

1 or 2 points for each answer that reflects high empathizing skills, thereby resulting in a possible score of 80 in which high scores represent high empathizing skills (Baron-Cohen & Wheelwright, 2004). In contrast, Baron-Cohen (2002) compiled further evidence for improved systemizing skills in individuals with increasing ASD severity, such as their preference for information that is rule based, structured, and factual; a higher likelihood of their undertaking science rather than arts degrees; and for males specifically, higher scores on the Systemizing Quotient for ASD groups compared with control groups. The Systemizing Quotient is a measure with a similar format to the Empathizing Quotient; for example, participants are asked to respond to items such as "I prefer to read non-fiction than fiction" (Baron-Cohen, Richler, Bisarya, Gurunathan, & Wheelwright, 2003).

Extreme Male Brain Theory: Prenatal Testosterone and ASD

The extreme male brain theory proposes that the highly systemizing behavior seen in ASD results from overexposure to androgens during prenatal development, which in turn produces a highly masculinized brain. This account is consistent with findings that have shown, on the basis of amniotic fluid samples, that high fetal androgen levels are associated with increased ASD occurrence (Auyeung et al., 2009; Ruta, Ingudomnukul, Taylor, Chakrabarti, & Baron-Cohen, 2011). The androgen overexposure postulated by extreme male brain theory predicts not only differences in sociocognitive style but also other effects downstream of androgen exposure. One example would be the relatively high male-to-female ratio in ASD; diagnoses are 3 to 5 times more common in males than in females (Bailey et al., 1995; Cosgrove & Riddle, 2004; Fombonne, 1999; Rodier, Bryson, & Welch, 1997). In a recently commissioned U.K. survey, in which adults were screened using a shortened version of the Autism-Spectrum Quotient (AQ; Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001), this ratio increased to 9:1 (Brugha et al., 2011). Some other effects linking testosterone-related symptoms with the occurrence of ASD have been documented. Ingudomnukul, Baron-Cohen, Wheelwright, and Knickmeyer (2007) found that females with ASD were more likely to report testosterone-related medical conditions, such as irregular menstrual cycles and an increased risk for polycystic ovary syndrome; in addition, they were more likely to be bisexual or asexual. Furthermore, females with congenital-adrenal hyperplasia, a genetic disorder that causes the overproduction of fetal androgens, exhibited an increase in masculine traits assessed by the AQ alongside a significantly higher overall score compared with control participants (Knickmeyer et al., 2006).

More generally, androgen exposure throughout the life span is associated with a variety of morphological as well as cognitive developments. In the present study, we explore whether facial masculinity is a visible phenotype correlated with autistic traits. We first examine whether there is a case that the prenatal androgen exposure assumed by extreme male brain theory may be correlated with effects on facial masculinity. We then present three studies in which we tested this correlation.

Facial Masculinity and Levels of Prenatal and Pubertal Androgens

It is well known that pubertal androgen levels affect facial masculinity. Facial masculinity in boys normally begins to develop at puberty, along with other rapid changes in body shape associated with sexual dimorphism, such as increases in body size, and this adolescent development is associated with a large increase in the androgen levels of boys (Tanner, 1989). The causal role of testosterone in facial masculinity is suggested by the fact that late-developing boys given testosterone treatment develop longer total mandibular length and anterior face height (Verdonck, Gaethofs, Carels, & de Zegher, 1999). Furthermore, facial-shape features used by observers to identify masculinity in pubertal adolescents include features correlated with circulating testosterone levels in male adolescents, such as increased breadth of the forehead, jaw, and chin (Mareckova et al., 2011).

However, the extreme male brain theory posits that it is effects of prenatal, not pubertal, androgens on brain development that promote ASD. What is less clear is whether there is an effect of prenatal androgens on facial masculinity. It is important to note that androgen levels are not stable during development but, instead, are better described by a two-stage process. According to the long-standing organizational-activational model (Phoenix, Goy, Gerall, & Young, 1959), sex hormones first establish primary sex characteristics and the "organization" of sexually dimorphic structures, including aspects of the brain, during perinatal development. Outside primary sex characteristics, this organization may not be readily apparent until "activation" during the second stage, which occurs at puberty (see Wallen, 2009, for a review of developments). It is therefore plausible that facial masculinity, like other sexual dimorphisms, such as body size (Chowen, Argente, Gonzalez-Parra, & Garcia-Segura, 1993), might be influenced by prenatal androgen levels during this organization period, even if the effect is dormant until activation during puberty.

We tested whether facial masculinity is associated with ASD, that is, whether facial masculinity is a downstream consequence of, or a correlate to, prenatal androgen exposure. A positive correlation between facial masculinity and

ASD would support a novel prediction of extreme male brain theory, thereby connecting behavior and morphology through linked hormonal effects. In addition, a positive correlation would illuminate what are currently poorly understood pathways for the effects of prenatal androgens on adult sexual dimorphism. Regardless of the mechanisms involved, identification of a visually salient phenotype relating to ASD, such as facial masculinity, may be important for better understanding the social perceptions that are linked to ASD.

Facial Morphology and ASD

Previous research has focused on the concept of facial traits associated with ASD. These studies have typically used the Waldrop scale, which lists specific abnormalities and corresponding weighted scores, to gain a standardized score for minor physical abnormalities (MPAs; Waldrop, Pedersen, & Bell, 1968). For example, Rodier et al. (1997) and Walker (1977) found a significant increase in MPA scores for children with ASD compared with control participants. In a comprehensive meta-analysis, Ozgen, Hop, Hox, Beemer, and van Engeland (2010) demonstrated a significant association between MPAs and ASD with a pooled effect size of 0.84. Ozgen et al. (2011) and Ozgen, Hellemann, de Jonge, Beemer, and van Engeland (2013) confirmed this association between ASD and MPA in subsequent, more controlled studies. These studies have not suggested that a specific anomaly or set of anomalies is diagnostic of ASD but, rather, that in the aggregate, a higher MPA score is more likely to be associated with ASD.

Other researchers have investigated specific facial features that may be diagnostic of autistic traits. Aldridge et al. (2011) found significant differences in the facial morphology of boys aged 8 to 12 years with and without autism, including increased breadth of mouth, orbits, and upper face; a flatter nasal bridge; and reduced height of the philtrum and maxillary region. Aldridge et al. suggested that as the developing face uses the brain as one foundation for craniofacial development, influences on brain development could produce corresponding, lasting impacts on facial appearance. Hammond et al. (2008) found greater facial asymmetry in boys with ASD compared with control participants, as well as the ability to discriminate, using pattern-matching algorithms, between the faces of boys with ASD and the faces of boys in the control group. These findings support the notion of a specific phenotype associated with ASD but do not link this type to perceptions of masculinity.

