British Journal of Psychology (2009), 100, 1–22

The British Psychological Society

www.bpsjournals.co.uk

Fetal testosterone and autistic traits

Bonnie Auyeung[1]*, Simon Baron-Cohen[1], Emma Ashwin[1], Rebecca Knickmeyer[1,2], Kevin Taylor[3] and Gerald Hackett[4]

[1]Autism Research Centre, Department of Psychiatry, University of Cambridge, Cambridge, UK
[2]Department of Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, USA
[3]Department of Clinical Biochemistry, Addenbrooke's Hospital, Cambridge, UK
[4]Department of Foetal Medicine, Rosie Maternity Hospital, Cambridge, UK

Studies of amniotic testosterone in humans suggest that fetal testosterone (fT) is related to specific (but not all) sexually dimorphic aspects of cognition and behaviour. It has also been suggested that autism may be an extreme manifestation of some male-typical traits, both in terms of cognition and neuroanatomy. In this paper, we examine the possibility of a link between autistic traits and fT levels measured in amniotic fluid during routine amniocentesis. Two instruments measuring number of autistic traits (the Childhood Autism Spectrum Test (CAST) and the Child Autism Spectrum Quotient (AQ-Child)) were completed by these women about their children ($N = 235$), ages 6–10 years. Intelligence Quotient (IQ) was measured in a subset of these children ($N = 74$). fT levels were positively associated with higher scores on the CAST and AQ-Child. This relationship was seen within sex as well as when the sexes were combined, suggesting this is an effect of fT rather than of sex *per se*. No relationships were found between overall IQ and the predictor variables, or between IQ and CAST or AQ-Child. These findings are consistent with the hypothesis that prenatal androgen exposure is related to children exhibiting more autistic traits. These results need to be followed up in a much larger sample to test if clinical cases of ASC have elevated fT.

Many clinical conditions occur in males more often than females, including autism, dyslexia, specific language impairment, attention-deficit hyperactivity disorder (ADHD), and early onset persistent antisocial behaviour (Rutter, Caspi, & Moffitt, 2003). Depression, anorexia, and the anxiety disorders do not show this male bias in sex ratio, raising the question of whether there are sex-linked or sex-limiting factors involved in the aetiology of those conditions that do exhibit this male bias. Autism in particular has been described as an extreme manifestation of certain sexually dimorphic

* *Correspondence should be addressed to Bonnie Auyeung, Department of Psychiatry, Autism Research Centre, Cambridge, Cambridgeshire CB2 8AH, UK (e-mail: ba251@cam.ac.uk).*

DOI:10.1348/000712608X311731

traits or as a consequence of an 'extreme male brain' (EMB; Baron-Cohen, 2002). Autism, high-functioning autism, Asperger syndrome, and pervasive developmental disorder (not otherwise specified; PDD/NOS) will be referred to as autism spectrum conditions (ASC), and are considered to lie on the same continuum. Individuals with an ASC diagnosis are impaired in reciprocal social interaction and communication, and show strongly repetitive behaviours and unusually narrow interests (APA, 1994). A recent epidemiological study in the UK reports that 1% of people have a diagnosis of ASC (Baird *et al.*, 2006).

The strong bias of ASC towards males has been well established (Bryson & Smith, 1998; Fombonne, 2005; Tidmarsh & Volkmar, 2003). Children with ASC have a sex ratio of 4:1 (male:female) across the full Intelligence Quotient (IQ) range (Chakrabarti & Fombonne, 2005) and the ratio is as high as 9:1 for Asperger syndrome (Scott, Baron-Cohen, Bolton, & Brayne, 2002a; Wing, 1981), the subgroup in which individuals have intact IQ and language development. The strikingly higher incidence of ASC in males might provide important clues to the aetiology of the condition (Baron-Cohen, 2002; Baron-Cohen, Knickmeyer, & Belmonte, 2005). ASC have a strong neurobiological and genetic component (Stodgell, Ingram, & Hyman, 2001), however the specific factors (hormonal, genetic, or environmental) that are responsible for the higher male incidence in ASC are unclear. Recent evidence is consistent with the idea that male sex hormones, and in particular, prenatal exposure to testosterone may be related to the development of autistic traits (Baron-Cohen, Lutchmaya, & Knickmeyer, 2004).

It has been proposed that some observable sex differences in human behaviour and cognition may be accounted for by sex differences in cerebral lateralization (Hines & Shipley, 1984). Geschwind and Galaburda (1985) hypothesized that fetal testosterone (fT) exposure facilitates the growth of certain areas in the right hemisphere of the brain while simultaneously inhibiting the growth of the same areas in the left hemisphere of the brain (Geschwind & Galaburda, 1985). The Geschwind and Galaburda hypothesis was questioned because of a lack of supportive evidence (Bryden, McManus, & Bulman-Fleming, 1994), though it should be noted that the majority of studies which aimed to test the hypothesis did not directly test whether fT affects sexually dimorphic neurobiology. Instead studies used indirect tests, such as examining associations between left-handedness and strongly sex dependent conditions like autism.

In support of the Geschwind and Galaburda hypothesis, studies investigating body asymmetry found that left-handedness and asymmetrical lateralization were associated with both being male, and with autism (Fein, Waterhouse, Lucci, Pennington, & Humes, 1985; McManus, Murray, Doyle, & Baron-Cohen, 1992; Satz, Soper, Orsini, Henry, & Zvi, 1985; Soper *et al.*, 1986). Analogies can also be drawn between sex differences in brain development and neuroanatomical characteristics found in autism. The typical male brain is heavier than the female brain and individuals with autism have heavier brains than typical males (Harden, Minshew, Mallikarjuhn, & Keshavan, 2001). The amygdala is also disproportionately large in boys compared to girls (Giedd *et al.*, 1996) and children with autism have enlarged amygdalae (Hazlett *et al.*, 2005). These findings are consistent with the EMB theory.

The ratio between the length of the 2nd and 4th digit (2D:4D) is sexually dimorphic, being lower in males than in females, and may be a useful proxy measure for fT production in humans during the first trimester of gestation (Manning, Scutt, Wilson, & Lewis-Jones, 1998). Studies researching fetal hand development have observed the sex difference in 2D:4D ratio in fetuses between 9 and 40 weeks of gestation (Malas, Dogan, Evcil, & Desdicioglu, 2006). The 2D:4D ratio has been found to be negatively associated

with the ratio of fT to fetal oestrogen (Lutchmaya, Baron-Cohen, Raggatt, Knickmeyer, & Manning, 2004). Lower (i.e. hyper-masculinized) digit ratios have been found in children with autism compared to typically developing children. This pattern was also found in the siblings and parents of children with autism, suggesting genetically based elevated fT levels in autism (Manning, Baron-Cohen, Wheelwright, & Sanders, 2001; Milne *et al.*, 2006). If 2D:4D ratio does reflect prenatal exposure to testosterone, this evidence suggests children with ASC may have been exposed to higher than average levels of fT.

The direct manipulation of fT levels is not possible in humans for ethical reasons. The investigation of prenatal hormone exposure and its relation to development in humans has therefore been investigated in naturally occurring abnormal environments such as in individuals with congenital adrenal hyperplasia (CAH) which is a genetic disorder that causes excess adrenal androgen production beginning prenatally in both males and females (New, 1998). Studies of individuals with CAH have generally found that girls with CAH show masculinization of performance in activities typically dominated by males such as spatial orientation, visualization, targeting, personality, cognitive abilities, and sexuality (Hampson, Rovet, & Altmann, 1998; Hines *et al.*, 2003; Resnick, Berenbaum, Gottesman, & Bouchard, 1986). Results from one study of girls with CAH suggest that they exhibit more autistic traits, measured using the adult version of the Autism Spectrum Quotient (AQ), compared to their unaffected sisters (Knickmeyer, Fane *et al.*, 2006). Whilst CAH provides an interesting opportunity to investigate the effects of additional androgen exposure, the relatively rare occurrence of CAH in conjunction with ASC makes it difficult to obtain large enough sample sizes for generalization of research findings to the wider population. Researchers have also suggested that CAH-related disease characteristics, rather than prenatal androgen exposure, could be responsible for the atypical cognitive profiles observed in this population (Fausto-Sterling, 1992; Quadagno, Briscoe, & Quadagno, 1977).

The relationship between fT exposure and postnatal development have also been examined using measures of fT levels in amniotic fluid, obtained during amniocentesis performed for other clinical reasons. In animal models, the critical period for sexual differentiation of the brain occurs when differences in serum testosterone are highest between sexes (Smith & Hines, 2000). The human fT surge is thought to occur between weeks 8 and 24 of gestation (Baron-Cohen *et al.*, 2004; Collaer & Hines, 1995; Hines, 2004). A major advantage of amniocentesis for examining hormone–behaviour relations is that it is typically performed during the second trimester of pregnancy, during a relatively narrow window of time (usually 14–20 weeks of gestation) that coincides with the serum testosterone peak period in male fetuses.

