

# **Heritability of testosterone levels in 12-year-old twins and its relation to pubertal development**

**Rosa A. Hoekstra<sup>\*</sup>, Meike Bartels, Dorret I. Boomsma**

Department of Biological Psychology, Vrije Universiteit,  
Van der Boechorststraat 1, 1081 BT Amsterdam, the Netherlands

Running title: Heritability of testosterone levels in early puberty

Published as: Rosa A. Hoekstra, Meike Bartels, Dorret I. Boomsma (2006). Heritability of testosterone levels in 12-year-old twins and its relation to pubertal development. *Twin Research and Human Genetics*, 9, 558-65.

<sup>\*</sup>To whom correspondence should be addressed.

Tel: 0031 20 5988363, Fax: 0031 20 5988832, E-mail: Ra.Hoekstra@psy.vu.nl.

## **Abstract**

The aim of this study was to estimate the heritability of variation in testosterone levels in 12-year-old children, and to explore the overlap in genetic and environmental influences on circulating testosterone levels and androgen dependent pubertal development. Midday salivary testosterone samples were collected on two consecutive days in a sample of 183 unselected twin pairs. Androgen induced pubertal development was assessed using self report Tanner scales of pubic hair development (boys and girls) and genital development (boys). A significant contribution of genetic effects to the variance in testosterone levels was found. Heritability was approximately 50% in both boys and girls. The remaining proportion of the variance in testosterone levels could be explained by non-shared environmental influences. The relatively high correlation between testosterone levels of opposite sex dizygotic twins suggests that sex differences in genes influencing variation in testosterone levels have not yet developed in pre- and early puberty. Variance in pubertal development was explained by a large genetic component, moderate shared environmental influences, and a small non-shared environmental effect. Testosterone levels correlated moderately ( $r = .31$ ) with pubertal development; the covariance between testosterone levels and pubertal development was entirely accounted for by genetic influences.

Puberty is the hallmark period in which a child undergoes the transition to adulthood. Driven by hormonal changes, major physical changes occur during this period of development. Puberty starts with the activation of the hypothalamic – pituitary-gonadal (HPG) axis by an increase of the pulsatile release of gonadotropin releasing hormone (GnRH) in the hypothalamus (Sisk & Foster, 2004; Grumbach & Styne, 2003). GnRH induces secretion of luteinizing hormone and follicle stimulating hormone in the pituitary, that in turn direct the testes and ovaries to produce sperm and eggs, and activate secretion of steroid hormones. Typically a few years prior to these hormonal changes, children experience the adrenarche, a process biologically independent from the activation of the HPG-axis. Adrenarche results from the activation of the hypothalamic-pituitary-adrenal axis, leading to secretion of dehydroepiandrosterone (DHEA), DHEA-sulphate and androstenedione by the adrenal glands (Auchus & Rainey, 2004; Grumbach & Styne, 2003). Extraglandular conversion of androstenedione leads to increasing levels of the androgens testosterone and dihydrotestosterone.

The increasing levels of circulating hormones lead to different physical changes in puberty. In girls, breast development is induced by increasing estrogen levels, excreted by the ovaries. The development of pubic hair is mainly influenced by androgens, originating from both the ovaries and the adrenal gland. In boys, both genital and pubic hair development are under androgen control (Grumbach & Styne, 2003). Most of the circulating testosterone levels in boys are secreted by the testes. In addition to direct secretion, a small amount of testosterone is derived from peripheral conversion of androstenedione secreted by the testes and adrenal glands (Grumbach & Styne, 2003).

Variation in pubertal timing and individual differences in circulating sex hormone levels are of considerable interest to both the fields of medicine and psychology. Early maturation in girls is found to be associated with internalizing symptoms and disorders, and

increased rates of alcohol, tobacco and substance use (Hayward & Sanborn, 2002; Graber et al., 2004; Grumbach & Styne, 2003). In boys, late maturation may be a risk factor for deviant behavior and substance abuse (Graber et al., 2004). Furthermore, individual differences in circulating sex hormone levels are thought to play a role in behavioral problems. An extensive line of research suggests that high testosterone concentrations are linked to (problem) behavior, particularly aggression (Archer, 1991; 2006; Book et al., 2001) and dominance (Mazur & Booth, 1998; Archer, 2006). The causality of this relationship is unknown. High testosterone levels could increase the risk of behavioral problems, but behavior itself can also influence circulating testosterone levels (Archer, 2006; Raine, 2002).

