Arch Sex Behav
DOI 10.1007/s10508-011-9864-8

ORIGINAL PAPER

Effects of Fetal Testosterone on Visuospatial Ability

Bonnie Auyeung · Rebecca Knickmeyer · Emma Ashwin · Kevin Taylor · Gerald Hackett · Simon Baron-Cohen

Received: 3 January 2011 / Revised: 22 September 2011 / Accepted: 22 September 2011

Abstract This study investigated whether fetal testosterone (FT) measured from second trimester amniotic fluid was related to specific aspects of visuospatial ability, in children aged 7–10 years (35 boys, 29 girls). A series of tasks were used: the children's Embedded Figures Test (EFT) (a test of attention to detail), a ball targeting task (measuring hand-eye coordination), and a computerized mental rotation task (measuring rotational ability). FT was a significant predictor for EFT scores in both boys and girls, with boys also showing a clear advantage for this task. No significant sex differences were observed in targeting. Boys scored higher than girls on mental rotation. However, no significant relationships were observed between FT and targeting or mental rotation. Girls' performance on the mental rotation and targeting tasks was significantly related to age, indicating that these tasks may have been too difficult for the younger children. These results indicate that FT has a significant role in some aspects of cognitive development but that further work is needed to understand its effect on the different aspects of visuospatial ability.

Keywords Spatial ability · Mental rotation · Sex differences · Fetal testosterone

Introduction

There is increasing evidence that prenatal hormones have a significant effect on sex-typical aspects of human behavior (Baron-Cohen, Knickmeyer, & Belmonte, 2005; Baron-Cohen, Lutchmaya, & Knickmeyer, 2004; Cohen-Bendahan, van de Beek, & Berenbaum, 2005; Hines, 2004). Behaviors showing clear sex differences are therefore considered the best candidates for studying the effects of prenatal hormones on later development (Cohen-Bendahan et al., 2005; Collaer & Hines, 1995; Hines, 2004). One of the most established areas where sex differences have been observed is the visuospatial domain, where research has shown that males outperform females on a range of tasks (Halpern, 2000; Kimura, 1999).

The largest sex difference in visuospatial ability is found in mental rotation (the ability to rotate mental images quickly and accurately) and targeting accuracy (the ability to aim projectiles accurately at a specified point in space). Effect sizes for sex differences were computed using Cohen's d where a d value of .2 is considered a small effect size, a d of .5 is considered a medium effect size, and a d greater than .8 is considered a large effect size (Cohen, 1988). A significant male advantage has been observed in both two-dimensional mental rotation ($d = .3$–$.5$) (Linn & Petersen, 1986; Voyer, Voyer, & Bryden, 1995) and targeting accuracy ($d = 1.9$) (Watson & Kimura, 1991). For mental rotation, a sex difference has been observed in preadolescent children ($d = .12$–$.42$, depending on the type of task used) (Kerns & Berenbaum, 1991; Levine, Huttenlocher, Taylor, & Langrock, 1999) which increases with age (Voyer et al., 1995). For targeting accuracy, a meta-analysis showed sex differences ($d = .96$) across age favoring males (Thomas & French, 1985). Sex

---

B. Auyeung (✉) · R. Knickmeyer · E. Ashwin · S. Baron-Cohen
Department of Psychiatry, University of Cambridge, Douglas House, 18B Trumpington Rd., Cambridge CB2 8AH, UK
e-mail: ba251@cam.ac.uk

R. Knickmeyer
Department of Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

K. Taylor
Department of Clinical Biochemistry, Addenbrooke's Hospital, Cambridge, UK

G. Hackett
Department of Foetal Medicine, Rosie Maternity Hospital, Robinson Way, Cambridge, UK

differences in other aspects of targeting performance, such as throwing velocity ($d = 1.98$) and throwing distance ($d = 2.18$), have also been found (Thomas & French, 1985).

Figure disembedding is another area where a male advantage in both children ($d = .49$) (Nebot, 1988) and adults ($d = .88$) (Parlee & Rajagopal, 1974) has been found. The Embedded Figures Test (EFT) requires the ability to segment a global image into component parts and to identify one local detail (Shah & Frith, 1983). It is variously considered a measure of field independence (Karp & Konstadt, 1963; Witkin, Oltman, Raskin, & Karp, 1971), segmentation (Shah & Frith, 1993), and attention to detail (Baron-Cohen, 2002).

Studies in non-human mammals have identified a clear link between exposure to hormones, such as testosterone, during critical periods for sexual differentiation of the brain and sex-typical behavior in later life (Hines, 2004), including sexually dimorphic cognitive skills, such as visuospatial ability (Williams, Barnett, & Meck, 1990; Williams & Meck, 1991). In humans, however, the manipulation or even direct measurement of prenatal hormone levels in healthy individuals would be unethical. Consequently, several approaches have been used to indirectly compare variation in prenatal hormone levels with behavior in later life.

One method of examining the effect of prenatal hormones on development is through the study of medical conditions leading to abnormal hormone environments, such as congenital adrenal hyperplasia (CAH). CAH is a genetic disorder affecting both sexes and causes excess adrenal androgen production, beginning prenatally (New, 1998). In general, behavioral studies of females with CAH show masculinization of characteristics typically associated with males compared to unaffected female siblings or sex-matched controls. These include spatial orientation, visualization, targeting, personality, sex-typed play behavior, cognitive abilities, and sexuality (Hampson, Rovet, & Altmann, 1998; Hines et al., 2003; Nordenstrom, Servin, Bohlin, Larsson, & Wedell, 2002; Pasterski et al., 2005; Resnick, Berenbaum, Gottesman, & Bouchard, 1986). In addition, research indicates that the more male-typical behavior is not due to parental influences (Nordenstrom et al., 2002; Pasterski et al., 2005).

There is conflicting evidence that girls with CAH have enhanced spatial ability. Some studies have reported significant associations (e.g., Hampson et al., 1998; Hines et al., 2003; Perlman, 1973; Puts, McDaniel, Jordan, & Breedlove, 2008; Resnick et al., 1986), whereas others show no effect (e.g., Baker & Ehrhardt, 1974; Hines et al., 2003; Malouf, Migeon, Carson, Petrucci, & Wisniewski, 2006). While CAH can provide useful insights into the effects of hormones on behavior, the complex nature of side effects of CAH make it difficult to conclude unequivocally that later behavior must reflect the role of particular hormones (Fausto-Sterling, 1992; Quadagno, Briscoe, & Quadagno, 1977).