Although the origins of hormones (and in particular androgens) found in amniotic fluid are not fully understood, the main source seems to be the fetus itself (Cohen-Bendahan, van de Beek, & Berenbaum, 2005). Hormones can enter the amniotic fluid in two ways: via diffusion through the fetal skin in early pregnancy and via fetal urine in later pregnancy (Judd, Robinson, Young, & Jones, 1976; Schindler, 1982). Thus, testosterone levels measured in amniotic fluid might be expected to be a good reflection of the levels in the fetus, providing an alternative to direct assay of fetal serum (Finegan, Bartleman, & Wong, 1991). A limitation is that amniocentesis is an invasive procedure conducted only when fetal anomalies are suspected (such as chromosomal trisomy 21 (Down syndrome)). This invasive nature and the small risk of miscarriage that amniocentesis entails, means that samples cannot be specifically collected for research purposes. As a result, women who undergo amniocentesis (about 6% of pregnancies) are not a random selection of all pregnant women. They tend to be above average in age

(because maternal age is a risk factor for having a child with Down syndrome) and above average in education, which may limit generalisability. Nevertheless, amniocentesis itself does not appear to have any negative effects on offspring behaviour (Finegan *et al.*, 1990). Maternal age can be studied as a variable in its own right (since women may have an amniocentesis at any age), as can maternal educational level (since this is measurable and can be entered as a covariate in any later analysis). Furthermore, studies can opt to include only those pregnancies that were normal and where the baby was born healthy and with no cytogenetic abnormalities. As such, the measurement of testosterone levels in amniotic fluid presents a unique opportunity for exploration of the fetal hormonal environment, and arguably the only ethical option for conducting such research in humans.

The EMB theory is an extension of the empathizing–systemizing (E–S) theory of typical psychological sex differences (Baron-Cohen, 2002, 2003) which proposes that females on average have a stronger drive to empathize (to identify another person's emotions and thoughts, and to respond to these with an appropriate emotion), while males have a stronger drive to systemize (to analyse or construct rule-based systems, whether mechanical, abstract, natural, etc.) (Baron-Cohen, 2003; Baron-Cohen *et al.*, 2005). Evidence has been reported showing a female advantage in empathy and a male advantage in systemizing (Baron-Cohen, O'Riordan, Stone, Jones, & Plaisted, 1999; Baron-Cohen, Wheelwright, & Hill, 2001; Davis, 1994; Lawrence, Shaw, Baker, Baron-Cohen, & David, 2004). Consistent with the EMB theory, individuals with ASC have been shown to have a stronger drive to systemize (Baron-Cohen, Richler, Bisarya, Gurunathan, & Wheelwright, 2003), but show impairment on tests of empathizing (Baron-Cohen *et al.*, 1999, 2001). It has been found that in typically developing children whose mothers had undergone amniocentesis, fT levels show a positive association with systemizing (Auyeung *et al.*, 2006), and a negative association with empathizing (Chapman *et al.*, 2006; Knickmeyer, Baron-Cohen, Raggatt, Taylor, & Hackett, 2006).

Empathy and systemizing are difficult to measure in infancy. Developmental precursors of empathy include social interest, indexed by behaviours such as eye-contact and attention to faces (Baron-Cohen, 1995; Johnson, Cheek, & Smither, 1983) and language development (Baron-Cohen, Baldwin, & Crowson, 1997). Developmental precursors to systemizing include attention to detail and narrow interests (Baron-Cohen, 2002). Reduced eye-contact, delayed language development, social difficulties and narrow interests are also characteristic of children with ASC (APA, 1994; Lutchmaya, Baron-Cohen, & Raggatt, 2002a; Rutter, 1978; Swettenham *et al.*, 1998). Results from the Cambridge fT Project showed that in typically developing children whose mothers had undergone amniocentesis, fT was inversely associated with frequency of eye-contact in males at 12 months old (Lutchmaya *et al.*, 2002a), and inversely predicted vocabulary development in children between ages 18 and 24 months (Lutchmaya, Baron-Cohen, & Raggatt, 2002b). At 4 years of age, high levels of fT were associated with poorer quality of social relationships and more narrow interests (Knickmeyer, Baron-Cohen, Raggatt, & Taylor, 2005). At 8 years old, fT was positively correlated with the child version of the systemizing Quotient (Auyeung *et al.*, 2006) and negatively correlated with the child version of the Empathy Quotient (Chapman *et al.*, 2006). These studies provide evidence that prenatal fT levels are associated with individual differences in sexually dimorphic behavioural traits relevant to the features of ASC.

In this paper, we report the most recent study from the Cambridge fT project, an ongoing longitudinal study investigating the relationship between fT levels and the

development of behaviours relating to ASC (Baron-Cohen *et al.*, 2004; Knickmeyer & Baron-Cohen, 2006). This is a unique longitudinal study, having measured fT levels in amniotic fluid for a sample of children whose behaviour has been examined postnatally at 1, 2, 4, and 8 years of age, using a variety of age appropriate tests. This project is ongoing and it is intended to follow these children through puberty and into adulthood. The key question behind this project is whether there is any relationship between individual variation in fT levels and later phenotypic traits.

The aim of the present study is to examine the relationship between autistic traits and fT exposure in this cohort of typically developing children who are between 6 and 10 years old. Two measures of autistic traits were employed: the Childhood Autism Spectrum Test (CAST; Scott, Baron-Cohen, Bolton, & Brayne, 2002b; Williams *et al.*, 2005), (formerly known as the Childhood Asperger Syndrome Test, renamed because it can be used for all subgroups on the autistic spectrum (Baron-Cohen *et al.*, 2008) and the Child Autism Spectrum Quotient (AQ-Child) (Auyeung, Baron-Cohen, Wheelwright, & Allison, in press). The CAST was administered because it is an established measure of autistic traits that has been validated on a large population in the UK (Williams *et al.*, 2005) and has been shown to be heritable (Ronald *et al.*, 2006). The AQ-Child was administered because research suggests that the AQ is a good measure of autistic traits across the life-span. The AQ also predicts clinical diagnosis in adults (Woodbury-Smith, Robinson, Wheelwright, & Baron-Cohen, 2005) and shows strong cross-cultural consistency (Hoekstra, Bartels, Cath, & Boomsma, in press; Wakabayashi, Baron-Cohen, Wheelwright, & Tojo, 2006) and substantial heritability in the general population (Hoekstra, Bartels, Verweij, & Boomsma, 2007). We also investigate the relationship between IQ and fT level in a subset of this typically developing sample, in order to explore if fT is related to general cognitive ability.

Methods

Participants

Participants were recruited from a longitudinal study on the effects of fT (Baron-Cohen *et al.*, 2004). Medical records of 950 patients who had undergone amniocentesis in the Cambridge (UK) region between 1996 and 2001 were examined. Women were excluded if: (a) the amniocentesis revealed a chromosomal abnormality; (b) there was a twin pregnancy; (c) the pregnancy ended in miscarriage, termination, or significant medical problems after birth; (d) relevant information was absent from medical records; or (e) medical practitioners indicated that contacting the family would be inappropriate.

The AQ-Child and CAST were sent to all mothers meeting inclusion criteria, resulting in 452 mothers contacted; 261 mothers completed the CAST; and 248 mothers completed the AQ-Child, resulting in a total of 235 (118 boys, 117 girls) children with complete data for both questionnaires. IQ (measured using the WASI (Wechsler, 1999)) was completed for a subgroup of these children ($N = 74$; 43 boys, 31 girls), whose mothers consented to bring them in for cognitive assessment.

Outcome variables

The Child Autism Spectrum Quotient

This is a 50-item parent-report questionnaire developed to detect autistic traits in children 4–11 years of age (Auyeung *et al.*, in press). Higher scores indicate a greater number of autistic traits. AQ-Child items are answered in a Likert format (definitely

agree, slightly agree, slightly disagree, and definitely disagree). The AQ-Child has shown good test–retest reliability ($r = .85$, $p < .001$), high sensitivity (95%) and high specificity (95%) (Auyeung *et al.*, in press). The AQ-Child was originally designed with five subscales to assess various domains of functioning: social skills; attention to detail; attention switching; communication; and imagination. Each domain is assessed by 10 questions, giving a maximum attainable score of 150. Principal components analysis suggests the AQ-Child has four empirically derived subscales: mind reading (16 items); attention to detail (9 items); social skills (15 items); and imagination (7 items), resulting in a new maximum score of 141. A score of 76 or above indicates a risk for ASC. These four subscales were found to be respectively highly correlated with the original AQ subscales of communication ($r = .97$, $p < .001$), attention to detail ($r = .95$, $p < .001$), social skills ($r = .97$, $p < .001$), and imagination ($r = .97$, $p < .001$) (Auyeung *et al.*, in press). Cronbach's α was calculated for total score, and results for this sample demonstrated good internal consistency ($\alpha = .91$). The internal consistency of the AQ-Child subscales were also satisfactory (mind reading = 0.80; attention to detail = 0.80; social skills = 0.87; and imagination = 0.75).