Finally, Bejerot et al. (2012) asked eight observers to assess the face and body of ASD and control individuals by using a 5-point coherence-typicality scale that coded for "gender typicality" at one end and "gender coherence" at the other using the following anchors: *very gender typical, gender typical, average, weak gender coherence, very weak gender coherence*. Compared with control participants, the female faces of ASD individuals were rated as significantly less gender coherent, but no significant differences were found for male faces. For our purposes, it is unfortunate that this scale entangles the typicality of traits with their dimorphism. For example, an average male face would, by statistical definition, be most typical of males (i.e., an extreme score of typicality), whereas a hypermasculine face would, again by statistical definition, be less typical of males (i.e., a less extreme typicality score). This coherence-typicality scale therefore does not seem to directly address the question of whether there is an association between autistic traits and facial dimorphism.

In the present study, we tested the specific hypothesis of association between facial masculinity and autistic traits. Because a nonclinical population was tested, autistic traits were measured using the AQ (Baron-Cohen et al., 2001; see the Measures section for a discussion of the AQ). We created two databases, each of which consisted of more than 200 participants' facial photographs and AQ scores. From each database, we selected the males and females with the highest and lowest AQ scores and made composite, or "average" images (e.g., the high-AQ male, the low-AQ female). These composites captured statistical regularities in facial appearance associated with high and low AQ while filtering out facial idiosyncrasies. We then used the composites to create stimuli that varied on facial properties associated with AQ score and measured observer responses for masculinity.

We hypothesized that correlates of prenatal testosterone may affect the masculinity of the adult face in men with high AQ scores. This prediction follows from studies, conducted with male participants, on the morphological consequences of testosterone exposure (e.g., Penton-Voak & Chen, 2004; Schaefer, Fink, Mitteroecker, Neave, & Bookstein 2005; Sisk, Schultz, & Zehr, 2003; Verdonck et al., 1999). For example, adolescent testosterone is known to affect masculinity of male faces (Mareckova et al., 2011). Predictions for female faces are much less clear because, at present, it is not known how testosterone might affect masculinity of the female face. We therefore included female faces in Experiments 1 and 2 as an exploratory comparison.

Experiment 1

Method

This experiment consisted of two phases. In Phase 1, we collected a data set of participant photographs alongside their AQ scores that we used to develop our face stimuli. In Phase 2, these stimuli were presented to observers to investigate perceptions of masculinity.

Phase 1: Stimulus creation. During this phase, we aimed to create male and female composite images that reflected the commonalities in face shape from individuals who scored in the extremes of the AQ. These images were then used to create "high-AQ" and "low-AQ" versions of individual faces.

Participants. A total of 225 Bangor University students (135 females, 90 males; mean age = 21.45 years, *SD* = 5.04) were paid £5 for their participation.

Measures. We used the AQ (Baron-Cohen et al., 2001) to assess the level of autistic traits. The AQ is not a diagnostic questionnaire; instead, it is used to identify the extent to which adults of normal intelligence show autistic traits and has been demonstrated to be a valid and reliable measure of autistic traits in an adult population. The AQ is formed from 50 items, which are split into five subscales that depict different symptom clusters: Communication, Social Skills, Attention to Detail, Attention Switching, and Imagination. Each subscale is represented by 10 items to which participants respond on a 4-point scale from *definitely agree* to *definitely disagree*. A point is given if the extreme answer (*definitely agree/definitely disagree*) correlated with autistic traits is selected, which results in a total possible questionnaire score of 50. A cutoff score of 32 has been proposed as a clinical threshold (Baron-Cohen et al., 2001). Participants' scores ranged from 3 to 41 for males (M = 17.31, SD = 6.56) and from 6 to 35 for females (M = 16.21, SD = 5.62).

Procedure. Participants were asked to remove all makeup and jewelry, tie back their hair, and adopt a neutral expression while a photograph was taken. This image was captured against a light background while the camera distance (2 m), height, flash, and zoom were kept constant. We controlled for racial characteristics and age in composite images by restricting photographs used in the study to participants who self-reported as White or White British ethnicity and were younger than 30 years old. Men with beards and moustaches were also removed, which resulted in a sample of 106 females and 48 males.

The 15 highest (female scores: M = 25.00, SD = 4.49; male scores: M = 21.47, SD = 4.16) and the 15 lowest (female scores: M = 8.33, SD = 1.35; male scores: M = 9.40, SD = 3.78) AQ scorers for each sex were then identified. Using the JPsychoMorph software (Tiddeman, Burt, & Perrett, 2001), we averaged the photographs of the individuals in these four groups to create composite images. This process resulted in four composite images: males with high AQ scores, males with low AQ scores, females with high AQ scores, and females with low AQ scores.

Next, these composites were used as anchor points to warp individual images to the two extremes they represented. Images from 20 males and 20 females who had average AQ scores (all within 1 *SD* of the mean) and who had consented to having their photo used individually were chosen at random from the data set. Each of these target faces was then warped 50% toward each of the two composite images, thereby resulting in 40 stimulus pairs: two images for each target face, which respectively represented the structural and surface features of the individuals who made up the high-AQ composite images and the individuals who made up the low-AQ composite images (see Fig. 1 for example stimuli). To create the warped images, we used a template of facial landmarks as anchor points to allow the individual images to adopt the structural morphology of the composite images, and the relevant color and texture information from each facial feature was also attributed to the new image (for a full description of this procedure, see Rowland & Perrett, 1995).

Phase 2: Perceived masculinity. In this phase, the images representing the high- and low-AQ versions of each target face were presented side by side, and observers were asked to pick the image they perceived to be more masculine.

Observers. A total of 38 observers (23 female, 15 male; mean age = 20.31 years, *SD* = 3.18) from Bangor University participated in exchange for course credit.

Procedure. On each trial, observers were presented with a stimulus pair that consisted of the same target face warped toward the high-AQ and low-AQ composite for its sex (position was counterbalanced across stimulus pairs). Observers were asked to indicate which image they thought was "more masculine in appearance" by clicking on the appropriate image with the mouse. This action then initiated the next trial. Stimulus pairs were presented in a different random order for each participant. Responses were not speeded and viewing distance was not set. Observers were unaware of the hypotheses that were being tested or the criteria used for creating the stimuli.