There is also a growing body of evidence that fetal hormone levels influence certain physical characteristics which can be observed after birth (Cohen-Bendahan et al., 2005; Kimura, 1999). These proxy measurements have been used in lieu of direct measures of prenatal androgen exposure. One such proxy measure is the ratio between the length of the 2nd and 4th digit (2D:4D) of each hand. This ratio is typically lower in males than females. Significant sex differences have been observed by week 14 of fetal life (Galis, Ten Broek, Van Dongen, & Wijnaendts, 2010; Malas, Dogan, Evcil, & Desdicioglu, 2006). It has been hypothesized that 2D:4D ratio might reflect fetal exposure to prenatal sex hormones in early gestation (Manning, 2002), and one study of amniotic measures of sex hormones in humans provides some support for this relationship (Lutchmaya, Baron-Cohen, Raggatt, Knickmeyer, & Manning, 2004).

Studies comparing 2D:4D ratios with performance on mental rotation tests have shown mixed results (Falter, Arroyo, & Davis, 2006; Hampson, Ellis, & Tenk, 2008; Manning & Taylor, 2001). The 2D:4D ratio has been found to predict performance on a computerized targeting task in both children and adults (Falter et al., 2006; Falter, Plaisted, & Davis, 2008). Falter et al. (2006) also investigated the relationship between performance on two computerized EFT tasks and 2D:4D ratios in adults. Results revealed a significant relationship between lower digit ratios on one EFT task, while the second task showed no significant correlation. A second study in children did not find a significant relationship between 2D:4D ratio and EFT scores (Falter et al., 2008). Results from a meta-analysis show little or no relationship between digit ratio and spatial ability in individuals with CAH (Puts et al., 2008).

A further method of investigating the effects of prenatal hormones is the sampling of amniotic fluid surrounding the fetus. The amniotic fluid sample is obtained for clinical (not research) reasons in pregnancies considered to be at risk, to detect genetic abnormalities in the fetus. This procedure (amniocentesis) is typically performed during a relatively narrow time period, coinciding with the hypothesized critical period for human sexual differentiation (approximately weeks 8–24 of gestation) (Hines, 2004; Smail, Reyes, Winter, & Faiman, 1981).

Finegan, Bartleman, and Wong (1989) were the first to propose that amniotic fluid could be assayed to investigate the effect of prenatal sex hormones on postnatal development. Finegan, Niccols, and Sitarenios (1992) conducted the first study which explored the relationship between prenatal hormone levels in amniotic fluid and later behavior on a broad range of cognitive functions. The findings were difficult to interpret since the authors used measures that did not show sex differences. However, the same children were followed up at 7 years of age and associations between mental rotation ability and FT were examined (Grimshaw, Sitarenios, & Finegan, 1995). A significant positive association between FT levels and faster performance on a mental rotation task was found in a subgroup of girls whose response times correlated with angle of rotation, indicating they had used a rotational strategy. A significant relationship between FT and the "attention to detail" subscale of the Autism Spectrum Quotient–Children's Version in 235 children between 6 and

10 years of age has been observed (Auyeung et al., 2009b). Exposure to FT has also been linked to more masculinized play behavior (Auyeung et al., 2009a) and systemizing (the drive to analyze, explore, and construct a system, measured using the children's version of the Systemizing Quotient) in both boys and girls (Auyeung et al., 2006). Intelligence measures were obtained and these results are reported elsewhere (Auyeung et al., 2009b). No relationships between FT level and full scale, performance, or verbal IQ scores were observed.

This study examined whether FT levels, measured in amniotic fluid from second trimester routine amniocentesis, predicted performance on mental rotation, targeting, and disembedding hidden figures—three visuospatial tasks where sex differences have previously been observed. Children whose FT levels had been measured prenatally via amniocentesis were tested on age-appropriate versions of mental rotation and targeting tasks as well as the EFT. Performance on these tasks was compared with amniotic FT level and a range of environmental control variables.

Method

Participants

Sixty-four children (35 boys, 29 girls) between 7 and 10 years of age ($M = 9.05$, $SD = .94$) took part. These children were recruited from a longitudinal study of the effects of fetal testosterone (FT) on child development. Their mothers had undergone routine amniocentesis in the Cambridge region and given birth to healthy singleton infants between 1996 and 1999 (Baron-Cohen et al., 2004).

Measures

Mental Rotation

This computer-based mental rotation task displayed two teddy bears simultaneously on the screen and was based on a previous methodology for testing children (Grimshaw et al., 1995). The bear on the right side of the screen was presented upright or rotated (30, 60, 90, 120, 150, 180°), while the bear on the left remained upright. The child was asked to indicate if the bears were holding up the same or different arm by pressing one of two buttons. Equal numbers of same and different items were presented in random order. The experimental procedure included examples of same/different judgments (4 trials), practice items on bear rotation (18 trials), and the experimental test (42 trials). Both the number of correct responses and the mean time the child took to respond (in milliseconds) were recorded.

Targeting

In this task, children were instructed to use an overhand throw to try to hit the center of a target mounted on a wall (elevated 150 cm and at a distance of 2 m) with a tennis-sized ball. This was an adapted ball-throwing task (for use with young children) where the ball sticks to a target board measuring 60 cm in diameter, and a similar task has been used in a study of CAH (Hines et al., 2003). Ten trials were performed with one hand followed by ten trials with the other hand. The order of the hands was counterbalanced across participants. A score of 3 was given for trials where the child hit the center of the target (20 cm in diameter), a score of 2 was given if the child hit the middle region (40 cm in diameter), a score of 1 was given if the child hit the outer region (60 cm in diameter), and a score of 0 was given in the case of completely missing the target. The child's handedness was recorded and there were 11 left-handed children (8 boys, 3 girls). Both the mean of the right and left hand total scores and the total score of the dominant (preferred) hand were recorded.

Children's EFT (Witkin et al., 1971)

This is a standardized task which requires the child to find a simple tent or house shape in progressively more complex drawings and to trace the shape to indicate its location. The tent series is administered before the house series and included four demonstration items, two practice items, and 11 scored items. If the child completed the tent series, then the more complex house series was administered, which consists of four demonstration items, one practice item, and 14 scored items. The total score represents the number of figures correctly located (maximum score of 25). Testing was discontinued after five consecutive incorrect responses.