Twin studies show that timing of puberty is influenced by both genetic and environmental effects, with heritability estimates ranging from 50 to 80% (Palmert & Boepple, 2006; van den Berg et al., 2006; Eaves et al., 2004; Mustanski et al., 2004). Molecular and animal studies suggest polygenic effects. Recent studies indicated that the kiss 1 gene, encoding for a family of neuropeptides called kisspeptines, could play a role in the regulation of GnRH (Popa et al., 2005). Also, the Oct-2 POU domain gene is reported to be associated with the onset of puberty (Terasawa & Fernandez, 2006). Furthermore, nutrition, social economic status (Grumbach & Styne, 2003), stress (Rieder & Coupey, 1999; Wierson, Long & Forehand, 1993) and exercise (Rogol, Clark & Roemmich, 2000) have shown to affect pubertal timing.

The enzymatic steps controlling testosterone biosynthesis and metabolism have been well characterized (see e.g. Griffin & Wilson, 2003), but the magnitude to which genetic or environmental effects contribute to the variation in circulating testosterone levels is not well established. Studies of heritability of testosterone have focused on adolescent and adult males (Meikle et al., 1986; 1988; Harris et al., 1998; Sluyter et al., 2000; Ring et al., 2005). Heritability estimates ranged from 26% (Meikle et al., 1986) to 66% (Harris et al., 1998) in

males, with the ages of the studied population varying from 14-21 year olds (Harris et al., 1998) to elderly men (Ring et al., 2005). The only study examining heritability in females, found that 41% of the variance in testosterone levels in 14-21 year old women and their mothers was explained by genetic factors (Harris et al., 1998). No studies into the heritability of testosterone levels in children have been reported.

The aim of our current study is to examine heritability of testosterone in 12-year-old children. At this age, variation in testosterone levels will be closely related to the variation in pubertal development. We examined the heritability of both circulating testosterone levels and the pubertal development that is thought to be under androgenic control. Furthermore, the overlap in genetic and environmental influences on testosterone levels and pubertal development is explored.

## **Methods**

### *Subjects*

This project is part of an ongoing longitudinal study on the development of cognition and behavioral problems in children. Subjects were recruited via the Netherlands Twin Register (NTR), kept by the Department of Biological Psychology at the Vrije Universiteit in Amsterdam (Boomsma et al., 1992; Boomsma et al., 2002). Out of the 209 twin pairs initially recruited in 1992, 183 twin pairs participated at age 12 (mean age = 12.02; SD = .08; for details see Bartels et al., 2002). Zygosity of the same sex twins was established by blood group (126 pairs) or DNA polymorphisms (19 pairs) and in four pairs by physical resemblance. The sample consisted of 34 monozygotic male (MZM), 39 dizygotic male (DZM), 40 monozygotic female (MZF), 36 dizygotic female (DZF), and 34 dizygotic pairs of opposite sex (DZMF). Because of difficulties during saliva collection (saliva volume too

limited for analysis, or sample contaminated with blood), or laboratory analyses, testosterone samples of 53 children could not be used, resulting in a total sample size of 313 children. Data on androgen dependent pubertal development were available for 172 girls and 165 boys. Since data analyses were performed on the raw data, all available data points were used.

#### *Salivary testosterone collection*

Saliva collection devices were sent to the twins by mail and samples were collected at home. Subjects were asked to collect their saliva by passive drool just before lunch, on two consecutive days, since reliability of salivary testosterone measurement is increased by using multiple samples (Dabbs, 1990). Twins from each pair collected their samples on the same days, and were instructed to do so on two school days to restrict their awakening time and time of sampling. Each participant was asked to write down the exact sampling time in a “saliva diary”, and to note exceptional events interfering with the daily routine. Subjects were instructed not to brush their teeth and not to eat or drink in the 30 minutes preceding the saliva collection and to thoroughly rinse their mouth with tap water before sampling. Saliva samples were stored in a refrigerator at the subject’s home until completion of the experimental protocol and were collected and brought to the laboratory by the research assistant. Samples were stored at -20°C until radioimmunoassay.