Results

The frequency with which observers chose the high-AQ face to be more masculine was recorded. No difference was found between male and female observers, $F(1, 36) = 0.86$, $p = .36$, so results presented are aggregated over observer sex. As shown in Figure 2a, high-AQ male faces were selected as more masculine 69% of the time, a rate significantly higher than that expected by chance, $t(37) = 6.07$, $p < .0001$, $r^2 = .49$. However, observers showed no bias for the high-AQ female faces, which were selected as

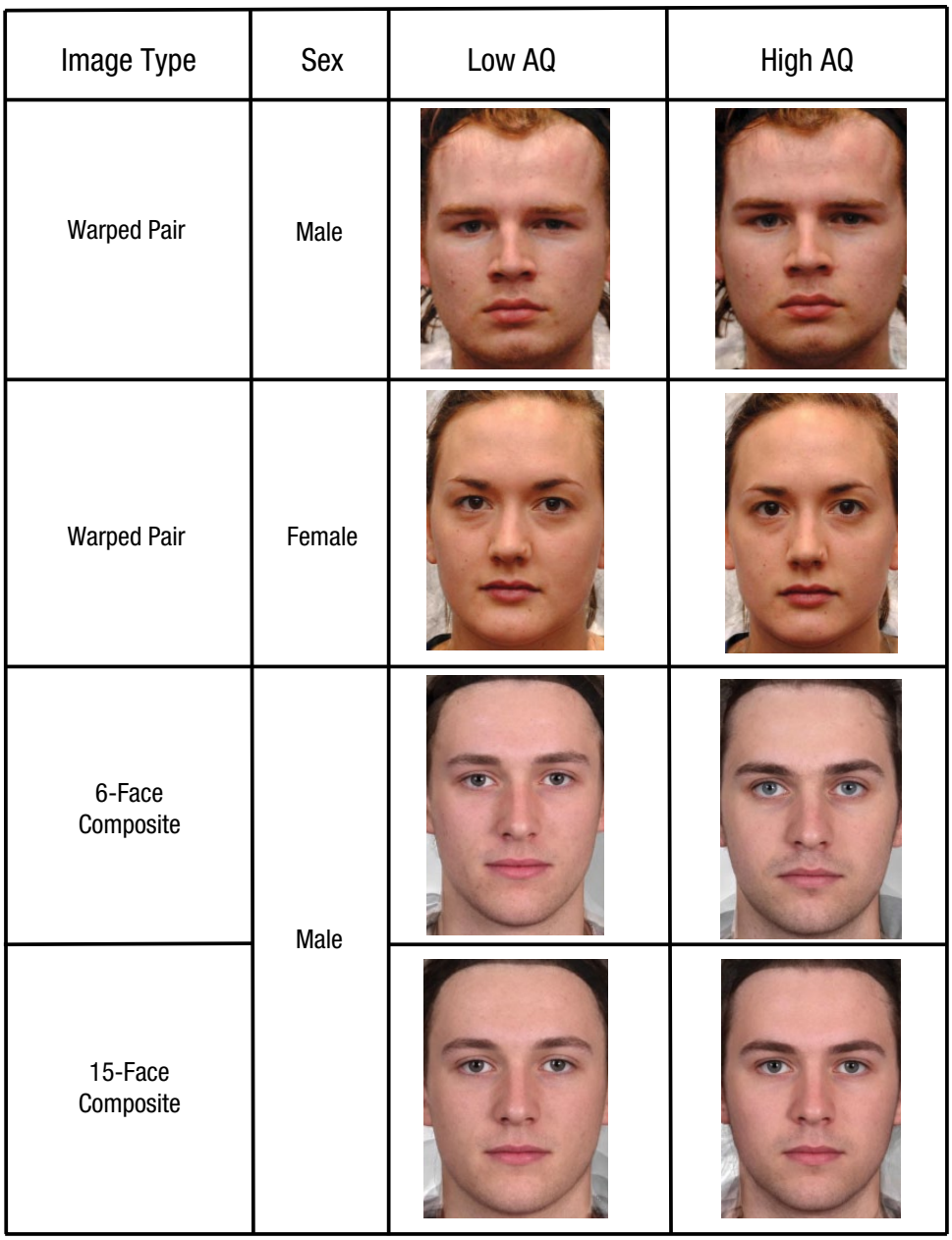

Fig. 1. Example stimuli: warped pairs presented in Experiment 1 and composite pairs presented in Experiment 2. AQ = Autism-Spectrum Quotient.

more masculine than the low-AQ female faces 47% of the time, a rate no different from chance, $t(37) = 0.90$, $p = .375$, $r^2 = .02$.

Experiment 2

Method

Results from Experiment 1 showed that high-AQ men had a more masculine facial appearance than did low-AQ men. We wanted to assess the validity and reliability of this finding. Our primary concern was whether unrepresentative faces in either AQ group might be driving our effects. We therefore tested the reliability of our results in a second experiment. Experiment 2 largely replicated the methods of Experiment 1 but with a few differences. We began by replicating the entire procedure anew; we developed a new stimulus set made from a second database of images and AQ scores, created from a different group of participants, which was then assessed