FT Levels

The major predictor in this study was FT level in amniotic fluid, measured by radioimmunoassay by the Department of Clinical Biochemistry, Addenbrooke's Hospital, Cambridge. Amniotic fluid was extracted with diethyl ether. The ether was evaporated to dryness at room temperature and the extracted material redissolved in an assay buffer. Testosterone was assayed by the DPC "Count-a-Coat" method (Diagnostic Products Corp, Los Angeles, CA 90045-5597), which uses an antibody to testosterone coated onto propylene tubes and a 125-I labelled testosterone analogue. Units of FT were expressed in nanomoles per litre (nmol/l). The detection limit of the assay using the ether-extraction method is approximately .05 nmol/l. The coefficient of variation (CV) for between batch imprecision is 19% at a concentration of .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 .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 included in all subsequent analyses: Gestational age at amniocentesis (in weeks), maternal age at the time of the child's birth, average level of education obtained by parents according to a 5-point scale rated from 1 (no formal qualifications) to 5 (postgraduate qualification), presence of older brothers in the home (or not), presence of older sisters in the home (or not), and child's age. These control variables aim to account for the some of the individual variations in the developmental environment (Auyeung et al., 2009b).

Procedure

Invitations for cognitive testing were sent to all 452 participating mothers. No significant differences were observed for the predictor or control variables between the children who participated in this study and the remainder of the group initially contacted (see Table 1).

This study was given ethical approval by the National Health Service Suffolk Research Ethics Committee. Written informed consent was obtained from General Practitioners and participating parents.

Results

Table 2 shows the means and *SD* for each sex separately, as well as combined for all variables. Effect sizes for sex differences were computed using Cohen's *d* (Cohen, 1988).

Inspection of the univariate distributions revealed that the distributions were not significantly skewed (skewness < 1). One male outlier and no female outliers were found for FT level. Analyses involving testosterone in males were conducted using the full dataset.[1] No outliers were observed for the other predictor or outcome variables. No significant sex differences were observed for any of the control variables.

Results showed a strong correlation between measured FT levels and sex ($r = .59$, $df = 62$, $p < .001$), with boys ($M = .79$, $SD = .37$) showing higher FT levels than girls ($M = .34$, $SD = .23$), $t(62) = 5.72$, $p < .001$, $d = 1.46$. Table 3 shows correlation coefficients for each variable for boys and girls separately.

Mental Rotation

Examination of the univariate distribution for mental rotation correct score and response times revealed that they were not significantly skewed (skewness < 1) for all cases together as well as in boys and girls separately.

A significant sex difference was found in the number of correct responses on the mental rotation task between boys ($M = 31.89$, $SD = 6.46$) and girls ($M = 26.93$, $SD = 9.70$), $t(62) = 2.44$, $p < .05$, equal variances, $d = .60$. No significant sex differences were found between boys ($M = 2775.21$, $SD = 766.67$) and girls ($M = 2847.36$, $SD = 888.48$) for mental rotation mean response time $t(62) < 1$, equal variances, $d = .09$.

Table 1 Comparison of participants versus non-participants

	Participants			Non-participants			*t*	*p*
	N	*M*	*SD*	*N*	*M*	*SD*		
FT level (nmol/l)	64	.59	.39	387	.60	.42	<1	ns
Gestational age	54	16.29	1.48	293	16.50	1.79	<1	ns
Mother age	64	35.39	4.30	353	35.60	4.71	<1	ns
Parental education level	61	3.37	.95	308	3.12	1.02	1.74	ns
Older sister	64	.08	.27	391	.11	.31	<1	ns
Older brother	64	.13	.33	391	.09	.29	<1	ns
Child number of siblings	61	1.48	1.03	312	1.33	.91	1.14	ns

Child sex ($r = .30$, $df = 62$, $p < .05$) and child age ($r = .41$, $df = 62$, $p < .001$) showed significant correlations with mental rotation scores. A hierarchical regression analysis was conducted including child age in the first stage using the enter method. Child sex and FT level were included in the second stage using the stepwise method (entry criterion $p < .05$, removal criterion $p > .10$) (Altman, 1991). The interaction between FT level and sex was tested for inclusion in the third stage using the stepwise method. The final regression model retained both child sex and child age, $F(2, 61) = 9.48$, $p < .001$, $R^2 = .24$. See Fig. 1 for a visual representation of the relationship between (a) mental rotation scores and child age and (b) mental rotation scores and FT levels.

In girls, the only significant correlation between the number of correct responses on the mental rotation task and the predictor variables was with child age ($r = .58$, $df = 27$, $p = .001$), which was a significant predictor of scores, $F(1, 27) = 13.73$, $p < .001$, $R^2 = .34$. In boys, no significant correlations were observed between mental rotation and any predictor variables. Table 4 shows the final regression model for correct scores on the mental rotation task for the sexes together as well as separately.

An analysis of the response times for the mental rotation task revealed no significant relationships with any of the predictor variables when the sexes were examined together or separately (all $ps > .05$).

Previous studies have compared an individual's response time with the angle of rotation for each item to establish whether the individual is using a rotational strategy (Brosnan, Daggar, & Collomosse, 2010; Falter et al., 2006, 2008; Grimshaw et al., 1995; Hooven, Chabris, Ellison, & Kosslyn, 2004). It has been suggested that the response time consists of a non-rotational component (which is the same for all items) and a further rotational component, which varies according to the angle of rotation. For each participant, response times for correct responses were regressed linearly to obtain a rotation slope (argued to represent the rotational aspects of the task) and the intercept (argued to represent the non-rotational aspects of the task) (Brosnan et al., 2010; Falter et al., 2006, 2008; Grimshaw et al., 1995; Hooven et al., 2004). No significant relationships were found

---

[1] The results excluding the one outlier in FT level did not differ significantly from those obtained with the full dataset. The full results of these analyses are available from the corresponding author upon request.