#### *Laboratory analysis*

The saliva samples of twins from the same pair were randomly distributed over different batches, the samples of the same subject were analyzed together in one batch. By doing so, any laboratory error would be reflected in variance unique to an individual, and not in variance common to a twin pair. All analyses were performed without knowledge of the zygosity of the twins. Testosterone level in saliva samples was measured after diethylether

extraction using an in-house competitive radioimmunoassay employing a polyclonal antitestosterone-antibody (Dr.Pratt AZG 3290). [1,2,6,7-3H]-Testosterone (TRK402, Amersham Nederland B.V.) was used as a tracer following chromatographic verification of its purity. The lower limit of detection was 10 pmol/L and inter-assay variation was 12.9, 8.3, and 10.6% at 60, 185 and 490 pmol/L respectively (n = 24).

### *Pubertal development*

Pubertal status was determined using the self report version of the Tanner scales (Marshall and Tanner 1969; 1970). The scales consist of schematic drawings of different pubertal stages of breast (girls), genital (boys) and pubic hair (both boys and girls) development. The scales range from 1 (pre-pubertal) to 5 (post-pubertal) for the assessment of breast and pubic hair development in girls. The pubic hair stages in boys ranged from 1 to 4, the genital development stages in boys ranged from 1 to 5. Additionally, girls were asked whether they had had their menarche yet. In this study, we focused on pubertal development that is known to be influenced by androgen levels. For girls, the pubic hair development score was used. For boys, the mean of the two developmental stages (genital development and pubic hair development) was taken and rounded off to the closest integer. If only one of the male Tanner stages was filled out (n = 5), the Tanner stage of this single measure was used as indicator of pubertal development. Studies into the reliability of self report measures of pubertal development have shown that the agreement between ratings of health professionals and self ratings decreases in later stages of puberty. Collapsing the Tanner stages of late puberty into one stage would increase the agreement (Coleman & Coleman, 2002). This finding, together with the fact that very few children reported late stages of puberty in our sample (see results section), made us decide to collapse Tanner stages 3, 4, and 5 into one stage.

### *Data analysis*

Descriptive statistics were calculated using SPSS/Windows 11.5. Testosterone measures were treated as continuous variables, pubertal development was treated as an ordinal variable with three categories. Because the testosterone variables were non-normally distributed, these variables were log transformed before statistical analysis. Analyses of variance were performed to examine testosterone level differences between girls who had had their first menstruation, and girls who had not had their menarche yet. Effects of day of sampling (day 1 vs. day 2), birth order (first born vs. second born), zygosity (MZ vs. DZ) and sex (males vs. females) on testosterone level means were examined using a bivariate saturated model in the computer program Mx (Neale et al., 2003). Phenotypic correlations between the two testosterone measures were estimated in Mx, polyserial correlations between testosterone and pubertal development were estimated in Mplus (Muthén & Muthén, 2006).

### *Genetic modeling*

In order to estimate the proportion of the variance arising from additive (A) genetic effects, dominance effects (D), shared (C) and non-shared (E) environmental influences, genetic modeling was performed in two steps, using both Mx (Neale et al., 2003) and Mplus (Muthén & Muthén, 2006). To get an impression of which variance components are of importance, first a saturated bivariate model for testosterone levels on the two measurement occasions was fitted using Mx. Based on the twin correlations, an ADE model was applied and evaluated using the -2 log likelihood (-2LL). The significance of sex differences on the magnitude of A, D, and E and the significance of variance components A and D was assessed by testing the deterioration in model fit after each component was dropped from the full model. The deterioration of the model was evaluated using the likelihood ratio test, the

difference between the -2LL under the two nested models, which is asymptotically distributed as a  $\chi^2$ . The degrees of freedom are given by the difference in the number of parameters estimated in the two models. A high increase in  $\chi^2$  against a low gain of degrees of freedom denotes a worse fit of the sub model compared to the full model. The most parsimonious model, with still a limited  $\chi^2$  was chosen as the best model.