The Childhood Autism Spectrum Test

This 37-item parent-report questionnaire was developed to detect ASC in 4–11 years old children (Scott *et al.*, 2002b). CAST items require a binary response ('yes/no') to 37 questions, 31 of which are scored (maximum score of 31). A validation study suggested that a score of 15 or above should be used to indicate risk for ASC (Scott *et al.*, 2002b; Williams *et al.*, 2005). The CAST has good test–retest reliability (Spearman's $\rho = .83$, $p < .001$) (Williams *et al.*, 2006), good positive predictive value (50%) and high specificity (97%) and sensitivity (100%) for ASC (Williams *et al.*, 2005). For the current sample, Cronbach's α showed acceptable internal consistency ($\alpha = .85$).

The Wechsler Abbreviated Scale of Intelligence (WASI)

This scale was used to measure IQ (Wechsler, 1999). The WASI provides scores for verbal IQ, performance IQ, and full scale IQ. We also examined the relationship between fT and the block design component of the WASI. Block design performance shows sexual dimorphism in adulthood (male advantage) (Lynn, 1998; Lynn, Raine, Venables, Mednick, & Irwing, 2005; Rönnlund & Nilsson, 2006) and individuals with ASC show superior performance on this subtest (Happe, 1994; Shah & Frith, 1993).

Predictor variables

fT levels

The major predictor in this study is fT level in amniotic fluid, measured by radioimmunoassay. Amniotic fluid was extracted with diethyl ether. The ether was evaporated to dryness at room temperature and the extracted material re-dissolved in an assay buffer. Testosterone was assayed by the DPC 'Count-a-Coat' method (Diagnostic Products Corp, Los Angeles, CA 90045-5597, USA), which uses an antibody to testosterone coated on to propylene tubes and a 125-I labelled testosterone analogue. The detection limit of the assay using the ether-extraction method is approximately 0.05 nmol/L. The coefficient of variation (CV) for between batch imprecision is 19% at a concentration of 0.8 nmol/L and 9.5% at a concentration of 7.3 nmol/L. The CV's for within batch imprecision are 15% at a concentration of 0.3 nmol/L and 5.9% at a concentration of 2.5 nmol/L. This method measures total extractable testosterone.

The following control variables were also included in our study:

Gestational age at amniocentesis

The amniocentesis procedure generally occurs between weeks 14 and 22. Therefore it is important to determine whether fT is related to gestational age. No significant relationships were found between fT levels and gestational age when both sexes were combined ($r = -.03, p > .05$), as well as when examining boys ($r = -.10, p > .05$) and girls ($r = .07, p > .05$) separately.

Maternal age

Maternal age was included because women undergoing amniocentesis have a higher mean age than the general child-bearing population.

Level of education obtained by the parents

Parental education level was measured according to a five-point scale: 1 = no formal qualifications; 2 = typical 16-year-old qualification; 3 = typical 18-year-old qualification; 4 = university degree; and 5 = postgraduate qualification. The average of the level of education obtained by the mother and father was computed to obtain parental education level.

Presence of older siblings

Older siblings have been found in previous research to have an impact on the social environment and influence child development (Nystul, 1981). Since the outcome of interest was sexually dimorphic, this variable was further split into two variables: older brothers present in the home (or not) and older sisters present in the home (or not).

Child's age

The children included in the analyses were between 6 and 10 years of age. Child age was therefore included as a control variable.

Results

Examination of the univariate distributions revealed that fT level was positively skewed, and was the only predictor variable with a distribution that deviated significantly from the Gaussian distribution. Four female outliers in fT levels (individuals who scored three or more standard deviations from the mean) were observed. These outlying values were replaced using a winsorizing procedure (Barnet and Lewis, 1978), where the extreme values are replaced by the highest observed level within three standard deviations from the mean (0.80 nmol/L). The winsorizing procedure was chosen because it is a compromise between the two goals of eliminating the strong influence of extreme values while at the same time utilizing all of the information. Winsorized fT levels showed no outliers and acceptable (skewness < 1) skewness statistics for boys and girls separately and combined, and are used in subsequent analyses.

No significant sex differences were found for any of the predictor variables except fT level. As expected, results showed that boys ($M = 0.84$, $SD = 0.41$) had higher fT levels than girls ($M = 0.32$, $SD = 0.20$), $t(167.97) = 12.92$, $p < .001$, equal variances not

assumed. Table 1 presents the means and standard deviations for predictor variables, AQ-Child and CAST scores.

AQ-Child scores

AQ-Child total

For AQ-Child total score, examination of the univariate distribution revealed that it was not skewed (skewness < 1) for all cases together as well as in boys and girls separately. Figure 1 shows the raw distribution of total AQ-Child scores. AQ-Child scores are normally distributed, as shown in Figure 1.

Scores on the AQ-Child showed significant sex-differences, $t(233) = 6.64$, $p < .001$ (equal variances assumed), with boys ($M = 48.75$, $SD = 17.96$) scoring higher than girls ($M = 34.42$, $SD = 15.01$).

A hierarchical multiple regression analysis was conducted. In the first block, any predictor variable that showed a significant correlation with AQ-Child scores at the $p < .20$ was entered into the analysis (Altman, 1991). In addition, the influence of suppressor variables (predictors that are highly correlated with other predictors in the model at $p < .01$) was investigated. The main effects of fT level and child sex were tested for inclusion in the second block using the stepwise method (entry criterion $p < .05$, removal criterion $p > .10$). The interaction between child sex and fT level was tested for inclusion in the third block using the stepwise method. Table 2 shows the correlation coefficients for both the predictor and outcome variables for boys and girls together.

The predictor variables that correlated with total AQ-Child scores at $p < .20$ were presence of older sisters ($r = -.19$, $p < .01$) and presence of older brothers ($r = -.14$, $p < .05$). The inclusion of fT level in the second block produced a significant *F*-change ($F = 46.35$, $p < .001$, $R^2 = .21$). Inclusion of child sex in the final regression model also produced a significant *F*-change ($F = 7.99$, $p < .05$). The interaction of sex and fT level was excluded as a predictor from the final regression model. See Figure 2 for a visual representation of the relationship between fT level and AQ-Child scores for males and females combined.

Within sex analyses were conducted to further investigate scoring patterns in boys and girls separately. For girls only, parent education level ($r = -.14$, $p < .01$), presence of older sisters ($r = -.17$, $p < .01$), and older brothers ($r = -.23$, $p < .05$) showed correlations at the $p < .20$ level and were entered into the first block using the enter method. fT level ($r = .27$, $p < .001$) was included in the second block using the stepwise method. A significant *F*-change ($F = 4.12$, $p < .01$) was observed when fT was entered into the regression in the second block. The predictor variables that correlated with AQ-Child scores at the $p < .20$ level for boys were presence of older sisters ($r = -.19$, $p < .01$) and brothers ($r = -.14$, $p < .05$). Presence of older sisters and brothers were included in the first block using the enter method. fT level ($r = .22$, $p < .001$) was entered in the second block using the stepwise method. The final model for boys included fT level, which showed a significant *F*-change ($F = 5.13$, $p < .05$). Residual analysis revealed acceptable plots and no outliers.