Table 2 Descriptive statistics

Variable	Girls				Boys				Cohen's *d*
	n	*M*	*SD*	Range	*n*	*M*	*SD*	Range	
FT level (nmol/l)[a],**	29	.34	.23	.05–1.00	35	.79	.37	.13–1.95	1.46
Gestational age	26	16.17	1.17	14–19	32	16.04	1.47	13–20	.01
Child age	28	8.97	1.03	7.11–10.42	35	9.11	.87	7.03–10.66	.15
Maternal age	28	35.41	4.65	25.55–45.66	33	35.42	4.20	29.57–43.95	.00
Parent education	29	3.21	.75	2–5	33	3.50	1.08	2–5	.31
EFT**	29	9.86	4.68	2–22	35	15.06	4.38	6–23	1.15
Mean targeting score	29	16.40	4.34	7–24	35	17.84	4.21	5–24	.34
Dominant hand total score	29	18.34	5.26	7–27	35	19.23	5.13	6–27	.17
Mental rotation*	29	26.93	9.70	5–42	35	31.89	6.46	18–40	.60
Mental rotation time (ms)	29	2847.36	888.48	962–4197	35	2775.21	766.67	700–4464	.09

[a] Indicates raw values

* Sex difference significant at $p<.05$

** Sex difference significant at $p<.01$

Table 3 Correlation matrix for girls (above diagonal) and boys (below diagonal)

	FT level	Gest. age	Child age	Maternal age	Parent education	Older sister	Older brother	EFT	Targeting	Targeting dominant	Mental rotation	Mental rotation time (ms)
FT level	–	−.17	−.17	−.08	−.16	.24	.51**	.71**	.03	.04	−.03	.00
Gest. age	−.17	–	.19	−.38	.14	−.17	−.33	−.21	.28	.39	.30	−.16
Child age	−.21	−.11	–	−.25	−.18	.12	−.29	.06	.55**	.60**	.58**	−.26
Maternal age	−.04	−.13	.22	–	−.14	.1	.15	.04	−.05	−.16	−.30	.19
Parent education	−.01	−.03	−.04	.01	–	.11	−.12	−.28	−.17	−.14	−.34	−.17
Older sister	.17	−.2	.14	.15	−.10	–	.29	.13	.25	.14	.03	−.11
Older brother	−.03	−.17	−.14	−.01	−.04	.53**	–	.36	−.06	−.20	−.29	.27
EFT	.54**	−.33	.09	.20	.21	.21	−.07	–	.09	.00	.13	.14
Targeting	−.04	.01	.30	.17	.02	−.08	.11	.12	–	.85**	.18	−.32
Targeting dominant	−.10	.24	.13	.18	.01	−.05	.16	−.03	.84**	–	.28	−.33
Mental rotation	−.10	−.05	.16	.12	.26	−.01	−.08	.15	−.03	−.03	–	−.17
Mental rotation time (ms)	−.12	.33	−.08	−.10	.24	−.15	−.23	.15	−.12	−.04	.29	–

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

between FT levels and the slopes ($r=.16$, $df=62$) or intercepts ($r=-.06$, $df=62$).

Finally, the correlation between reaction time (for correct responses) and angle of rotation was calculated for each participant, indicating the extent to which an individual used a rotational strategy. Examination of the distribution of correlation coefficients did not reveal the clear distinction between individuals using rotational or non-rotational strategies observed by Grimshaw et al. (1995). Median split of the response time/angle of rotation correlation coefficients was used to identify individuals who may have used a more rotational strategy. No sex difference was observed in the number of individuals using a rotational strategy, $t(62)<1$. There were also no significant correlations between response times and any of the predictor variables for this group.

Targeting

Examination of the univariate distributions for targeting revealed that they were not significantly skewed (skewness < 1) for all cases together as well as in boys and girls separately. For mean right and left hand targeting score, no significant sex differences were found between boys ($M=17.84$, $SD=4.21$) and girls ($M=16.40$, $SD=4.34$), $t(62)<1$, equal variances, $d=.34$. There were also no significant correlations with any of the

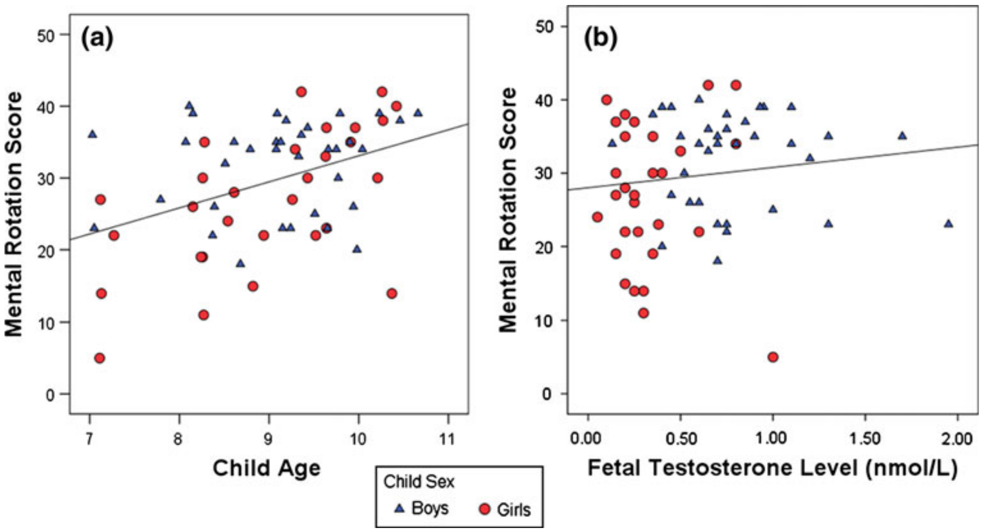

Fig. 1 Scatter plots showing the relationship between mental rotation scores and child age (**a**) and mental rotation scores and FT levels (**b**)

Table 4 Final regression model for number of correct responses on the mental rotation task

Outcome	Predictors	Final regression model					
		R^2	ΔR^2	B	SE	*β*	*p*
Group							
Mental rotation	Child age	.17	.17	3.45	1.00	.39	.001
	Child sex	.24	.07	2.23	.94	.27	<.05
Girls only							
Mental rotation	Child age	.34	.34	5.47	1.48	.58	.001
Boys only							
Mental rotation	No significant predictors						

predictor variables except child age ($r = .43$, $df = 62$, $p < .001$). There was a significant negative correlation between age and the number of missed trials, demonstrating that the younger the child was, the more missed trials they tended to have ($r = -.41$, $df = 62$, $p < .001$). See Fig. 2 for a visual representation between mean targeting score, child's age, and FT levels.

Dominant (preferred) hand targeting score also showed no significant sex difference between boys ($M = 19.23$, $SD = 5.13$) and girls ($M = 18.34$, $SD = 5.26$), $t(62) < 1$, equal variances, $d = .17$. Dominant hand targeting showed no significant correlations with any of the variables except child age ($r = .36$, $df = 62$, $p < .001$), therefore regression analyses were not conducted.