Secondly, a trivariate model consisting of two continuous variables (testosterone levels on two measurement occasions) and one ordinal variable (pubertal development) was fitted in Mplus (Muthén & Muthén, 2006; see Prescott (2004) for specifying twin models in Mplus). A liability model with thresholds was fitted for pubertal development, assuming that pubertal development is a gradual process and subjects who exceed an underlying threshold are in a more advanced stage of puberty than subjects who score below this threshold. All models were fitted allowing parameter estimates for means and thresholds to differ across males and females. Initially, sex differences in magnitude of the variance components were also allowed. A series of models were fitted to the data, using weighted least squares with mean and variance adjusted chi-squares (WLSMV). When comparing nested models using WLSMV in Mplus, the  $\chi^2$  and degrees of freedom are adjusted (see the technical appendix of Mplus or the Mplus discussion board, both on [www.statmodel.com](http://www.statmodel.com)). The only value that is interpretable is the p-value. Therefore, only the p-values are mentioned in the results section. To test whether different genes affect pubertal development in males and females, we freely estimated the correlation between the genetic effects for opposite sex DZ twins. An estimated correlation significantly lower than 0.5 would indicate that opposite sex DZ twins share less genetic variance than same sex DZ twins. The estimate of the correlation was constrained to fall within the range of biologically plausible values (0 - 0.5) first, and consequently constrained to the value of 0.5. Adjusted chi-squared difference testing using WLSMV is not available in Mplus 4.0 in combination with boundary constraints. Therefore weighted least

squares with mean adjusted chi-squares (WLSM) was used in this case, for which a simplified version of the adjusted chi-squared difference test was presented in Satorra & Bentler (1999; see also the technical appendix of Mplus, page 22, on [www.statmodel.com](http://www.statmodel.com)). We estimated the simplified chi-squared difference to test if the estimated opposite sex DZ correlation differed significantly from 0.5. Similarly, we tested whether the correlation of shared environmental effects on pubertal development was lower in opposite sex twins than in same sex twins. An estimated correlation lower than 1.0 would indicate that opposite sex twins have less shared environmental influences in common than same sex twins. The estimate of the correlation was bound to fall between 0 and 1. Subsequently it was tested whether the shared environmental correlation in opposite sex twins was significantly different from zero.

## **Results**

Table 1 shows the scores on the Tanner scales in boys and girls. The majority of the children reported to be in the early to mid stages of puberty: 81.9% (genital development) and 98.2% (pubic hair growth) of the boys reported to be in early to mid-pubertal stages. Of the girls, respectively 88.5% and 78.5 % reported to be in early to mid stages of breast development and pubic hair growth. Twenty girls (both from the MZ and DZ twin groups) reported that they had experienced their first menstruation. Testosterone levels in these girls were significantly increased compared to those who had not yet had their first menstruation,  $F(1, 149) = 6.144, p = .014$ . No outliers were detected; all testosterone levels fell within a range of 2 SDs from the mean. Therefore it was decided to retain all data in the analysis. In subsequent analyses using the androgen dependent Tanner scales, stage 3 and above were collapsed into one stage.

Table 1 and 2 about here

Descriptive statistics on the testosterone measures for the two samples are presented in Table 2. Sample size, mean testosterone levels and standard deviations are shown for both boys and girls. Moreover, testosterone levels are shown per stage of androgen dependent pubertal development. Testosterone levels were significantly correlated with pubertal development:  $r = .31$  (95% CI = .20 - .43), this correlation was equal for both testosterone measurements, and equal in boys and girls. Within person correlation for the testosterone samples taken on the two consecutive days was  $.55$  (95% CI = .46 - .64) There were no birth order effects (differences in testosterone level between first- and second borns,  $\chi^2_4 = 3.028$ ,  $p = .553$ ) or zygosity effects (mean testosterone level DZ vs. DZMF twins,  $\chi^2_2 = 1.465$ ,  $p = .481$ ; MZ vs. DZ twins,  $\chi^2_2 = 0.545$ ,  $p = .761$ ). Testosterone levels were higher in girls than in boys ( $\chi^2_1 = 20.773$ ,  $p < .001$ ).