A score of 76 or higher was considered to indicate risk for ASC (Auyeung *et al.*, in press), and the above analyses were repeated excluding those who scored at or above this cut-off ($N = 10$; 1 girl, 9 boys) to again ensure that the results were not affected by the high scorers on this measure. See Table 3 for AQ-Child regression results with and without the high scorers. The first block included presence of older sisters ($r = -.18$, $p < .01$) and brothers ($r = -.13$, $p = .06$). The second block tested for the

Table 1. Means, standard deviations, and ranges for each sex separately as well combined

	Combined group ($N = 235$)			Girls ($N = 117$)			Boys ($N = 118$)			
Variable	M	*SD*	Range	*M*	*SD*	Range	*M*	*SD*	Range	*t*
fT level (nmol/L)**	0.58	0.42	0.05–2.05	0.32	0.20	0.05 – 0.80	0.84	0.41	0.10 – 2.05	12.39**
Gestational age	16.49	1.44	13–22	16.57	1.48	14 – 22	16.40	1.40	13–20	0.74
Child age	8.91	0.95	6.97–10.68	8.80	0.97	7.01–10.68	9.02	0.92	6.97–10.66	1.76
Maternal age	35.77	4.40	23.68 – 45.90	35.88	4.22	23.68 – 45.66	35.66	4.59	24.28 – 45.90	0.37
Parental education	3.24	1.01	1–5	3.18	0.87	2–5	3.30	1.14	1–5	0.88
CAST total[a]*	4.65	3.87	0 – 22	4.08	3.24	0 – 18	5.22	4.35	0 – 22	2.12*
AQ-Child total**	41.62	18.02	6 – 103	34.42	15.01	6 – 80	48.75	17.96	16 – 103	6.64**

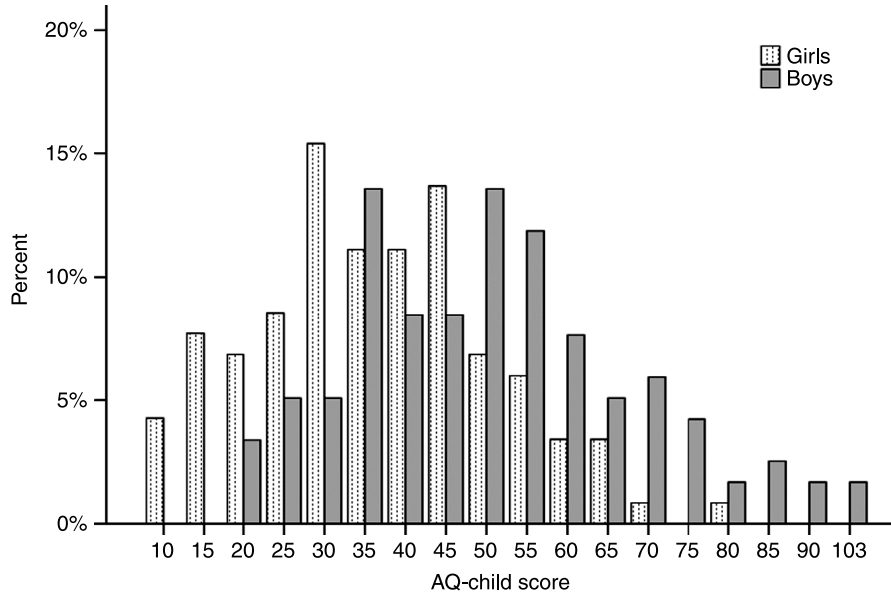
*$p < .05$; **$p < .01$.

[a] Square-root transformation carried out prior to analyses.

Table 2. Correlation matrix showing relationships between predictor variables and CAST and AQ-Child total scores

	fT level	Sex	Gestational age	Child age	Maternal age	Parent education	Older sister	Older brother	CAST total
Sex	.63**	–	–	–	–	–	–	–	–
Gestational age	−.04	−.06	–	–	–	–	–	–	–
Child age	.03	.12	−.05	–	–	–	–	–	–
Maternal age	−.02	−.03	−.28**	−.06	–	–	–	–	–
Parent education	.07	.06	−.09	−.05	.16*	–	–	–	–
Older sister	−.07	−.08	−.11	−.10	−.03	−.07	–	–	–
Older brother	−.04	−.12	−.10	.01	.09	−.15*	.36**	–	–
CAST total	.25**	.14*	.03	.08	−.13	−.02	−.01	−.07	–
AQ-Child total	.41**	.40**	.01	−.01	−.01	−.06	−.19**	−.14*	.25**

*$p < .05$; **$p < .01$.



Note: Maximum score = 141, cut-off score = 76

Figure 1. Distribution of AQ-Child scores.

inclusion of fT level ($r = .43, p < .001$) and child sex ($r = .37, p < .001$). Results show that, when excluding the high scorers, fT level ($F = 48.69, p < .001$) and sex ($F = 4.20$, $p < .05$) both produce a significant *F*-change, and is consistent with results including the high scorers. Residual analysis revealed acceptable plots and no outliers.

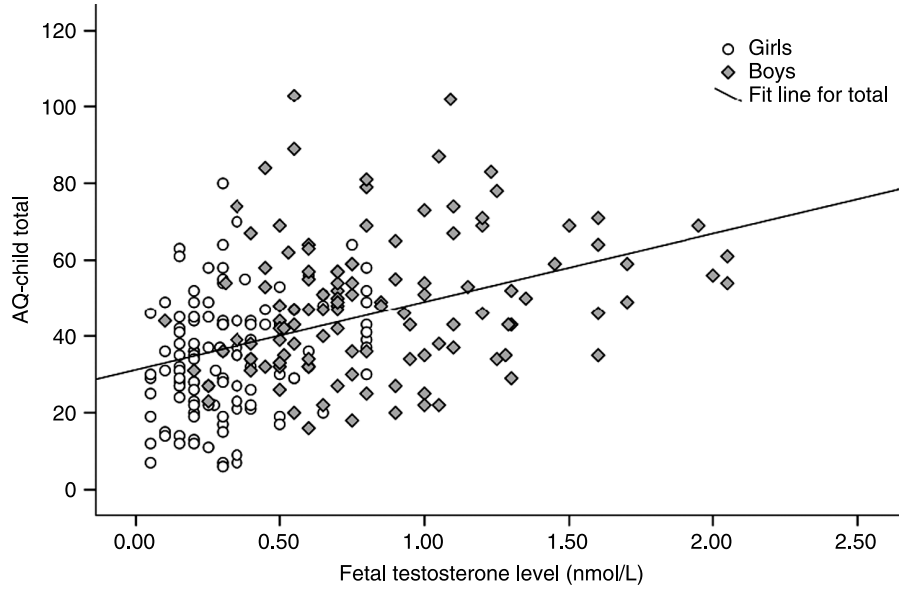


Figure 2. Relationship between fT level and AQ-Child scores.

Table 3. Final regression model for the AQ-Child

Outcome	Predictor variable	R^2	ΔR^2	*B*	*SE*	β
AQ-Child total[a]	Constant	.04	.04	36.02	2.19	
	Older sister			7.36	3.45	0.13*
	Older brother			3.53	4.03	0.06
	fT level	.20	.16	11.61	3.23	0.27**
	Child sex	.23	.03	3.82	1.35	0.21**
AQ-Child total[b]	Constant	.04	.04	33.86	1.92	
	Older sister			5.91	2.99	0.13*
	Older brother			2.81	3.48	0.05
	fT level	.21	.17	11.92	2.84	0.32**
	Child sex	.22	.01	2.44	1.19	0.16*

*$p < .05$; **$p < .01$.
[a] AQ-Child regression results including all participants.
[b] AQ-Child regression results excluding individuals scoring ≥76.

AQ-Child subscales

The AQ-Child subscales were next examined to test if the fT relationships were consistent across all four subscales. Due to the uneven number of items per subscale, a raw score was calculated for each of the four subscales and then divided by the number of items, allowing for comparison between subscales.

Sex differences were explored among the four empirical AQ-Child subscales (see Table 4). All four subscales showed significant sex differences (all $p < .001$) with boys scoring higher than girls.

Table 4. Examination of weighted AQ-Child subscale scores by sex

	Girls ($N = 117$)		Boys ($N = 118$)		
Variable	*M*	*SD*	*M*	*SD*	*t*
Mind reading	0.83	0.46	1.09	0.50	4.14**
Attention to detail	1.05	0.56	1.37	0.58	4.26**
Social skills	0.58	0.39	0.89	0.50	5.30**
Imagination	0.43	0.43	0.82	0.53	6.15**
AQ-Child total	34.42	15.01	48.75	17.96	6.64**

*$p < .05$; **$p < .01$.

Mind reading, attention to detail, social skills, and imagination were significantly associated with fT level and sex. These results are consistent with those observed in AQ-Child total when all participants are examined together (see Table 5).

CAST scores

Examination of univariate distributions indicated that the distribution for CAST scores was positively skewed. Figure 3 shows the raw distribution of CAST scores. CAST scores were transformed by adding one and taking the square root of each score, resulting in a distribution that was not significantly skewed.