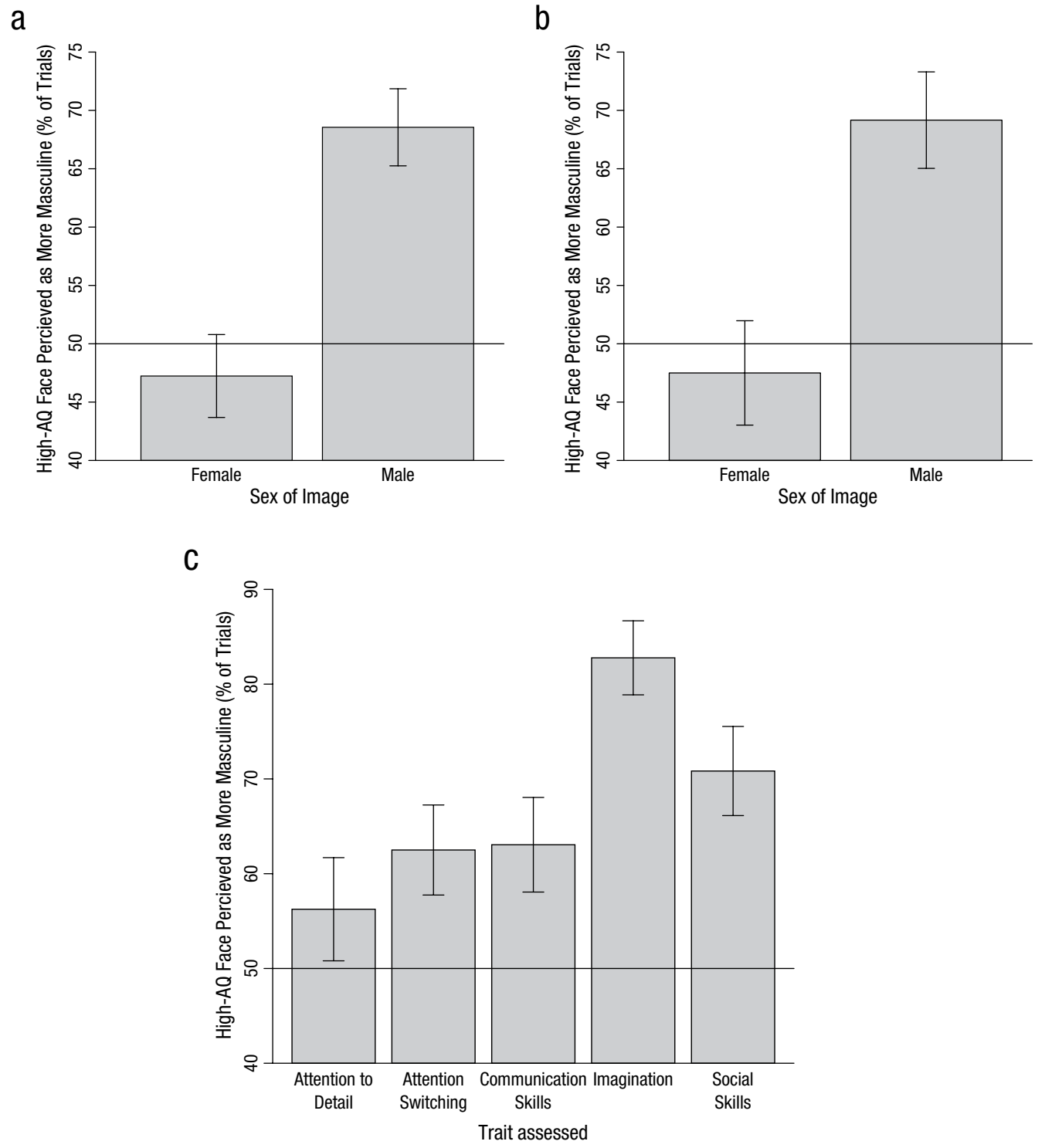

Fig. 2. Results from Experiments 1 through 3 showing percentage of trials in which high-AQ faces were perceived as more masculine. In Experiment 1 (a), high- and low-AQ warped pairs were presented, and observers indicated the more masculine image. Experiment 2 (b) replicated Experiment 1 using a new stimulus set of high and low composite pairs and new observers. In Experiment 3 (c), high- and low-AQ composites made for each of the AQ subscales were presented, and observers indicated the more masculine image. Error bars represent 95% confidence intervals. AQ = Autism-Spectrum Quotient.

by a different set of observers. Next, we changed the procedure for making the stimulus pairs. We wanted to ensure that our previous findings were not idiosyncratic to the single pair of anchors used for each sex. In Experiment 1, the high and low images (for a given sex) were warped in accordance with the differences between the single anchor pair. In Experiment 2, we increased the number of comparisons by dividing our high and low scorers into smaller subsets, thereby creating three composites for each group. We then presented observers with all combinations of the high and the low composites.

Phase 1: Stimuli development. The procedures of Experiment 1 were largely replicated except as noted in the following discussion. Changes were mainly in the techniques used to create the new stimuli.

Participants. A new set of 221 Bangor University students (130 females, 91 males; mean age = 21.65 years, *SD* = 5.09) were paid £6 for their participation.

Procedure. Photographs and AQ scores were collected as before. Participants' AQ scores were comparable with those in Experiment 1 (female: *M* = 17.30, *SD* = 6.72, range = 3–41; male: *M* = 17.74, *SD* = 5.76, range = 3–30). Again, photographs from the 15 highest (female score: *M* = 28.20, *SD* = 5.19; male score: *M* = 24.87, *SD* = 2.45) and lowest (female score: *M* = 7.53, *SD* = 2.10; male score: *M* = 10.20, *SD* = 2.60) scorers from each sex were used to create composite 15-face images, as in Experiment 1.

To ensure that results were not driven by the appearance of a few individuals used in the composite images, we created additional composites. The 18 highest (overall female score: *M* = 27.06, *SD* = 5.40; overall male score: *M* = 24.17, *SD* = 2.75) and lowest (overall female score: *M* = 7.94, *SD* = 2.13; overall male score: *M* = 10.50, *SD* = 2.62) scores for each sex were selected. These groups of 18 were randomly divided into three subsets of 6 faces each, and a 6-face composite was made from each subset. An individual appeared in only one 6-face composite, and all individuals in the 15-face composite appeared in one of the 6-face composites.

In total, there were 16 composite images. For each sex, we created 1 high and 1 low 15-face composite and 3 high and 3 low 6-face composites (see the examples in Fig. 1). These images were then assessed for masculinity in Phase 2.

Phase 2: Observer impressions of masculinity. Selection of masculine faces was assessed similarly to Experiment 1; differences are noted in the following discussion.

Observers. A new set of 48 Bangor University students (36 female, 12 male; mean age = 20.54 years, *SD* = 3.05) participated in exchange for course credit.

Procedure. The main procedure of Experiment 1 was repeated with each trial consisting of a high- and a low-AQ image presented side by side. Each possible combination of the three high-AQ and three low-AQ 6-face composites was presented once for a total of 9 trials using male faces and 9 trials using female faces. In addition, the high- and low-AQ versions of the 15-face composites were presented once with the male faces and once with the female faces. This procedure resulted in 20 trials overall.

Results

The frequency with which observers picked the high-AQ face as more masculine was recorded. Again, no differences in choices made by female and male observers were found, $F(1, 46) = 1.66$, $p = .67$, so results were aggregated over observer sex. The results were comparable with Experiment 1. As shown in Figure 2b, over all trials, observers picked the high-AQ male face as more masculine 69% of the time, a level significantly higher than chance, $t(47) = 6.61$, $p < .00001$, $r^2 = .48$. With the female faces, there was again no significant difference; the high-AQ image was selected as more masculine 48% of the time, $t(47) = 0.79$, $p = .43$, $r^2 = .01$. Results from the 6-face and 15-face composites were comparable. With the male composites, observers picked the high-AQ face as more masculine 77% of the time when viewing 6-face composites and 68% of the time when viewing the 15-face composite; both results were significantly different from chance—6-face male composites: $t(47) = 6.00$, $p < .00001$, $r^2 = .44$; 15-face male composite: $t(47) = 4.42$, $p = .000058$, $r^2 = .29$. For the 6-face female composites, the high-AQ face was picked as more masculine 47% of the time, a rate not different from chance, $t(47) = 1.03$, $p = .31$, $r^2 = .02$; for the 15-face composite, the high-AQ face was picked 54% of the time, which again was not different from chance, $t(47) = 0.57$, $p = .57$, $r^2 < .001$. In addition, results were generally consistent within the combination trials of 6-face composites; the high-AQ male image was chosen as more masculine between 66% and 73% of the time, and the high-AQ female image was chosen between 40% and 56% of the time.