Children's EFT

For EFT score, examination of the univariate distribution revealed that it was not significantly skewed (skewness < 1) for all participants together as well as for boys and girls separately. A significant sex difference was found in EFT scores with boys ($M = 15.06$, $SD = 4.38$) scoring higher than girls ($M = 9.86$, $SD = 4.68$), $t(62) = 4.58$, $p < .001$, equal variances, $d = 1.15$.

A hierarchical multiple regression analysis was used to explore the contributions of the predictor variables to variation in the outcome variables. In the first stage, any predictor variable that showed a significant correlation with the outcome at $p < .20$ was entered into the analysis (Altman, 1991). In addition, the influence of suppressor variables (predictors that were 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 stage using the stepwise method (entry criterion $p < .05$, removal criterion $p > .10$). The interaction between FT level and child sex was tested for inclusion in the third stage using the stepwise method.

The predictor variables that correlated with EFT scores at $p < .20$ were gestational age ($r = -.26$, $df = 62$, $p < .20$) and the presence of older sister(s) ($r = .16$, $df = 62$, $p < .20$) and brother(s) (suppressor). These were included in the regression analysis using the enter method in the first stage. The inclusion of FT level in the second stage produced a significant *F*-change, $F(1, 43) = 31.51$, $p < .001$, $\Delta R^2 = .38$. Child sex and the sex and FT level interaction were excluded from the final regression model. Figure 3 shows the relationship between (a) EFT scores and child age and (b) EFT scores and FT levels.

Within sex analyses showed that FT levels were significantly related to EFT scores in both girls ($r = .71$, $df = 27$, $p < .001$) and boys ($r = .54$, $df = 33$, $p = .001$). The regression analysis for girls included parent education level ($r = -.28$, $df = 26$, $p < .20$) and the presence of older brother(s) ($r = .36$, $df = 27$, $p < .20$) and sister(s) (suppressor) in the first stage using the enter method. The inclusion of FT level in the second stage produced a significant *F*-change, $F(1, 23) = 12.99$, $p = .001$, $\Delta R^2 = .28$. The first stage of the regression analysis for boys included gestational age ($r = -.33$, $df = 25$, $p < .20$) and results showed a significant *F*- change, $F(1, 24) = 5.05$, $p < .05$, $\Delta R^2 = .16$ when FT level was included in the second stage. Results from the EFT regression analysis are shown in Table 5.

Fig. 2 Scatter plots showing the relationship between mean targeting scores and child age (**a**) and mean targeting scores and FT levels (**b**)

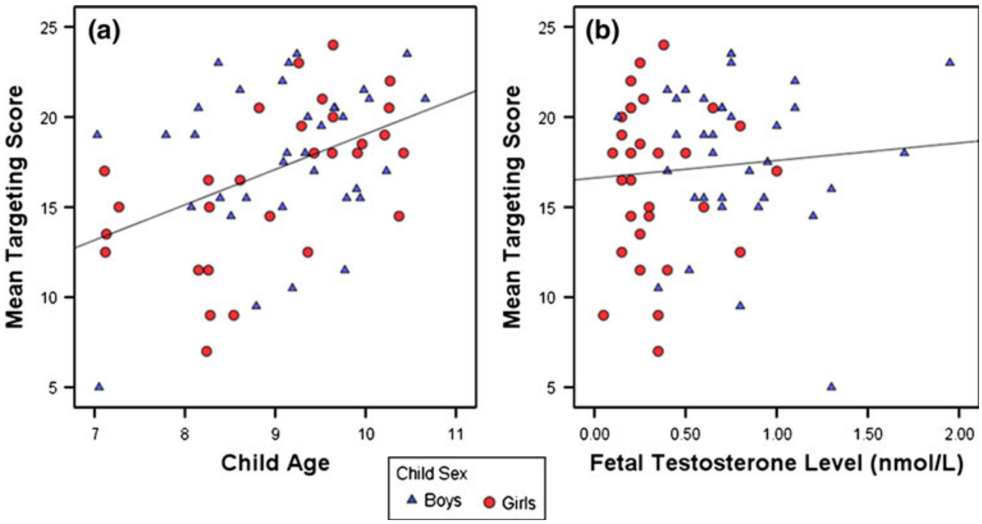

Fig. 3 Scatter plots showing the relationship between the children's EFT scores and child age (**a**) and the children's EFT scores and FT levels (**b**)

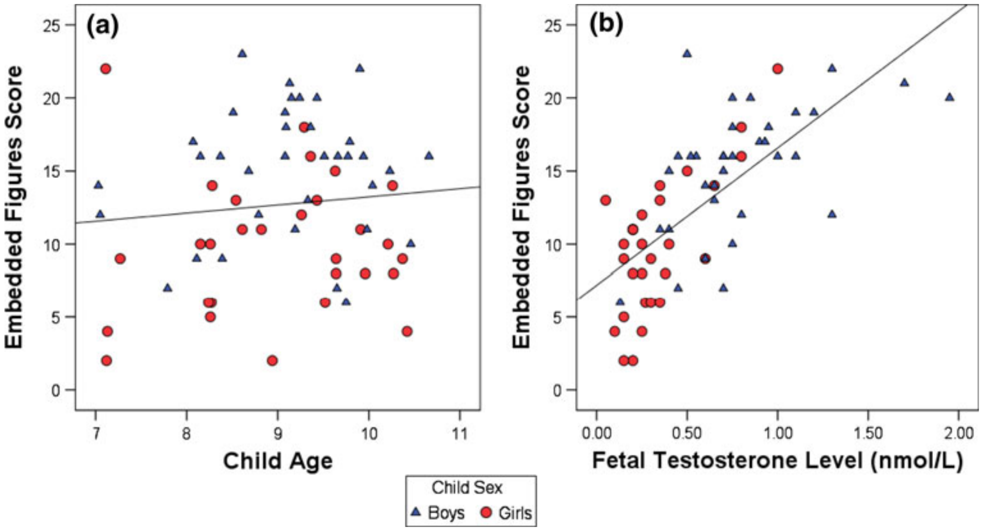

Table 5 Final regression model for the children's EFT scores

Outcome	Predictors	Final regression model					
		R^2	ΔR^2	B	SE	β	p
Group							
EFT	Gestational age	.11	.11	−.61	.46	.15	ns
	Older sister			2.51	3.12	.09	ns
	Older brother			−.69	2.03	.04	ns
	FT level	.48	.38	11.69	1.47	.63	<.001
Girls only							
EFT	Parent education	.22	.22	−1.07	.93	.17	ns
	Older sister			−.36	2.78	.02	ns
	Older brother			−.18	2.39	.01	ns
	FT level	.50	.28	14.32	3.97	.67	<.001
Boys only							
EFT	Gestational age	.11	.11	−.78	.52	−.27	ns
	FT level	.27	.16	4.53	2.02	.40	<.05