Table 5. Correlations for fT level and AQ-Child subscales

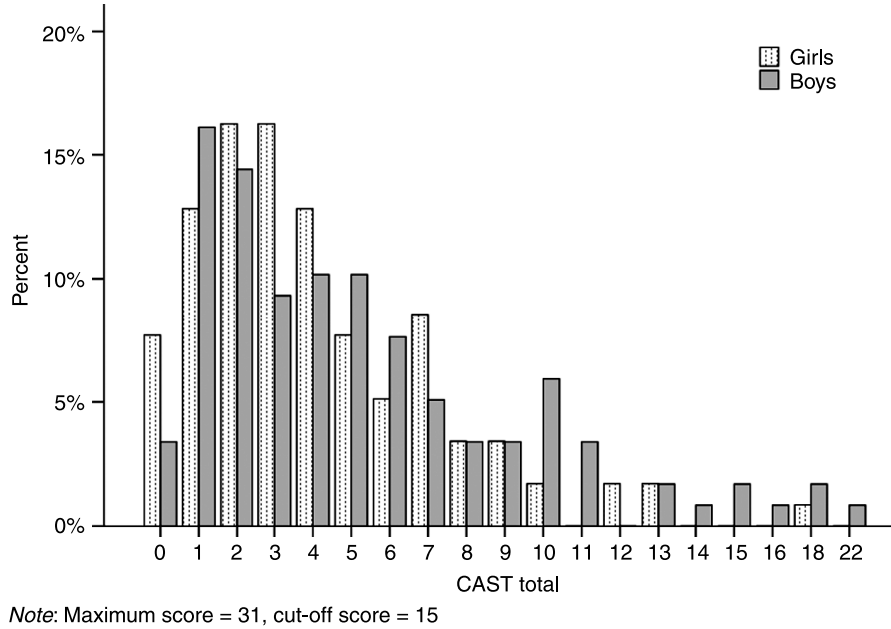
	fT level	Mind reading	Attention detail	Social skills	Imagination
Mind reading	.30**	–	–	–	–
Attention to detail	.27**	.35**	–	–	–
Social skills	.33**	.62**	.31**	–	–
Imagination	.38**	.38**	.17**	.38**	–
AQ-Child total	.41**	.86**	.61*	.83**	.57**

*$p < .05$; **$p < .01$.

Transformed CAST scores showed significant sex-differences, $t(226.55) = 2.12$, $p < .05$, equal variances not assumed, with boys ($M = 2.36$, $SD = 0.82$) scoring higher than girls ($M = 2.15$, $SD = 0.69$).

For the hierarchical regression analysis the predictor variables that correlated with CAST scores at the $p < .20$ level, were sex ($r = .14$, $p < .05$), fT ($r = .25$, $p < .001$), and maternal age ($r = -.13$, $p = .06$). No suppressor variables were observed. Inclusion of fT level in the second block produced a significant *F*-change ($F = 10.72$, $p < .01$, $R^2 = .07$). The main effect of sex was excluded as a predictor. Residual analysis showed no outliers and acceptable plots. See Figure 4 for a visual representation of the relationship between fT level and CAST scores.

In addition, to further investigate whether the results might be due to a sex difference (not necessarily involving fT), we analysed the relationship between these scores and fT within each sex. For boys, maternal age ($r = -.15$, $p = .12$) and presence of older brothers ($r = .14$, $p = .15$) met criteria for entry into the analysis ($r = .14$, $p < .001$). Presence of older sisters was included as a suppressor variable due to its high correlation with the presence of older brothers ($r = .32$, $p < .001$). Inclusion of fT level



Note: Maximum score = 31, cut-off score = 15

Figure 3. Distribution of CAST scores.

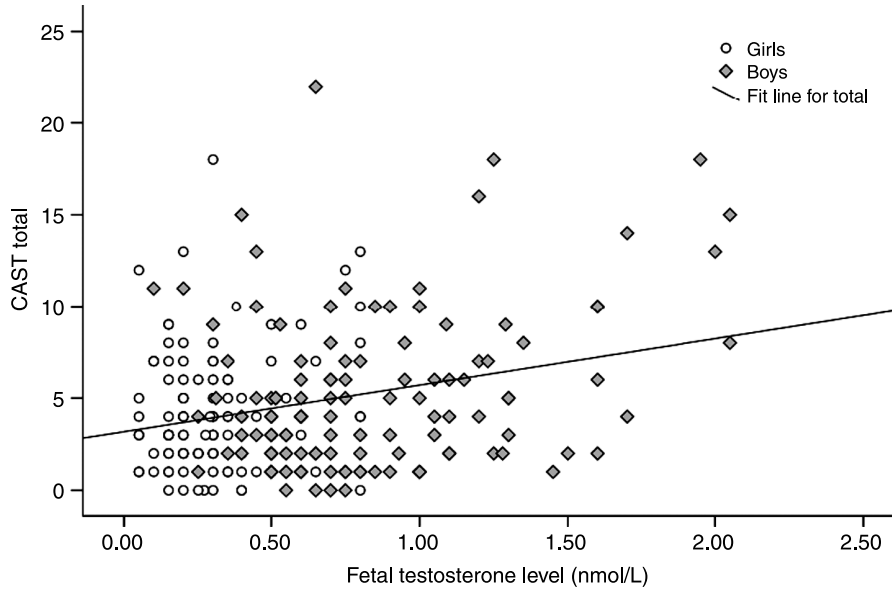


Figure 4. Relationship between fT level and CAST scores.

in the second block produced a significant *F*-change ($F = 6.57$, $p < .05$, $R^2 = .12$). For girls alone, no significant relationship was found between CAST scores and fT levels, therefore regression analyses were not conducted.

A raw score of 15 or higher was considered to indicate risk for ASC (Scott *et al.*, 2002b). As for the AQ-Child, the above analyses were repeated excluding those who scored at or above this cut-off ($N = 7$; 1 girl, 6 boys). Table 6 shows regression results for the CAST with and without high scorers. Maternal age ($r = -.12, p < .20$) and presence of older brothers ($r = -.10, p < .20$) were entered into the first block using the enter method. Presence of older sisters was also included in the first block due to its positive relationship with older brothers ($r = .37, p < .001$). The second block included fT level ($r = .32, p < .001$) and child sex ($r = .32, p < .001$). Inclusion of fT level in the second

Table 6. Final hierarchical regression model for the CAST

Outcome	Predictor variable	R^2	ΔR^2	*B*	*SE*	β
CAST[a]	Constant	.02	.02	2.76	0.42	
	Mother age			0.02	0.01	0.13
	fT level	.07	.05	11.61	3.23	0.22**
CAST[b]	Constant	.03	.03	2.60	0.38	
	Mother age			0.02	0.01	0.11
	Older sister			0.25	0.17	0.11
	Older brother			0.15	0.14	0.08
	fT level	.05	.02	0.25	0.12	0.14*

Note. Square-root transformation was conducted before analysis.
*$p < .05$; **$p < .01$.
[a] CAST regression results including all participants.
[b] CAST regression results excluding individuals scoring >15.

block produced a significant *F*-change ($F = 4.29$, $p < .05$, $R^2 = .05$), while child sex and the interaction of sex and fT level were excluded as predictors from the final regression model. These results suggest that children with higher fT levels score higher on the CAST.

IQ scores

Table 7 shows the means, standard deviations, and *t* statistics for all variables for this subset of children. No significant sex-differences were found between boys and girls for full scale IQ, performance IQ, verbal IQ, or block design scores (using the Bonferroni correction for multiple comparisons). No significant correlations were found between any of the predictor variables and full scale IQ, performance IQ, verbal IQ, and block design scores. Regression analyses were not conducted for these variables (see Table 8).

Table 7. Means and standard deviations for subset of children by sex

	Girls ($N = 31$)		Boys ($N = 43$)		
Variable	*M*	*SD*	*M*	*SD*	*t*
fT level (nmol/L)**	0.35	0.23	0.79	0.35	6.09**
Gestational age	16.44	1.25	16.15	1.63	0.65
Child age	9.23	0.80	9.16	0.77	0.35
Maternal age	34.79	4.70	35.71	4.70	0.81
Parental education	3.30	0.81	3.60	1.13	1.22
CAST total[a]*	3.84	3.31	4.70	3.76	1.10
AQ-Child total**	32.35	13.03	46.67	17.54	3.84**
Full scale IQ	100.61	13.86	108.95	15.64	2.37
Verbal IQ	94.90	12.82	100.26	15.52	1.51
Performance IQ	108.13	17.39	116.26	16.70	2.03
Block design	16.26	10.08	19.14	9.41	1.26

*$p < .05$; **$p < .01$.
[a] Square-root transformation carried out prior to analyses.

Table 8. Correlations for IQ scales and fT level

	fT level	Full IQ	Verbal IQ	Performance IQ
Full IQ	.13	–	–	–
Verbal IQ	.02	.83**	–	–
Performance IQ	.19	.84**	.41**	–
Block design	.16	.68**	.27*	.85**

*$p < .05$; **$p < .01$.