Experiment 3

Method

In Experiments 1 and 2, we demonstrated a consistent association between high AQ scores and perceived facial

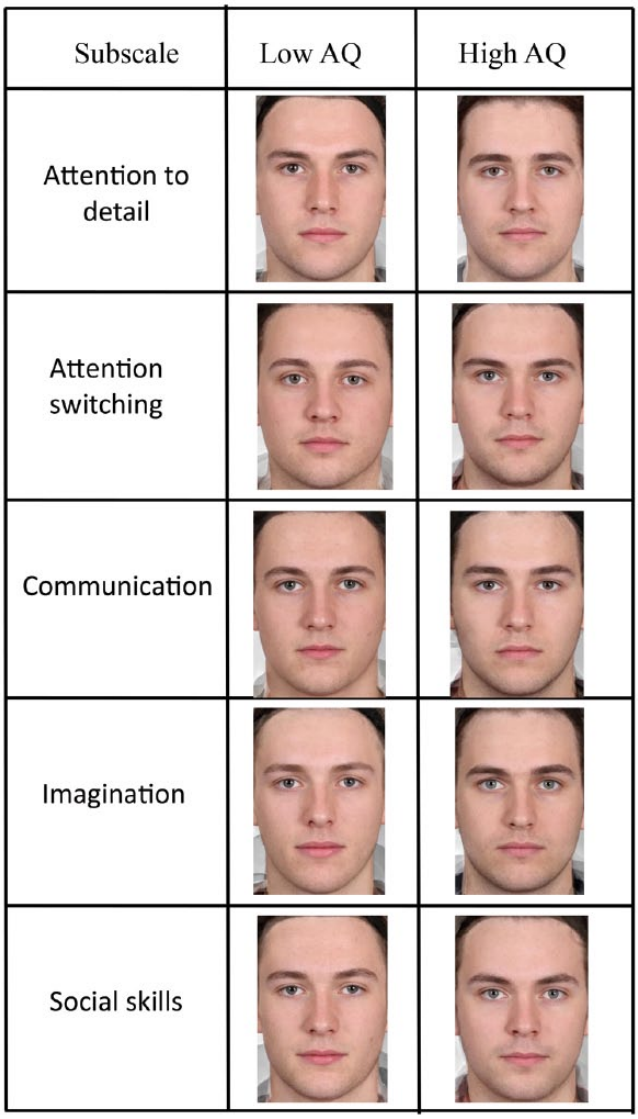

Fig. 3. Example stimuli: high- and low-AQ composite pairs presented for each of the five AQ subscales in Experiment 3. AQ = Autism-Spectrum Quotient.

masculinity in males. The AQ is composed of five subscales: Social Skills, Communication Skills, Attention Switching, Attention to Detail, and Imagination. In Experiment 3, we looked in more detail at different components of the AQ score. In particular, we wondered whether the association we had found between AQ and facial masculinity in men was identifiable in all the AQ subscales. Or was the general association with AQ driven by a specific correlation of masculinity with one of the subscales? We therefore created composite images from the high and low scorers for each of the five AQ subscales. Given that observations of female images had been at chance levels for previous studies, only male images were used in this experiment.

Phase 1: Stimuli development. Stimuli were developed from the same photo database as for Experiment 2; changes to stimuli development are noted in the following discussion.

Procedure. Photographs from the 18 highest and lowest scorers for each of the five subscales were selected: Social Skills (high AQ: $M = 4.50$, $SD = 1.86$; low AQ: $M = 0.28$, $SD = 0.46$), Attention Switching (high AQ: $M = 7.11$, $SD = 1.13$; low AQ: $M = 2.39$, $SD = 0.61$), Attention to Detail (high AQ: $M = 8.27$, $SD = 1.00$; low AQ: $M = 2.78$, $SD = 1.11$), Communication Skills (high AQ: $M = 4.78$, $SD = 1.11$; low AQ: $M = 0.78$, $SD = 0.73$), and Imagination (high AQ: $M = 3.94$, $SD = 1.26$; low AQ: $M = 0.72$, $SD = 0.46$). As in Experiment 2, these groups of 18 were then randomly divided into subgroups of 6, and 6-face composites were made from each subset. This procedure resulted in 30 composite images, examples of which are shown in Figure 3.

Phase 2: Observer impressions of masculinity. The procedure of Experiment 2 was largely replicated; differences are noted in the following discussion.

Observers. A new set of 40 observers was collected through opportunistic sampling (30 female, 10 male; mean age = 24.70 years, $SD = 4.27$).

Procedure. As in Experiment 2, each possible combination of the three high- and low-AQ images for each subscale was presented side by side to observers. This resulted in 9 trials per subscale and 45 trials in total. Once again, observers were asked to select the image they perceived to be more masculine in appearance.

Results

We found no differences in ratings given by male or female observers for any of the subscales (all ts < 0.08, all ps > .37, all r^2s < .02); thus, results were collapsed over observer sex. The frequency with which observers picked the high-AQ face as more masculine was recorded, and this frequency compared with a 50% chance rate. As shown in Figure 2c, observers consistently picked the high-AQ face as more masculine in appearance across all subscales—Attention to Detail: 59% of trials, $t(39) = 3.52$, $p = .001$, $r^2 = .24$; Attention Switching: 61% of trials, $t(39) = 4.43$, $p = .0002$, $r^2 = .34$; Communication Skills: 63% of trials, $t(39) = 3.86$, $p = .0008$, $r^2 = .28$; Imagination: 83% of trials, $t(39) = 14.14$, $p < .00001$, $r^2 = .83$; Social Skills: 71% of trials, $t(39) = 6.32$, $p < .00001$, $r^2 = .50$ (results corrected for multiple comparisons).

Post hoc analysis showed that observers rated the high-AQ image as more masculine more frequently in the

Imagination subscale than in all others (all ts > 4.98, all ps < .00001, all r^2s > .38). In addition, the high-AQ Communication Skills images were rated as more masculine with a higher frequency than were the Attention Switching, $t(39) = 3.32$, $p = .01$, $r^2 = .22$, and Attention to Detail, $t(39) = 3.34$, $p = .009$, $r^2 = .22$, images (results corrected for multiple comparisons).

Discussion

Our results are consistent with a shared hormonal influence on behavior and on appearance. In our Experiments 1 and 2, we found an association between facial masculinity and total AQ scores of men such that high AQ was associated with more masculine faces. Using different data sets, participants, and stimulus creation methods, we found very similar results across both experiments. No association was found in female faces. In Experiment 3, we found that the association of facial masculinity and AQ score in men held across all five AQ subscales. In other words, facial masculinity was correlated with (or a consequence of) the cause of high scores for each subscale. The simplest account of this pattern is to hypothesize a shared, sex-related basis for scores on each subscale. Post hoc analysis showed that the high-AQ image was rated as more masculine more often in the Imagination subscale than in any other subscale, which suggests a more robust association with masculinity, although all subscales reached highly significant levels. Consistent with extreme male brain theory, androgen exposure may be the underlying shared basis for these effects and the association of facial masculinity and AQ scores.