Table 3 about here

Twin correlations for testosterone levels (see Table 3) showed that MZ correlations were higher than DZ correlations in both sexes, indicating additive genetic effects (A). In boys, the DZ correlation for testosterone levels was less than half the MZ correlation, suggesting dominance effects. The relatively high correlation in DZ opposite sex twins ( $r = .23$  for sample 1 and  $r = .53$  for sample 2) suggests overlap in genetic expression in boys and girls. A bivariate ADE model of testosterone levels on two measurement occasions was fitted. Constraining the parameters that represent the influence of A, D and E to be the same in boys and girls did not significantly worsen the fit ( $\chi^2_9 = 8.373$ ,  $p = .497$ ), suggesting that the relative effects of these components were equal in both sexes. Dropping component D

from the model did not result in a worse model fit ( $\chi_3^2 = 1.661$ ,  $p = .646$ ). The best fitting parsimonious model was a model in which variation of testosterone levels was accounted for by additive genetic influences and non-shared environmental effects.

Twin correlations for androgen dependent pubertal development (see Table 3) showed very high MZ correlations and rather high DZ correlations, suggesting both genetic and shared environmental effects. The point estimate for the correlation of opposite sex DZ twins was low, but the confidence interval for this correlation was rather wide. A trivariate model incorporating both testosterone measures and pubertal development was fitted in Mplus. Based on the outcome of the analyses in Mx, an AE model was tested for the testosterone variables. Based on the twin correlations, the variance unique to pubertal development (i.e. the variance that could not be explained by additive genetic or non-shared environmental variance in testosterone) was partitioned into components A, C and E. Genetic model fitting showed that the genetic correlation of pubertal development in opposite sex DZ twins could be equaled to 0.5, similar to same sex DZ twins ( $\chi_1^2 = .069$ ,  $p = .792$ ). Constraining the parameters that represent the influence of A and E on testosterone levels to be the same in boys and girls did not significantly worsen the fit ( $p = .823$ ), indicating that the relative effects of these components are equal in both sexes. Similarly, constraining the variance components A, C and E on the residual variance of pubertal development to be equal in both sexes, did not significantly worsen the fit ( $p = .221$ ). The shared environmental influences on the variance unique to pubertal development were of significant importance, dropping the C component from the model lead to a significant deterioration of the model ( $p = .024$ ). Constraining the influence of E on the covariance between the variables to be zero did not significantly affect the fit of the model ( $p = .304$ ), indicating that the non-shared environmental influences were unique for all measurements. The genetic influences on both measurement occasions of testosterone could be constrained to be from the same source ( $p =$

.265), indicating that all genetic influences on the testosterone measurements were shared between the two samples. The genetic effects on the covariance between testosterone and pubertal development were of significant importance, dropping this component from the model resulted in a significant deterioration of the fit ( $p < .001$ ). Also, the genetic influences on the variance unique to pubertal development were significant ( $p = .016$ ). The best fitting model incorporating both testosterone measures and androgen dependent pubertal development is given in Figure 1.

Insert Figure 1 about here

Heritability estimates were calculated by squaring the standardized estimates of the path loadings. In the best fitting model, additive genetic effects explained  $(.720)^2 = 52\%$  of the variance in testosterone levels on both measurement days. The non-shared environmental influences on T-levels were unique to both measurement occasions, and could explain  $(.694)^2 = 48\%$  of the variance. Fifteen percent of the variance  $(.391)^2$  in pubertal development was shared with variance in testosterone levels, and this covariance was entirely accounted for by genetic effects. The variance unique to pubertal development was accounted for by genetic effects (45%); shared environmental effects (37%) and a small non-shared environmental effect (3%). Lastly, it was tested whether the correlation of shared environmental influences might be lower in opposite sex twins. The shared environmental correlation in opposite sex twins was not significantly different from zero ( $\Delta\chi_1 = 0.037$ ;  $p = .847$ ), suggesting that opposite sex twins may have less environmental variance in common than same sex twins.

## Discussion

This is the first study to report heritability estimates of testosterone levels in early adolescence. A significant contribution of additive genetic effects to midday salivary testosterone levels was found. The relative contribution of genetic and environmental influences on testosterone level variation was the same in both sexes. Genetic influences explained 52% of the variation. The remaining proportion of the variance was accounted for by non-shared environmental influences. Moreover, a significant overlap in genetic influences was found between testosterone levels and androgen dependent pubertal development. Fifteen percent of the variance in pubertal development was shared with the variance in testosterone levels. The overlap was entirely explained by genetic effects.