Discussion

This study examines the relationship between fT levels and the later development of autistic traits as measured by the CAST and the AQ-Child. fT levels were found to be positively associated with number of autistic traits. The significant positive relationship between fT levels was observed across CAST total score, AQ-Child total score, as well as

in the four subscales of the AQ-Child. Results remained consistent when excluding individuals who scored above the established cut-offs for the CAST and AQ-Child.

The finding that boys scored higher on both the AQ-Child and CAST coupled with the significant positive associations observed between fT levels and these measures provides support for the EMB theory at both psychometric and biological levels. While this dual-level convergence of evidence has been reported previously using indirect biological measures such as the 2D:4D ratio (assumed to be a proxy for fT) (Manning *et al.*, 2001), or brain activity using fMRI (Baron-Cohen *et al.*, 2006), this is first time that such a relationship has been reported using direct measures. We can assume that the observed positive association between fT levels and autistic traits may reflect a direct causal effect of fT on neural development, but this remains speculative due to the correlational design of this study. It is for example possible that fT is serving as an index for an unknown third variable.

Scores from the AQ-Child showed a significant positive relationship with fT levels when the sexes were combined as well as when they were examined independently. The CAST, however, was found to have a significant positive relationship with fT levels when the sexes were combined and in boys only. The relationship was not found between CAST and fT levels when girls were examined alone. fT levels predicted about 7% of the variation in CAST scores, and 20% of the variation in AQ-Child scores. These results suggest that the AQ-Child is a better instrument for measuring the variance in autistic traits in typically developing children. This may be because the AQ-Child is a wider scale with a less skewed distribution, relative to the CAST.

Although the correlation between the CAST and AQ-Child ($r = .25$, $p < .001$) was statistically significant, the low coefficient does not suggest high convergent validity. This may be because of the differing response format of each questionnaire. The CAST requires a binary response, where parents indicate whether their child presents a particular behaviour. Naturally, one would expect to see many children who have low scores, because the majority of children are not at risk for ASC. In contrast, the AQ-Child is answered in Likert format, and was designed to measure the extent to which the parents agree their child exhibits certain behaviours, resulting in a larger variation in scores.

It is of interest to note that most other research examining the role of testosterone in human psychosexual development has produced more supportive evidence in females than in males (Hines, 2004, 2006). In the current study, we found a significant relationship between fT levels and AQ-Child scores in boys and girls separately. For CAST scores, a significant correlation with fT levels in boys was observed, but no significant relationship was seen with fT levels in girls. However, compared to boys, girls showed less variability in CAST scores as well as fT levels possibly limiting the ability to detect associations.

No relationships were found between full scale IQ and the predictor variables (including fT) or the CAST or AQ-Child scores. Our prediction that fT level would be significantly positively related to Block Design scores was not supported ($r = .16$, $p > .05$). Our results suggest that fT level does not account for individual variation in full scale IQ, verbal IQ, performance IQ, or Block Design scores. No sex differences were found between these variables. It is possible that testosterone only plays a role in a subset of behaviours where sex differences are found. However, an alternative explanation could be that the strength of the relationship between fT levels and outcome may vary for different behaviours. If the effect of normal variation in fT level on IQ is small, then only studies with large sample sizes will reveal it. It is also possible that

fT levels contribute to IQ but does so at a different time period than that examined in this study. However, the lack of significant sex differences in IQ scores in the current study suggest that even in a larger sample, no relationship between fT levels and IQ would have been observed. It would nevertheless be beneficial for future studies to examine the relationship between fT levels and IQ scores (in particular block design performance) in a larger sample of children. In addition, it would be interesting to investigate if these relationships remain consistent throughout adolescence and adulthood, since sex differences in block design performance have been seen in adults (Lynn, 1998; Lynn *et al.*, 2005; Rönnlund & Nilsson, 2006).

Whilst we have obtained evidence that higher fT levels are associated with a child exhibiting more autistic traits, it remains unknown if postnatal testosterone (pT) level would also show such an association. Addressing this question would help to narrow down the window of time in which testosterone exerts its effects on neural development. It is possible that for testosterone to have such effects, these need to occur within a 'critical period' (during the second trimester). In addition to the surge that occurs in the second trimester of pregnancy, a second peak in circulating testosterone occurs in human male infants a few days after birth. The levels of pT are strikingly high: they remain in a pubertal range for a few months and then usually drop to barely detectable levels by 4–6 months of age (Smail, Reyes, Winter, & Faiman, 1981). It has been proposed that pT levels in humans influence later spatial skills such as mental rotation ability (Hines *et al.*, 2003), but there is a clear need for further research into the roles of fT and pT in relation to behavioural development before conclusions can be reached.

It is known that hormones fluctuate during the day and across days, even in fetuses (Seron-Ferre, Ducsay, & Valenzuela, 1993; Walsh, Ducsay, & Novy, 1984). It is ethically unacceptable to attempt to obtain repeated samples of fT during 1 day or across days, during gestation. Therefore the representativeness of a single sample of fT remains unclear. Given the reported time course of testosterone secretion (Smail *et al.*, 1981), it is likely the most promising time to be measuring fT is prenatal weeks 8–24 (Baron-Cohen *et al.*, 2004; Collaer & Hines, 1995; Hines, 2004), but this is still a relatively wide range. Since the decision to perform amniocentesis is understandably based on clinical–medical factors, rather than purely scientific ones, the inferences we can draw about the single measure of fT will necessarily be limited.

While the factors contributing to individual variation in fT levels are not fully understood, adolescent and adult levels of testosterone are heritable (Harris, Vernon, & Boomsma, 1998; Hoekstra, Bartels, & Boomsma, 2006). There is also significant heritability of autism in clinical samples (Rutter, 2000) and of autistic traits measured in the general population (Hoekstra *et al.*, 2007; Ronald *et al.*, 2006). It may therefore be likely that genetic effects underlie the relationship between fT levels and number of autistic traits. A recent study found both diagnosis of ASC, and number of autistic traits, are significantly associated with specific candidate genes that regulate testosterone production and synthesis (Chakrabarti *et al.*, 2008). It is not yet clear though if the effects reported here are purely genetic, if high fT levels also serve as an endocrine environmental risk factor for autism, or if there are gene–environment interaction effects, such that high fT only increases risk for ASC in genetically susceptible individuals.

Our studies suggest that variations in fT levels are related to autistic traits as well as other specific aspects of sexually dimorphic behaviour and cognition in typically developing children (Auyeung *et al.*, 2006; Baron-Cohen *et al.*, 2004; Chapman *et al.*,

2006; Grimshaw, Sitarenios, & Finegan, 1995; Knickmeyer *et al.*, 2005; Knickmeyer, Baron-Cohen *et al.*, 2006; Lutchmaya *et al.*, 2002a, 2002b), but caution needs to be taken when extrapolating these results to individuals with a formal diagnosis of ASC. The current study is too small a sample to be able to test if fT levels are elevated in formally diagnosed cases of ASC, since these have a prevalence rate of about 1% (Baird *et al.*, 2006). We are currently conducting a large-scale collaboration with the Danish Biobank so as to increase our sample size sufficiently to compare fT levels of cases versus controls.

Conclusions

There is converging evidence from a range of methods showing the masculinizing effects of fT on individual differences in the following sexually dimorphic traits: social skills; language development; empathy; systemizing; and visual-analytic skills. If, according to the EMB theory, ASC is an extreme of male-typical behaviour, exposure to elevated levels of fT could be one important factor that is involved with the development of the condition. This study was conducted to examine the effects of fT levels on the development of autistic traits. fT levels were positively associated with higher scores (indicating greater number of autistic traits) on the CAST and the AQ-Child. No relationships were found between IQ and the predictor variables. Further work is needed to test for any link between fT and later development of ASC, focusing on individuals with a clinical diagnosis.

Acknowledgements

This work was supported by the Nancy Lurie-Marks Family Foundation and the Medical Research Council (MRC). Bonnie Auyeung was supported by a scholarship from Trinity College, Cambridge. We are grateful to the families who have taken part in this longitudinal study over many years and to Melissa Hines, Rosa Hoekstra, Bhismadev Chakrabarti, Malcolm Bang, Svetlana Lutchmaya, Greg Davis, and Ieuan Hughes for valuable discussions.

References

Altman, D. G. (1991). *Practical statistics for medical research*. London: Chapman and Hall.

APA (1994). *DSM-IV Diagnostic and Statistical Manual of Mental Disorders* (4th ed.). Washington, DC: American Psychiatric Association.

Auyeung, B., Baron-Cohen, S., Chapman, E., Knickmeyer, R. C., Taylor, K., & Hackett, G. (2006). Foetal testosterone and the child systemizing quotient. *European Journal of Endocrinology*, *155*, S123–S130.

Auyeung, B., Baron-Cohen, S., Wheelwright, S., & Allison, C. (in press). Development of the autism spectrum quotient – children's version (AQ-Child). *Journal of Autism and Developmental Disorders*.