These experiments were motivated by findings and hypotheses concerning the role of androgens in ASD and on facial masculinity at different points in development. There is a step in the causal chain that is needed to link prenatal androgen levels, posited by extreme male brain theory as a driver for ASD, to pubertal androgen levels, a known driver for facial masculinity. Our preferred account at present is based on the organizational-activational hypothesis (Phoenix et al., 1959) and decades of nonhuman animal research that has shown that prenatal androgens can affect sensitivity to activational hormones at subsequent stages of development (e.g., Wallen, 2009). Postnatal sexual dimorphism, both in behavior (e.g., Grady, Phoenix, & Young, 1965; Phoenix et al., 1959) and in morphology (e.g., body size; Chowen et al., 1993), depends on perinatal androgens in nonhuman mammals, including primates (Thornton, Zehr, & Loose, 2009). By this account, facial masculinity in men would be a joint result of both early organizational effects of prenatal androgens and later activational effects of pubertal hormones.

The robust association we found for autistic traits and masculinity of male faces did not hold for female faces. In fact, if organizational effects of prenatal androgens were responsible for the association between autistic traits and facial appearance, it would be surprising if they had the identical long-lasting effect on female and male facial development, given the sex difference in absolute levels and the contributing factors from other hormones during development. Previous studies of the association between masculinity and testosterone have been largely or entirely limited to males (e.g., Penton-Voak & Chen, 2004; Schaefer et al., 2005; Sisk et al., 2003; Verdonck et al., 1999). Research into the correlations between prenatal, pubertal, and adult testosterone levels in humans would further understanding of such sex differences.

Our findings are consistent with those from some previous investigations on the effects of perinatal androgens in humans. These findings are based on the ratio between the lengths of the second and fourth fingers (2D:4D) as a marker of prenatal androgen levels (Manning, Scutt, Wilson, & Levis-Jones, 1998). The 2D:4D ratio has frequently been shown to be sexually dimorphic, with the ratio lower in males than in females (Manning et al., 1998). Some recent reviews have suggested that the 2D:4D ratio is a specific consequence of prenatal androgens (Hönekopp, Bartholdt, Beier, & Liebert, 2007; McIntyre, 2006), a view that is supported by observations of a stable 2D:4D dimorphism as early as 9 weeks after gestation (Malas, Dogan, Evcil, & Desdicioglu, 2006). Although the reliability of the 2D:4D ratio as a correlate of masculine traits and behaviors is not always high (Putz, Gaulin, Sporter, & McBurney, 2004), it has been previously correlated with ASD. A recent meta-analysis by Teatero and Netley (2013) showed that individuals with ASD had mean 2D:4D ratios from 0.10 to 0.77 *SD* lower than did control individuals. In addition, systemizing scores were negatively correlated with the 2D:4D ratio, whereas empathizing scores showed a positive correlation. In this context of the 2D:4D ratio as an early androgen marker, it is therefore relevant that the adult 2D:4D ratio is negatively correlated with adult facial masculinity in adult men (Neave, Laing, Fink, & Manning, 2003) and, using the same database, with facial shape (Schaefer et al., 2005). This correlation is present even in prepubertal boys, which is consistent with a possible direct influence of prenatal androgens on facial masculinity (Meindl, Windhager, Wallner, & Schaefer, 2012).

Our findings do not agree with those of Bejerot et al. (2012), who concluded that ASD was not characterized by masculinization of facial features. There are participant differences between the studies: Bejerot et al. compared individuals with ASD and control individuals, rather than high- and low-AQ individuals, and used only eight observers. However, we suggest the important difference is the

rating task for observers: We asked observers to judge sexual dimorphism, whereas Bejerot et al. asked for ratings of sex typicality-coherence. Bejerot et al. used a scale in which high sex typicality was at one extreme, low sex coherence was at the other extreme, and average typicality-coherence was in the middle. Bejerot et al. did not further elaborate on or describe the terms of their scale; however, we note that, as discussed earlier, sexual dimorphism is not the same as sex typicality, that is, certainly not in the statistical sense. High levels of sexual dimorphism are, by statistical definition, not typical, and highly typical faces for a given sex are, by statistical definition, not the most dimorphic. We therefore suggest that a typicality-coherence scale is not well suited for assessing issues relating strictly to dimorphism, given that extreme dimorphism would be rated as lower in typicality than would average dimorphism. By this interpretation, the faces of male control participants in Bejerot et al. were (nonsignificantly) more typical, not more dimorphic, than were those of male ASD participants. Similar reasoning applies to their results on body typicality-coherence. In contrast, our measure directly addresses issues relating to sexual dimorphism; we asked observers to choose the more masculine of two faces, one based on the statistical qualities of high-AQ scorers and one based on the statistical qualities of low-AQ scorers.

As highlighted by Hammond et al. (2008), the phenotype and etiology of ASD are heterogeneous, which results in problems identifying the cause of and developing effective treatments for the disorder. Hammond et al. argued that analysis of the ASD phenotype could help to identify homogenous subgroups, which could inform a more incisive analysis of the genes and mechanisms involved in its pathogenesis. Our results suggest that phenotypes related to facial masculinity may be an interesting line of exploration here. That is, facial masculinity might be useful in parceling a more homogenous subgroup from a heterogeneous population of ASD issues.

Facial masculinity in men is an important social cue, and our findings suggest a relatively unexplored possibility of a mismatch between masculine appearance and behavior in high-AQ men. Studies have produced mixed findings in terms of the social implications of perceptions of masculinity, as reviewed by Scott, Clarke, Boothroyd, and Penton-Voak (2012). Scott et al. highlighted the lack of consistent evidence for popular theories of facial masculinity, such as the immunocompetency theory, in which facial masculinity is proposed to signal high genetic fitness due to the ability to overcome the negative immunosuppressive effects of testosterone. As an alternative, Scott et al. proposed that female preferences for masculinity reflect preferences for competitive mates; that is, more masculine males are more likely to rate higher in intrasexual competition as a result of aggressiveness and dominance. These masculine traits of aggression and dominance are not necessarily characteristic of high AQ. There may therefore be a mismatch between observer expectancies, based on facial masculinity, and the actual behavior of high-AQ men. This mismatch, namely, the perception of individuals with high AQ scores as having social traits that they do not exhibit, could lead to social problems.

The results from the present study could be applied to theories further afield from the extreme male brain theory, such as the diametric hypothesis of Crespi, Stead, and Elliot (2010). Crespi et al. proposed that ASD and psychotic spectrum disorders, such as schizophrenia and bipolar disorder, form the opposing ends of a single continuum. In addition, it has been suggested that ASD is the consequence of paternal genetic imprinting, whereas psychotic spectrum disorders are caused by maternal imprinting (Crespi & Badcock, 2008). Taking into account our findings that high-AQ males are perceived to be more masculine in facial appearance, we believe that further research could be conducted to see whether individuals with schizophrenia spectrum disorders are perceived to be more feminine in appearance, thereby developing support for the combination of these theories.