Discussion

This study examined the relationship between prenatal hormone levels measured from amniotic fluid (collected during the second trimester of pregnancy) and performance in three areas of visuospatial ability which have previously shown significant sex differences. FT levels were significantly higher in boys, where the average level was approximately 2.3 times that found in girls. Removing a single male outlier did not significantly change these results. The sex difference in FT levels was in line with previous studies using amniotic measures of testosterone (Beck-Peccoz et al., 1991; Bergman, Glover, Sarkar, Abbott, & O'Connor, 2010; Sarkar, Bergman, Fisk, O'Connor, & Glover, 2007; van de Beek, Thijssen, Cohen-Kettenis, van Goozen, & Buitelaar, 2004; van de Beek, van Goozen, Buitelaar, & Cohen-Kettenis, 2008).

For the mental rotation test, a significant sex difference favoring males was found in the number of correct items the child attained, but no significant sex difference was found in the mean time the child took to respond. Within sex analysis showed no significant predictors for boys. Age was a significant predictor of mental rotation scores in girls and results showed an advantage for older children.

No significant relationship was observed between mental rotation ability and FT levels when all participants were examined together or when examined separately by sex or rotation strategy. These findings were consistent with previous studies of mental rotation showing no relationships with 2D:4D ratio (Falter et al., 2006; Hampson et al., 2008). The current findings were also consistent with evidence suggesting that girls with elevated exposure to FT as a result of CAH were not faster or more accurate at mental rotation (Baker & Ehrhardt, 1974; Hines et al., 2003; Malouf et al., 2006). However, other studies have found significantly enhanced performance on mental rotation tests in girls with CAH (Hampson et al., 1998; Perlman, 1973; Resnick et al., 1986). These findings underline the importance of task method and participant selection for this task, including factors such as control group used, age at testing, and sample sizes (for detailed reviews, see Hines, 2004; Puts et al., 2008).

Targeting ability was found to increase significantly with age and no significant sex differences or associations with FT were found. Another study by Hines et al. (2003) found that females exposed to elevated FT levels as a result of CAH performed better than unaffected relatives in targeting. While the task used in the current study was similar to that employed by Hines et al. (2003), participants in the latter study were between 12-45 years of age (*M* age, 19.5). Other studies which have shown clear sex differences in targeting ability have also focused on older age groups. Future studies looking at targeting ability in children may wish to focus on a more narrow age range of participants in order to reduce the overall effect on scores. In addition, a continuous measure of targeting accuracy (e.g., cm from target) could be used to provide a more sensitive assessment than the discrete scoring system used in this study. Such a measure would help identify any correlations with predictor variables.

EFT scores showed a clear advantage for boys, consistent with previous research showing superior male performance on this task (Nebot, 1988), although not all studies of this age group have found this sex difference (Bigelow, 1971).

FT levels were found to be a significant predictor of EFT scores (see Fig. 3) for all participants together and also when boys and girls were examined separately. The task did not show a significant correlation with age or with any of the other predictor variables.

The link between FT from amniotic fluid and performance on the EFT has not previously been examined. However, the current findings were consistent with measurements indicating a link between higher FT levels and increased attention to detail as well as increased systemizing ability in children (Auyeung et al., 2006, 2009b). These findings also support results by Resnick et al. (1986), who found that girls with CAH showed enhanced performance on the Hidden Figures test. However, they were not consistent with a study showing girls with CAH scored lower than unaffected women on the “attention to detail” subscale of the Autism Spectrum Quotient (AQ) (Knickmeyer et al., 2006). The “attention to detail” subscale of the AQ contains items specifically pertaining to interest in numbers and dates, whereas there are few items tapping visual details which may be more relevant to the EFT. Together, these studies warrant further investigation into which cognitive abilities are involved in EFT performance.

The finding that amniotic fluid FT level was significantly related to EFT score but not the other tasks examined in this study also requires further investigation. At first glance, all three of these tasks examine areas of visuospatial ability which have previously shown sex differences. However, the design of tasks for children is often very difficult. In the case of the EFT, the specific task used in this study has been widely employed and is non-computerized. The other tasks, however, were adapted versions of measures used by other researchers.

A study by Falter et al. (2006) highlighted the difficulty of examining cognitive abilities by comparing performance on two slightly different measures of figure disembedding with respect to 2D:4D ratio. The two tasks only differed in their use of color and visual prompts. A significant relationship was observed using one task and a non-significant relationship with the second task.

Both the targeting and mental rotation tasks used in this study showed positive correlations with age, indicating that the adapted tasks were too difficult for the younger children. In addition, the age range of participants was relatively large. This is an inherent problem with studies looking at amniotic fluid samples because of the difficulty in recruiting large sample sizes within a narrow time frame, since only a small proportion of pregnant women will undergo amniocentesis.

The computerized nature of the mental rotation task may also have been a factor. Previous studies which found a significant sex difference in preadolescent children for mental rotation have used non-computerized tasks (Johnson & Meade, 1987; Kerns & Berenbaum, 1991; Levine et al., 1999). Future studies may wish to identify specific tasks which have shown a clear sex difference and compare performance on these tasks with prenatal hormone levels.

The assessment of other factors that may influence visuospatial ability was beyond the scope of this study and future studies could further investigate other influences, such as effects of increased videogame usage, which has been shown to reduce gender differences in spatial ability (Feng, Spence, & Pratt, 2007; Terlecki, Newcombe, & Little, 2008).

Finally, it should be noted that the measurement of FT levels via amniotic fluid has some limitations. Amniotic fluid samples do not provide an indication of overall fetal exposure to testosterone throughout the pregnancy. Methods using this approach are generally restricted to measurement at a single point in time due to the hazardous nature of the procedure. While it would ideally be possible to make direct measurements of testosterone at regular intervals throughout gestation and into postnatal life, amniocentesis itself carries a risk of causing miscarriage (of about 1%) (d'Ercole et al., 2003; Sangalli, Langdana, & Thurlow, 2004) and should therefore only be carried out for clinical purposes.