Barnet, V., & Lewis, T. (1978). *Outliers in statistical data.* New York: Wiley.

Baird, G., Simonoff, E., Pickles, A., Chandler, S., Loucas, T., Meldrum, D., *et al.* (2006). Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: The Special Needs and Autism Project (SNAP). *Lancet*, *368*, 210–215.

Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind. Learning, development, and conceptual change* (p. 171) Cambridge, MA: The MIT Press.

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

Baron-Cohen, S. (2003). *The essential difference: Men, women and the extreme male brain*. London: Penguin.

Baron-Cohen, S., Baldwin, D. A., & Crowson, M. (1997). Do children with autism use the speaker's direction of gaze strategy to crack the code of language? *Child Development*, *68*, 48–57.

Baron-Cohen, S., Knickmeyer, R., & Belmonte, M. K. (2005). Sex differences in the brain: Implications for explaining autism. *Science*, *310*, 819–823.

Baron-Cohen, S., Lutchmaya, S., & Knickmeyer, R. (2004). *Prenatal testosterone in mind*. Cambridge, MA: The MIT Press.

Baron-Cohen, S., O'Riordan, M., Stone, V., Jones, R., & Plaisted, K. (1999). Recognition of faux pas by normally developing children and children with Asperger syndrome or high-functioning autism. *Journal of Autism and Developmental Disorders*, *29*, 407–418.

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

Baron-Cohen, S., Ring, H., Chitnis, X., Wheelwright, S., Gregory, L., Williams, S., *et al.* (2006). fMRI of parents of children with Asperger syndrome: A pilot study. *Brain Cognition*, *61*, 122–130.

Baron-Cohen, S., Scott, F. J., Allison, C., Williams, J., Bolton, P., Matthews, F. E., *et al.* (2008). *Estimating autism spectrum prevalence in the population: A school based study from the UK*. Manuscript submitted for publication.

Baron-Cohen, S., Wheelwright, S., & Hill, J. (2001). The 'Reading the mind in the eyes' test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychiatry and Psychiatry*, *42*, 241–252.

Bryden, M. P., McManus, I. C., & Bulman-Fleming, M. B. (1994). Evaluating the empirical support for the Geschwind-Behan-Galaburda model of cerebral lateralization. *Brain and Cognition*, *26*, 103–167.

Bryson, S. E., & Smith, I. M. (1998). Epidemiology of autism: Prevalence, associated characteristics, and implications for research and service delivery. *Mental Retardation and Developmental Disabilities Research Reviews. Special Issue: Autism*, *4*, 97–103.

Chakrabarti, B., Hill-Cawthorne, G., Dudbridge, F., Wheelwright, S., Allison, C., & Baron-Cohen, S. (2008). *Candidate genes associated with Asperger syndrome and four measures of autistic traits and the 'broader autism phenotype'*. Manuscript submitted for publication.

Chakrabarti, S., & Fombonne, E. (2005). Pervasive developmental disorders in preschool children: Confirmation of high prevalence. *American Journal of Psychiatry*, *162*, 1133–1141.

Chapman, E., Baron-Cohen, S., Auyeung, B., Knickmeyer, R., Taylor, K., & Hackett, G. (2006). Fetal testosterone and empathy: Evidence from the empathy quotient (EQ) and the 'Reading the mind in the eyes' test. *Social Neuroscience*, *1*, 135–148.

Cohen-Bendahan, C. C., van de Beek, C., & Berenbaum, S. A. (2005). Prenatal sex hormone effects on child and adult sex-typed behavior: Methods and findings. *Neuroscience and Biobehavioral Reviews*, *29*, 353–384.

Collaer, M. L., & Hines, M. (1995). Human behavioural sex differences: A role for gonadal hormones during early development? *Psychological Bulletin*, *118*, 55–107.

Davis, M. H. (1994). *Empathy: A social psychological approach*. Boulder, CO: Westview Press.

Fausto-Sterling, A. (1992). *Myths of gender*. New York: Basic Books.

Fein, D., Waterhouse, L., Lucci, D., Pennington, B., & Humes, M. (1985). Handedness and cognitve functions in pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, *15*, 323–333.

Finegan, J., Bartleman, B., & Wong, P. Y. (1991). A window for the study of prenatal sex hormone influences on postnatal development. *Journal of Genetic Psychology*, *150*, 101–112.

Finegan, J. A., Quarrington, B. J., Hughes, H. E., Mervyn, J. M., Hood, J. E., Zacher, J. E., *et al.* (1990). Child outcome following mid-trimester amniocentesis: Development, behaviour, and physical status at age 4 years. *British Journal of Obstetrics and Gynaecology*, *97*, 32–40.

Fombonne, E. (2005). The changing epidemiology of autism. *Journal of Applied Research in Intellectual Disabilities*, *18*, 281–294.

Geschwind, N., & Galaburda, A. M. (1985). Cerebral lateralization: Biological mechanisms, associations, and pathology. *Archives of Neurology*, *42*, 428–654.

Giedd, J. N., Snell, J. W., Lange, N., Rajapakse, J. C., Casey, B. J., Kozuch, P. L., *et al.* (1996). Quantitative magnetic resonance imaging of human brain development: Ages 4–18. *Cerebral Cortex*, *6*, 551–560.

Grimshaw, G. M., Sitarenios, G., & Finegan, J. K. (1995). Mental rotation at 7 years: Relations with prenatal testosterone levels and spatial play experiences. *Brain and Cognition*, *29*, 85–100.

Hampson, E., Rovet, J. F., & Altmann, D. (1998). Spatial reasoning in children with congenital adrenal hyperplasia due to 21-hydroxylase deficiency. *Developmental Neuropsychology*, *14*, 299–320.

Happe, F. G. (1994). Wechsler IQ profile and theory of mind in autism: A research note. *Journal of Child Psychology and Psychiatry*, *35*, 1461–1471.

Harden, A. Y., Minshew, N. J., Mallikarjuhn, M., & Keshavan, M. S. (2001). Brain volume in autism. *Journal of Child Neurology*, *16*, 421–424.

Harris, J. A., Vernon, P. A., & Boomsma, D. I. (1998). The heritability of testosterone: A study of Dutch adolescent twins and their parents. *Behavior Genetics*, *28*, 165–171.

Hazlett, H. C., Poe, M., Gerig, G., Smith, R. G., Provenzale, J., Ross, A., *et al.* (2005). Magnetic resonance imaging and head circumference study of brain size in autism: Birth through age 2 years. *Archives of General Psychiatry*, *62*, 1366–1376.

Hines, M. (2004). *Brain gender*. New York: Oxford University Press.

Hines, M. (2006). Prenatal testosterone and gender-related behaviour. *European Journal of Endocrinology*, *155*, S115–S121.

Hines, M., Fane, B. A., Pasterski, V. L., Matthews, G. A., Conway, G. S., & Brook, C. (2003). Spatial abilities following prenatal androgen abnormality: Targeting and mental rotations performance in individuals with congenital adrenal hyperplasia. *Psychoneuroendocrinology*, *28*, 1010–1026.

Hines, M., & Shipley, C. (1984). Prenatal exposure to diethylstilbestrol (DES) and the development of sexually dimorphic cognitive abilities and cerebral lateralization. *Developmental Psychology*, *20*, 81–94.

Hoekstra, R., Bartels, M., & Boomsma, D. I. (2006). Heritability of testosterone levels in 12-year-old twins and its relation to pubertal development. *Twin Research and Human Genetics*, *9*, 558–565.

Hoekstra, R. A., Bartels, M., Cath, D. C., & Boomsma, D. I. (in press). Factor structure, reliability and criterion validity of the Autism-spectrum Quotient (AQ): A study in Dutch population and patient groups. *Journal of Autism and Developmental Disorders*.

Hoekstra, R. A., Bartels, M., Verweij, C. J., & Boomsma, D. I. (2007). Heritability of autistic traits in the general population. *Archives of Pediatric and Adolescent Medicine*, *161*, 372–377.

Johnson, J. A., Cheek, J. M., & Smither, R. (1983). The structure of empathy. *Journal of Personality and Social Psychology*, *45*, 1299–1312.

Judd, H. L., Robinson, J. D., Young, P. E., & Jones, O. W. (1976). Amniotic fluid testosterone levels in midpregnancy. *Obstetrics and Gynecolology*, *48*, 690–692.

Knickmeyer, R. C., & Baron-Cohen, S. (2006). Fetal testosterone and sex differences. *Early Human Development*, *82*, 755–760.