Previous studies have shown that environmental influences, such as participating in sports competitions (and especially winning a competition), could lead to changes in testosterone levels (Archer, 2006). Apart from such 'real' environmental influences, non-shared environmental influences also include measurement error. In this respect it is important to note that both the MZ correlation for testosterone levels ( $r$  ranging from .51 to .76) and the MZ cross twin cross sample correlation (cross  $r = .35 - .49$ ) were as high as the within person correlation for both testosterone samples ( $r = .55$ ). This finding suggests that, if measurements were corrected for daily fluctuations and measurement error, variance in testosterone levels would practically be entirely explained by genetic effects. This is also reflected in the observation that all covariance between the two testosterone samples is of genetic origin, and all unique variance is non-shared environmental.

Harris et al. (1998) found a heritability estimate of 66% in 14- to 21-year-old males, and an estimate of 41% in females (14-21 year old twins and their mothers). Furthermore, they found zero correlation between testosterone levels of opposite sex DZ twins. In our study, neither sex specific genes nor sex differences in heritability estimates were detected.

The opposite sex DZ twin correlation for testosterone level was .23 and .53 in our sample, suggesting an overlap in genetic expression in boys and girls. Our subjects were younger than the participants in Harris' study, and were still in the early to mid-stages of puberty. Concentrations of circulating testosterone levels are subject to large changes over the lifetime, particularly during puberty. After a surge of testosterone levels in male fetuses, levels are equally low in boys and girls after birth, and remain suppressed until the reactivation of the GnRH neurons, marking the onset of puberty (Grumbach & Styne, 2003). During pubertal development, sex differences in testosterone levels arise, from then onwards testosterone levels are higher in males than in females (Grumbach & Styne, 2003). In our sample, testosterone levels were found to be significantly higher in girls than in boys, indicating that the majority of the boys was still prepubescent. A study in adolescent twins and their parents found no correlations in plasma testosterone levels between fathers and sons, suggesting that different genetic mechanisms may influence testosterone concentrations during the life span (Harris et al., 1998). The results of our study suggest that in pre- and early puberty, there are no sex differences in genes influencing variation in testosterone levels. Neither differences in the magnitude of the variance components, nor significant influences of dominance effects or shared environmental factors were found on the variation in testosterone levels. However, our sample size was relatively small and a contribution of shared environmental or genetic dominance effects, or subtle differences in the magnitude of A and E on the variance in testosterone levels cannot completely be excluded, due to statistical power considerations.

Variance in androgen dependent pubertal development was largely (60%) explained by genetic effects. Shared environmental effects could explain 37% of the variation in pubertal development. Nutrition and other factors related to socio-economic status have been shown to influence pubertal development (Grumbach & Styne, 2003), these factors could

explain the non-genetic familial clustering. The twins completed the puberty questionnaire at home. Therefore, it cannot be ruled out that the twin resemblance was inflated because twins discussed their ratings. This explanation is further supported by the finding that the shared environmental correlation in opposite sex twins (who could not discuss their ratings because they filled out different questionnaires) did not significantly differ from zero, suggesting less shared environmental variance in pubertal development in opposite sex twins than in same sex twins.

In summary, this study provides evidence that the variation in salivary testosterone levels in 12-year-old children is explained by genetic and non-shared environmental effects in both boys and girls. The relatively high correlation between testosterone levels in opposite sex DZ twins suggests that sex differences in gene expression arise in more advanced stages of puberty. At this age, fifteen percent of the variance in androgen induced pubertal development is shared with the variance in testosterone levels. This overlap is entirely accounted for by genetic influences.

## **Acknowledgments**

Financial support was given by The Netherlands Organization for Scientific Research (NWO, grant numbers 575-25-012 and NWO/SPI 56-464-14192). Dr Bartels is financially supported by NWO (VENI: 451-04-034). The authors would like to thank Marleen de Moor and Irene Rebollo for their assistance with the Mplus analyses.