The finding that high-AQ male faces had features that were consistently perceived as more masculine lends some support for the extreme male brain theory (Baron-Cohen, 2002). That is, for men, there seems to be a correlation between appearance and behavior, which could plausibly be explained by hormonal factors. However, it must be noted that a student, not a clinical, sample was used in the current study, and scores were generally below the proposed cutoff for clinically relevant traits. For this reason, our findings should be considered, first, for their potential insight into issues surrounding high AQ, including a possible mismatch of expectations and behavior regarding masculinity and, second, as a justification for further research on facial masculinity in clinical samples of ASD, which would better develop understanding into the possible social implications of observer impressions of masculinity in ASD, particularly in adults.

Author Contributions

N. J. Scott and R. Ward developed the study concept. All authors contributed to the study design. Testing and data collection were performed by N. J. Scott, A. L. Jones, and R. S. S. Kramer, and N. J. Scott performed the data analysis and interpretation under the supervision of R. Ward. N. J. Scott drafted the manuscript, and R. Ward, R. S. S. Kramer, and A. L. Jones provided critical revisions. All authors approved the final version of the manuscript for submission.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Funding

N. J. Scott is under a PhD studentship funded by the Economic and Social Research Council (Grant ES/J500197/1).

References

Aldridge, K., George, I. D., Cole, K. K., Austin, J. R., Takahashi, T. N., Duan, Y., & Miles, J. H. (2011). Facial phenotypes in subgroups of prepubertal boys with autism spectrum disorders are correlated with clinical phenotypes. *Molecular Autism*, *2*, 15. Retrieved from http://www.molecularautism.com/content/2/1/15

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

Bailey, A., Le Couteur, A., Gottesman, I., Bolton, P., Simonoff, E., Yuzda, E., & Rutter, M. (1995). Autism as a strongly genetic disorder: Evidence from a British twin study. *Psychological Medicine*, *25*, 63–77. doi:10.1017/S0033291700028099

Baron-Cohen, S. (2002). The extreme male brain theory of autism. *Trends in Cognitive Sciences*, *6*, 248–254. doi:10.1016/S1364-6613(02)01904-6

Baron-Cohen, S., Richler, J., Bisarya, D., Gurunathan, N., & Wheelwright, S. (2003). The Systemizing Quotient: An investigation of adults with Asperger syndrome or high-functioning autism, and normal sex differences. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *358*, 361–374. doi:10.1098/rstb.2002.1206

Baron-Cohen, S., & Wheelwright, S. (2004). The Empathy Quotient: An investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders*, *34*, 163–175. doi:10.1023/B:JADD.0000022607.19833.00

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. *Journal of Autism and Developmental Disorders*, *31*, 5–17. doi:10.1023/A:1005653411471

Bejerot, S., Eriksson, J. M., Bonde, S., Carlstrom, 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*, 116–123. doi:10.1192/bjp.bp.111.097899

Brugha, T. S., McManus, S., Bankart, J., Scott, F., Purdon, S., Smith, J., . . . Meltzer, H. (2011). Epidemiology of autism spectrum disorders in adults in the community in England. *Archives of General Psychiatry*, *68*, 459–466. doi:10.1001/archgenpsychiatry.2011.38

Chowen, J. A., Argenta, J., Gonzalez-Parra, S., & Garcia-Segura, L. M. (1993). Differential effects of the neonatal and adult sex steroid environments on the organization and activation of hypothalamic growth hormone-releasing hormone and somatostatin neurons. *Endocrinology*, *133*, 2792–2802. doi:10.1210/en.133.6.2792

Cosgrove, L., & Riddle, B. (2004). Gender bias and sex distribution of mental disorders in *DSM-IV-TR*. In P. J. Caplan & L. Cosgrove (Eds.), *Bias in psychiatric diagnosis* (pp. 127–140). New York, NY: Rowman & Littlefield.

Crespi, B., & Badcock, C. (2008). Psychosis and autism as diametrical disorders of the social brain [Target article and commentaries]. *Behavioral & Brain Sciences*, *31*, 241–320. doi:10.1017/S0140525X08004214

Crespi, B., Stead, P., & Elliot, M. (2010). Comparative genomics of autism and schizophrenia. *Proceedings of the National Academy of Sciences, USA*, *107*, 1736–1741. doi:10.1073/pnas.0906080106

Fombonne, E. (1999). The epidemiology of autism: A review. *Psychological Medicine*, *29*, 769–786. doi:10.1017/S0033291799008508

Grady, K. L., Phoenix, C. H., & Young, W. C. (1965). Role of the developing rat testis in differentiation of the neural tissues mediating mating behaviour. *Journal of Comparative and Physiological Psychology*, *59*, 176–182. doi:10.1037/h0021824

Hammond, P., Forster-Gibson, C., Chudler, A. E., Allanson, J. E., Hutton, T. J., Farrell, S. A., . . . Lewis, M. E. S. (2008). Face-brain asymmetry in autism spectrum disorders. *Molecular Psychiatry*, *13*, 614–623. doi:10.1038/mp.2008.18

Hönekopp, J., Bartholdt, L., Beier, L., & Liebert, A. (2007). Second to fourth digit length ratio (2D:4D) and adult sex hormone levels: New data and a meta-analytic review. *Psychoneuroendocrinology*, *32*, 313–321. doi:10.1016/j.psyneuen.2007.01.007

Ingudomnukul, E., Baron-Cohen, S., Wheelwright, S., & Knickmeyer, R. (2007). Elevated rates of testosterone-related disorders in women with autism spectrum conditions. *Hormones and Behavior*, *51*, 597–604. doi:10.1136/jech.2009.093823

Knickmeyer, R., Baron-Cohen, S., Fane, B. A., Wheelwright, S., Mathews, G. A., Conway, G. S., . . . Hines, M. (2006). Androgens and autistic traits: A study of individuals with congenital adrenal hyperplasia. *Hormones and Behavior*, *50*, 148–153. doi:10.1016/j.yhbeh.2006.02.006

Lawson, J., Baron-Cohen, S., & Wheelwright, S. (2004). Empathising and systemising in adults with and without Asperger syndrome. *Journal of Autism and Developmental Disorders*, *34*, 301–310. doi:10.1023/B:JADD.0000029552.42724.1b

Malas, M. H., Dogan, S., Evcil, E. H., & Desdicioglu, K. (2006). Fetal development of the hand, digits and digit ratio (2D:4D). *Early Human Development*, *82*, 469–475. doi:10.1016/j.earlhumdev.2005.12.002