The time point at which amniocentesis takes place could have a significant impact on research findings. Given the reported time course of testosterone secretion, the most promising time to measure FT is probably at prenatal weeks 8 to 24 (Smail et al., 1981), but this is still a relatively wide range. In addition, research on non-human primates has shown that androgens masculinize different behaviors at different times during gestation, suggesting different behaviors may also have different sensitive periods for development (Goy, Bercovitch, & McBrair, 1988). For all these reasons, the inferences we can draw about the single measurement of FT are necessarily limited. Given these limitations, however, it is probable that any significant correlation observed between amniotic FT and behavior will be conservative.

This sample of children whose mothers had undergone amniocentesis is also not necessarily representative of the wider population (e.g., maternal age is usually higher in such samples, though no relationship with the outcome variables was found in this study). Amniocentesis has a number of strengths, however, which mainly concern its timing and direct, quantifiable measurement of the fetal environment while avoiding unnecessary additional risk.

In conclusion, the results from this study highlight a clear link between FT measured in amniotic fluid and performance on an EFT at ages 7–10 years. This task also showed a large sex difference, with males obtaining higher scores. These correlations were in line with previous studies linking both amniotic and proxy measures of FT to superior performance on tasks which show a clear male advantage. These findings support the theory that FT has an organizational effect in influencing sexually dimorphic cognitive behavior in later life.

The measures of visuospatial ability in mental rotation and targeting included in this study did not, however, show any significant variation with amniotic FT level. These tasks were adapted from previous studies and did not display the large effect size for sex seen on the EFT. For girls, performance on both of these tasks was significantly affected by age.

Future studies need to examine the influence of amniotic measures of FT on other aspects of cognition as well as how prenatal testosterone might interact with testosterone exposure at other points in the life cycle (e.g., pre- and post-puberty) to explain variation in visuospatial ability. It will also be important to explore the underlying causes of the variation in prenatal hormone levels observed between individuals, which are likely to be genetic (Chakrabarti et al., 2009). Other genetic factors may include hormone receptor sensitivity and distribution or the rate of hormone metabolism. Models which combine FT as well as genetic variant measurements may also allow more powerful examination of how testosterone exposure explains variation in visuospatial abilities.

Acknowledgments This work was supported by grants to SBC from the MRC and the Nancy Lurie-Marks Family Foundation. BA 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, Greg Davis, and Ieuan Hughes for valuable discussions. This study was conducted in association with the NIHR CLAHRC for Cambridgeshire and Peterborough NHS Foundation Trust.

References

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

Auyeung, B., Baron-Cohen, S., Ashwin, E., Knickmeyer, R., Taylor, K., Hackett, G., et al. (2009a). Fetal testosterone predicts sexually differentiated childhood behavior in girls and in boys. *Psychological Science, 20*, 144–148.

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

Auyeung, B., Baron-Cohen, S., Chapman, E., Knickmeyer, R., Taylor, K., & Hackett, G. (2009b). Fetal testosterone and autistic traits. *British Journal of Psychology, 100*, 1–22.

Baker, S. W., & Ehrhardt, A. A. (1974). Prenatal androgen, intelligence and cognitive sex differences. In R. C. Friedman, R. M. Richart, & R. L. Van de Wiele (Eds.), *Sex differences in behavior* (pp. 53–76). New York: Wiley.

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

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.

Beck-Peccoz, P., Padmanabhan, V., Baggiani, A. M., Cortelazzi, D., Buscaglia, M., Medri, G., et al. (1991). Maturation of hypothalamic-pituitary-gonadal function in normal human fetuses: Circulating levels of gonadotropins, their common alpha-subunit and free testosterone, and discrepancy between immunological and biological activities of circulating follicle-stimulating hormone. *Journal of Clinical Endocrinology and Metabolism, 73*, 525–532.

Bergman, K., Glover, V., Sarkar, P., Abbott, D. H., & O'Connor, T. G. (2010). In utero cortisol and testosterone exposure and fear reactivity in infancy. *Hormones and Behavior, 57*, 306–312.

Bigelow, G. (1971). Field dependence-field independence. *Journal of Educational Research, 64*, 397–400.

Brosnan, M., Daggar, R., & Collomosse, J. (2010). The relationship between systemising and mental rotation and the implications for the extreme male brain theory of autism. *Journal of Autism and Developmental Disorders, 40*, 1–7.

Chakrabarti, B., Dudbridge, F., Kent, L., Wheelwright, S., Hill-Cawthorne, G., Allison, C., et al. (2009). Genes related to sex steroids, neural growth, and social-emotional behavior are associated with autistic traits, empathy, and Asperger syndrome. *Autism Research, 2*, 157–177.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

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.

d'Ercole, C., Shojai, R., Desbriere, R., Chau, C., Bretelle, F., Piechon, L., et al. (2003). Prenatal screening: Invasive diagnostic approaches. *Child's Nervous System, 19*, 444–447.

Falter, C. M., Arroyo, M., & Davis, G. J. (2006). Testosterone: Activation or organization of spatial cognition? *Biological Psychology, 73*, 132–140.

Falter, C. M., Plaisted, K. C., & Davis, G. (2008). Visuo-spatial processing in autism—testing the predictions of extreme male brain theory. *Journal of Autism and Developmental Disorders, 38*, 507–515.

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

Feng, J., Spence, I., & Pratt, J. (2007). Playing an action video game reduces gender differences in spatial cognition. *Psychological Science, 18*, 850–855.

Finegan, J., Bartleman, B., & Wong, P. Y. (1989). A window for the study of prenatal sex hormone influences on postnatal development. *Journal of Genetic Psychology, 150*, 101–112.

Finegan, J. K., Niccols, G. A., & Sitarenios, G. (1992). Relations between prenatal testosterone levels and cognitive abilities at 4 years. *Developmental Psychology, 28*, 1075–1089.

Galis, F., Ten Broek, C. M., Van Dongen, S., & Wijnaendts, L. C. (2010). Sexual dimorphism in the prenatal digit ratio (2D:4D). *Archives of Sexual Behavior, 39*, 57–62.

Goy, R. W., Bercovitch, F. B., & McBrair, M. C. (1988). Behavioral masculinization is independent of genital masculinization in prenatally androgenized female rhesus macaques. *Hormones and Behavior, 22*, 552–571.

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.