Knickmeyer, R., Baron-Cohen, S., Raggatt, P., & Taylor, K. (2005). Foetal testosterone, social relationships, and restricted interests in children. *Journal of Child Psychology and Psychiatry*, *46*, 198–210.

Knickmeyer, R., Baron-Cohen, S., Raggatt, P., Taylor, K., & Hackett, G. (2006). Fetal testosterone and empathy. *Hormones and Behavior*, *49*, 282–292.

Knickmeyer, R. C., Fane, B. A., Mathews, G., Wheelwright, S., Baron-Cohen, S., Conway, G., *et al.* (2006). Autistic traits in people with congenital adrenal hyperplasia: A test of the fetal testosterone theory of autism. *Hormones and Behavior*, *50*, 148–153.

Lawrence, E. J., Shaw, P., Baker, D., Baron-Cohen, S., & David, A. S. (2004). Measuring empathy: Reliability and validity of the empathy quotient. *Psychological Medicine*, *34*, 911–919.

Lutchmaya, S., Baron-Cohen, S., & Raggatt, P. (2002a). Foetal testosterone and eye contact in 12 month old infants. *Infant Behavior and Development*, *25*, 327–335.

Lutchmaya, S., Baron-Cohen, S., & Raggatt, P. (2002b). Foetal testosterone and vocabulary size in 18- and 24-month-old infants. *Infant Behavior and Development*, *24*, 418–424.

Lutchmaya, S., Baron-Cohen, S., Raggatt, P., Knickmeyer, R., & Manning, J. T. (2004). 2nd to 4th digit ratios, fetal testosterone and estradiol. *Early Human Development*, *77*, 23–28.

Lynn, R. (1998). Sex differences in intelligence: Data from a Scottish standardisation of the WAIS-R. *Personality and Individual Differences*, *24*, 289–290.

Lynn, R., Raine, A., Venables, P. H., Mednick, S. A., & Irwing, P. (2005). Sex differences on the WISC-R in Mauritius. *Intelligence*, *33*, 527–533.

Malas, M. A., 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.

Manning, J. T., Baron-Cohen, S., Wheelwright, S., & Sanders, G. (2001). The 2nd to 4th digit ratio and autism. *Developmental Medicine and Child Neurology*, *43*, 160–164.

Manning, J. T., Scutt, D., Wilson, J., & Lewis-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.

McManus, I. C., Murray, B., Doyle, K., & Baron-Cohen, S. (1992). Handedness in childhood autism shows a dissociation of skill and preference. *Cortex*, *28*, 373–381.

Milne, E., White, S., Campbell, R., Swettenham, J., Hansen, P., & Ramus, F. (2006). Motion and form coherence detection in autistic spectrum disorder: Relationship to motor control and 2:4 digit ratio. *Journal of Autism and Developmental Disorders*, *36*, 225–237.

New, M. I. (1998). Diagnosis and management of congenital adrenal hyperplasia. *Annual Review of Medicine*, *49*, 311–328.

Nystul, M. S. (1981). Effects of siblings' sex composition on self-concept. *Journal of Psychology: Interdisciplinary and Applied*, *108*, 133–136.

Quadagno, D. M., Briscoe, R., & Quadagno, J. S. (1977). Effects of perinatal gonadal hormones on selected nonsexual behavior patterns: A critical assessment of the nonhuman and human literature. *Psychological Bulletin*, *84*, 62–80.

Resnick, S. M., Berenbaum, S. A., Gottesman, I. I., & Bouchard, T. J. (1986). Early hormonal influences on cognitive functioning in congenital adrenal hyperplasia. *Developmental Psychology*, *22*, 191–198.

Ronald, A., Happe, F., Bolton, P., Butcher, L. M., Price, T., Wheelwright, S., *et al.* (2006). Genetic heterogeneity between the three components of the autism spectrum: A twin study. *Journal of the American Academy of Child and Adolescent Psychiatry*, *45*, 691–699.

Rönnlund, M., & Nilsson, L. (2006). Adult life-span patterns in WAIS-R Block Design performance: Cross-sectional versus longitudinal age gradients and relations to demographic factors. *Intelligence*, *34*, 63–78.

Rutter, M. (1978). Diagnosis and definition. In I. M. R. E. Schopler (Ed.), *Autism: A reappraisal of concepts and treatment* (pp. 1–26). New York: Plenum Press.

Rutter, M. (2000). Genetic studies of autism: From the 1970s into the millennium. *Journal of Abnormal Child Psychology*, *28*, 3–14.

Rutter, M., Caspi, A., & Moffitt, T. E. (2003). Using sex differences in psychopathology to study causal mechanisms: Unifying issues and research strategies. *Journal of Child Psychology and Psychiatry*, *44*, 1092–1115.

Satz, P., Soper, H., Orsini, D., Henry, R., & Zvi, J. (1985). Handedness subtypes in autism. *Psychiatric Annals*, *15*, 447–451.

Schindler, A. E. (1982). Hormones in human amniotic fluid. *Monographs on Endocrinology*, *21*, 1–158.

Scott, F. J., Baron-Cohen, S., Bolton, P., & Brayne, C. (2002a). Brief report: Prevalence of autism spectrum conditions in children aged 5–11 years in Cambridgeshire, UK. *Autism*, *6*, 231–237.

Scott, F. J., Baron-Cohen, S., Bolton, P., & Brayne, C. (2002b). The CAST (Childhood Asperger Syndrome Test): Preliminary development of a UK screen for mainstream primary-school-age children. *Autism*, *6*, 9–13.

Seron-Ferre, M., Ducsay, C. A., & Valenzuela, G. J. (1993). Circadian rhythms during pregnancy. *Endocrine Reviews*, *14*, 594–609.

Shah, A., & Frith, C. (1993). Why do autistic individuals show superior performance on the block design task? *Journal of Child Psychology and Psychiatry*, *34*, 1351–1364.

Smail, P. J., Reyes, F. I., Winter, J. S. D., & Faiman, C. (1981). The fetal hormonal environment and its effect on the morphogenesis of the genital system. In S. J. Kogan & E. S. E. Hafez (Eds.), *Pediatric andrology* (pp. 9–19). Boston, MA: Martinus Nijhoff.

Smith, L. L., & Hines, M. (2000). Language lateralization and handedness in women prenatally exposed to diethylstilbestrol (DES). *Psychoneuroendocrinology*, *25*, 497–512.

Soper, H., Satz, P., Orsini, D., Henry, R., Zvi, J., & Schulman, M. (1986). Handedness patterns in autism suggest subtypes. *Journal of Autism and Developmental Disorders*, *16*, 155–167.

Stodgell, C. J., Ingram, J. I., & Hyman, S. L. (2001). The role of candidate genes in unraveling the genetics of autism. *International Review of Research in Mental Retardation*, *23*, 57–81.

Swettenham, J., Baron-Cohen, S., Charman, T., Cox, A., Baird, G., Drew, A., *et al.* (1998). The frequency and distribution of spontaneous attention shifts between social and non-social stimuli in autistic, typically developing, and non-autistic developmentally delayed infants. *Journal of Child Psychology and Psychiatry*, *9*, 747–753.

Tidmarsh, L., & Volkmar, F. R. (2003). Diagnosis and epidemiology of autism spectrum disorders. *Canadian Journal of Psychiatry*, *48*, 517–525.

Wakabayashi, A., Baron-Cohen, S., Wheelwright, S., & Tojo, Y. (2006). The autism-spectrum quotient (AQ) in Japan: A cross-cultural comparison. *Journal of Autism and Developmental Disorders*, *36*, 263–270.

Walsh, S. W., Ducsay, C. A., & Novy, M. J. (1984). Circadian hormonal interactions among the mother, fetus, and amniotic fluid. *Americal Journal of Obstetrics and Gynecology*, *150*, 745–753.

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence*. San Antonio, TX: The Psychological Corporation and Harcourt Brace.

Williams, J., Allison, C., Scott, F., Stott, C., Bolton, P., Baron-Cohen, S., *et al.* (2006). The Childhood Asperger Syndrome Test (CAST): Test-retest reliability. *Autism*, *10*, 415–427.

Williams, J., Scott, F., Stott, C., Allison, C., Bolton, P., Baron-Cohen, S., *et al.* (2005). The CAST (Childhood Asperger Syndrome Test): Test accuracy. *Autism*, *9*, 45–68.

Wing, L. (1981). Asperger's syndrome: A clinical account. *Psychological Medicine*, *11*, 115–129.

Woodbury-Smith, M. R., Robinson, J., Wheelwright, S., & Baron-Cohen, S. (2005). Screening adults for Asperger syndrome using the AQ: A preliminary study of its diagnostic validity in clinical practice. *Journal of Autism and Developmental Disorders*, *35*, 331–335.

Received 11 January 2007; revised version received 3 April 2008