Manning, J. T., Scutt, D., Wilson, J., & Levis-Jones, D. I. (1998). The ratio of 2nd to 4th digit length: A predictor of sperm numbers and concentrations of testosterone, luteinizing hormone and oestrogen. *Human Reproduction*, *13*, 3000–3004. doi:10.1093/humrep/13.11.3000

Mareckova, K., Weinbrand, Z., Chakravarty, M. M., Lawrence, C., Aleong, R., Leonard, G., . . . Paus, T. (2011). Testosterone-mediated sex differences in the face shape during adolescence: Subjective impressions and objective features. *Hormones and Behavior*, *60*, 681–690. doi:10.1016/j.yhbeh.2011.09.004

McIntyre, M. H. (2006). The use of digit ratios as markers for perinatal androgen action. *Reproductive Biology and Endocrinology*, *4*, 10. Retrieved from http://www.rbej.com/content/4/1/10

Meindl, K., Windhager, S., Wallner, B., & Schaefer, K. (2012). Second-to-forth 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*, 2457–2463. doi:10.1098/rspb.2011.2351

Neave, N., Laing, S., Fink, B., & Manning, J. T. (2003). Second to forth digit ratio, testosterone and perceived male dominance. *Proceedings of the Royal Society B: Biological Sciences*, *270*, 2167–2172. doi:10.1098/rspb.2003.2502

Ozgen, H., Hellemann, G. S., de Jonge, M. V., Beemer, F. A., & van Engeland, H. (2013). Predictive value of morphological features in patients with autism versus normal controls. *Journal of Autism and Developmental Disorders*, *43*, 147–155. doi:10.1007/s10803-012-1554-4

Ozgen, H. M., Hellemann, G. S., Stellato, R. K., Lahuis, B., van Daalen, E., Staal, W. G., . . . van Engeland, H. (2011). Morphological features in children with autism spectrum disorders: A matched case-control study. *Journal of Autism and Developmental Disorders*, *41*, 23–31. doi:10.1007/s10803-010-1018-7

Ozgen, H. M., Hop, J. W., Hox, J. J., Beemer, F. A., & van Engeland, H. (2010). Minor physical anomalies in autism: A meta-analysis. *Molecular Psychiatry*, *15*, 300–307. doi:10.1038/mp.2008.75

Penton-Voak, I. S., & Chen, J. Y. (2004). High salivary testosterone is linked to masculine male facial appearance in humans. *Evolution and Human Behavior*, *25*, 229–241. doi:10.1016/j.evolhumbehav.2004.04.003

Phoenix, C. H., Goy, R. W., Gerall, A. A., & Young, W. C. (1959). Organizing action of prenatally administered testosterone proportionate on the tissues mediating mating behaviour in the female guinea pig. *Endocrinology*, *65*, 369–382. doi:10.1210/endo-65-3-369

Putz, D. A., Gaulin, S. J. C., Sporter, R. J., & McBurney, D. H. (2004). Sex hormones and finger length: What does 2D:4D indicate? *Evolution and Human Behavior*, *25*, 182–199. doi:10.1016/j.evolhumbehav.2004.03.005

Rodier, P. M., Bryson, S. E., & Welch, J. P. (1997). Minor malformations and physical measurements in autism: Data from Nova Scotia. *Teratology*, *55*, 319–325. doi:10.1002/(SICI)1096-9926(199705)55:5<319::AID-TERA4>3.3.CO;2-K

Rowland, D. A., & Perrett, D. I. (1995). Manipulating face appearance through shape and color. *IEEE Computer Graphics and Applications*, *15*, 70–76. doi:10.1109/38.403830

Ruta, L., Ingudomnukul, E., Taylor, K., Chakrabarti, B., & Baron-Cohen, S. (2011). Increased serum androstenedione in adults with autism spectrum conditions. *Psychoneuroendocrinology*, *36*, 1154–1163. doi:10.1016/j.psyneuen.2011.02.007

Schaefer, K., Fink, B., Mitteroecker, P., Neave, N., & Bookstein, F. L. (2005). Visualizing facial shape regression upon 2nd to 4th digit ratio and testosterone. *Collegium Antropologicum*, *29*, 415–419.

Scott, I. M. L., Clarke, A. P., Boothroyd, L. G., & Penton-Voak, I. S. (2012). Do men's faces really signal heritable immunocompetence? *Behavioral Ecology*, *24*, 579–589. doi:10.1093/beheco/ars092

Sisk, C. L., Schulz, K. M., & Zehr, J. L. (2003). Puberty: A finishing school for male social behaviour. *Annals of the New York Academy of Sciences*, *1007*, 189–198.

Stauder, J. E. A., Cornet, L. J. M., & Ponds, R. W. H. M. (2011). The extreme male brain theory and gender role behavior in persons with an autism spectrum condition. *Research in Autism Spectrum Disorders*, *5*, 1209–1214. doi:10.1016/j.rasd.2011.01.008

Tanner, J. M. (1989). *Foetus into man: Physical growth from conception to maturity*. Hertfordshire, England: Castlemead.

Teatero, M. L., & Netley, C. (2013). A critical review of the research on the extreme male brain theory and digit ratio (2D:4D). *Journal of Autism and Developmental Disorders*, *43*, 2664–2676. doi:10.1007/s10803-013-1819-6

Thornton, J., Zehr, J. L., & Loose, M. D. (2009). Effects of prenatal androgens on rhesus monkeys: A model system to explore the organizational hypothesis in primates. *Hormones and Behavior*, *55*, 633–644. doi:10.1016/j.yhbeh.2009.03.015

Tiddeman, B., Burt, M., & Perrett, D. (2001). Prototyping and transforming facial textures for perception research. *IEEE Computer Graphics and Applications*, *21*, 42–50. doi:10.1109/38.946630

Verdonck, A., Gaethofs, M., Carels, C., & de Zegher, F. (1999). Effect of low-dose testosterone treatment on craniofacial growth in boys with delayed puberty. *European Journal of Orthodontics*, *21*, 137–143. doi:10.1093/ejo/21.2.137

Waldrop, M. F., Pedersen, F. A., & Bell, R. Q. (1968). Minor physical anomalies and behaviour in preschool children. *Child Development*, *39*, 391–400. doi:10.2307/1126953

Walker, H. A. (1977). Incidence of minor physical anomaly in autism. *Journal of Autism and Childhood Schizophrenia*, *7*, 165–176. doi:10.1007/BF01537727

Wallen, K. (2009). The organizational hypothesis: Reflections on the 50th anniversary of the publication of Phoenix, Goy, Gerall and Young (1959). *Hormones and Behavior*, *55*, 561– 565. doi:10.1016/j.yhbeh.2009.03.009