Halpern, D. F. (2000). *Sex differences in cognitive abilities* (3rd ed.). Mahwah, NJ: Erlbaum.

Hampson, E., Ellis, C. L., & Tenk, C. M. (2008). On the relation between 2D:4D and sex- dimorphic personality traits. *Archives of Sexual Behavior, 37*, 133–144.

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.

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

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.

Hooven, C. K., Chabris, C. F., Ellison, P. T., & Kosslyn, S. M. (2004). The relationship of male testosterone to components of mental rotation. *Neuropsychologia, 42*, 782–790.

Johnson, E. S., & Meade, A. C. (1987). Developmental patterns of spatial ability: An early sex difference. *Child Development, 58*, 725–740.

Karp, S. A., & Konstadt, N. L. (1963). *Manual for the Children's Embedded Figures Test*. Brookland, NY: Cognitive Tests.

Kerns, K. A., & Berenbaum, S. A. (1991). Sex differences in spatial ability in children. *Behavior Genetics, 21*, 383–396.

Kimura, D. (1999). *Sex and cognition*. Cambridge, MA: The MIT Press.

Knickmeyer, R., Baron-Cohen, S., Fane, B. A., Wheelwright, S., Mathews, G. A., Conway, G. S., et al. (2006). Androgens and autistic traits: A study of individuals with congenital adrenal hyperplasia. *Hormones and Behavior, 50*, 148–153.

Levine, S. C., Huttenlocher, J., Taylor, A., & Langrock, A. (1999). Early sex differences in spatial skill. *Developmental Psychology, 35*, 940–949.

Linn, M. C., & Petersen, A. C. (1986). Gender differences in spatial ability: Implications for mathematics and science achievement. In J. S. Hyde & M. C. Linn (Eds.), *The psychology of gender: Advances through meta-analysis* (pp. 67–101). Baltimore: Johns Hopkins University Press.

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.

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.

Malouf, M. A., Migeon, C. J., Carson, K. A., Petrucci, L., & Wisniewski, A. B. (2006). Cognitive outcome in adult women affected by congenital adrenal hyperplasia due to 21-hydroxylase deficiency. *Hormone Research, 65*, 142–150.

Manning, J. T. (2002). *Digit ratio: A pointer to fertility, behavior and health*. New Brunswick, NJ: Rutgers University Press.

Manning, J. T., & Taylor, R. P. (2001). Second to fourth digit ratio and male ability in sport: Implications for sexual selection in humans. *Evolution and Human Behavior, 22*, 61–69.

Nebot, T. K. (1988). Sex differences among children on embedded tasks. *Perceptual and Motor Skills, 67*, 972–974.

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

Nordenstrom, A., Servin, A., Bohlin, G., Larsson, A., & Wedell, A. (2002). Sex-typed toy play behavior correlates with the degree of prenatal androgen exposure assessed by the *CYP21* genotype in girls with congenital adrenal hyperplasia. *Journal of Clinical Endocrinology & Metabolism, 87*, 5119–5124.

Parlee, M. B., & Rajagopal, J. (1974). Sex differences on the Embedded-Figures Test: A cross-cultural comparison of college students in India and in the United States. *Perceptual and Motor Skills, 39*, 1311–1314.

Pasterski, V. L., Geffner, M. E., Brain, C., Hindmarsh, P., Brook, C., & Hines, M. (2005). Prenatal hormones and postnatal socialization by parents as determinants of male-typical toy play in girls with congenital adrenal hyperplasia. *Child Development, 76*, 264–278.

Perlman, S. M. (1973). Cognitive abilities of children with hormone abnormalities: Screening by psychoeducational tests. *Journal of Learning Disabilities, 6*, 26–34.

Puts, D. A., McDaniel, M. A., Jordan, C. L., & Breedlove, S. M. (2008). Spatial ability and prenatal androgens: meta-analyses of congenital adrenal hyperplasia and digit ratio (2D:4D) studies. *Archives of Sexual Behavior, 37*, 100–111.

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.

Sangalli, M., Langdana, F., & Thurlow, C. (2004). Pregnancy loss rate following routine genetic amniocentesis at Wellington Hospital. *New Zealand Medical Journal, 117*, U818.

Sarkar, P., Bergman, K., Fisk, N. M., O'Connor, T. G., & Glover, V. (2007). Amniotic fluid testosterone: Relationship with cortisol and gestational age. *Clinical Endocrinology, 67*, 743–747.

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.

Shah, A., & Frith, U. (1983). An islet of ability in autistic children: A research note. *Journal of Child Psychology and Psychiatry, 24*, 613–620.

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: Martinus Nijhoff.

Terlecki, M. S., Newcombe, N. S., & Little, M. (2008). Durable and generalized effects of spatial experience on mental rotation: Gender differences in growth patterns. *Applied Cognitive Psychology, 22*, 996–1013.

Thomas, J. R., & French, K. E. (1985). Gender differences across age in motor performance a meta-analysis. *Psychological Bulletin, 98*, 260–282.

van de Beek, C., Thijssen, J. H. H., Cohen-Kettenis, P. T., van Goozen, S. H., & Buitelaar, J. K. (2004). Relationships between sex hormones assessed in amniotic fluid, and maternal and umbilical cord blood: What is the best source of information to investigate the effects of fetal hormonal exposure? *Hormones and Behavior, 46*, 663–669.

van de Beek, C., van Goozen, S. H. M., Buitelaar, J. K., & Cohen-Kettenis, P. T. (2008). Prenatal sex hormones (maternal and amniotic fluid) and gender-related play behavior in 13-month-old infants. *Archives of Sexual Behavior, 38*, 6–15.

Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables. *Psychological Bulletin, 117*, 250–270.

Watson, N. V., & Kimura, D. (1991). Nontrivial sex differences in throwing and intercepting: Relation to psychometrically-defined spatial functions. *Personality and Individual Differences, 12*, 375–385.

Williams, C. L., Barnett, A. M., & Meck, W. H. (1990). Organizational effects of early gonadal secretions on sexual differentiation in spatial memory. *Behavioral Neuroscience, 104*, 84–97.

Williams, C. L., & Meck, W. H. (1991). The organizational effects of gonadal steroids on sexually dimorphic spatial ability. *Psychoneuroendocrinology, 16*, 155–176.

Witkin, H. A., Oltman, P. K., Raskin, E., & Karp, S. (1971). *A Manual for the Embedded Figures Test*. Palo Alto, CA: Consulting Psychology Press.