## References

- Archer, J. (1991). The influence of testosterone on human aggression. *British Journal of Psychology*, 82, 1-28.
- Archer, J. (2006). Testosterone and human aggression: an evaluation of the challenge hypothesis. *Neuroscience and Biobehavioral Reviews*, 30, 319-345.
- Auchus, R. J. & Rainey, W. E. (2004). Adrenarche - physiology, biochemistry and human disease. *Clinical Endocrinology*, 60, 288-296.
- Bartels, M., Rietveld, M. J. H., Van Baal, G. C. M. & Boomsma, D. I. (2002). Genetic and environmental influences on the development of intelligence. *Behavior Genetics*, 32, 237-249.
- Book, A. S., Starzyk, K. B. & Quinsey, V. L. (2001). The relationship between testosterone and aggression: a meta-analysis. *Aggression and violent behavior*, 6, 579-599.
- Boomsma, D. I., Orlebeke, J. F. & Van Baal, G. C. M. (1992). The Dutch Twin Register: growth data on weight and height. *Behavior Genetics*, 22, 247-251.
- Boomsma, D. I., Vink, J. M., Van Beijsterveldt, C. E. M., de Geus, J. C. N., Beem, A. L., Mulder, E. J., Derks, E. M., Riese, H., Willemsen, A. H. M., Bartels, M., Van den Berg, M., Kupper, N. H., Polderman, T. J. C., Posthuma, D., Rietveld, M. J. H., Stubbe, J. H., Knol, L. I., Stroet, T. & Van Baal, G. C. M. (2002). Netherlands Twin Register: a focus on longitudinal research. *Twin Research*, 5, 401-406.

- Coleman, L. & Coleman, J. (2002). The measurement of puberty: a review. *Journal of Adolescence*, 25, 535-550.
- Dabbs, J. M. Jr. (1990). Salivary testosterone measurements: reliability across hours, days, and weeks. *Physiology and Behavior*, 48, 83-86.
- Eaves, L., Silberg, J., Foley, D., Bulik, C., Maes, H., Erkanli, A., Angold, A., Costello, E. J. & Worthman, C. (2004). Genetic and environmental influences on the relative timing of pubertal change. *Twin Research*, 7, 471-481.
- Graber, J. A., Seeley, J. R., Brooks-Gunn, J. & Lewinsohn, P. M. (2004). Is pubertal timing associated with psychopathology in young adulthood? *Journal of the American Academy of Child and Adolescent Psychiatry*, 43, 718-726.
- Griffin, J. E. & Wilson, J. D. (2003). Disorders of the testes and the male reproductive tract. In P. R. Larsen, H. M. Kronenberg, S. Melmed, & K. S. Polonski (Eds.), *Williams Textbook of Endocrinology*. (pp. 709-769). W.B. Saunders, Philadelphia.
- Grumbach, M. & Styne, D. (2003). Puberty: ontogeny, neuroendocrinology, physiology, and disorders. In P. R. Larsen, H. M. Kronenberg, S. Melmed, & K. S. Polonski (Eds.), *Williams Textbook of Endocrinology*. (pp. 1115-1286). W.B. Saunders., Philadelphia.
- 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.

Hayward, C. & Sanborn, K. (2002). Puberty and the emergence of gender differences in psychopathology. *Journal of Adolescent Health, 30*, 49-58.

Marshall, W. A. & Tanner, J. M. (1969). Variations in pattern of pubertal changes in girls. *Archives of Disease in Childhood, 44*, 291-303.

Marshall, W. A. & Tanner, J. M. (1970). Variations in the pattern of pubertal changes in boys. *Archives of Disease in Childhood, 45*, 13-23.

Mazur, A. & Booth, A. (1998). Testosterone and dominance in men. *Behavioral and Brain Sciences, 21*, 353-363.

Meikle, A. W., Bishop, D. T., Stringham, J. D. & West, D. W. (1986). Quantitating genetic and nongenetic factors that determine plasma sex steroid variation in normal male twins. *Metabolism, 35*, 1090-1095.

Meikle, A. W., Stringham, J. D., Bishop, D. T. & West, D. W. (1988). Quantitating genetic and nongenetic factors influencing androgen production and clearance rates in men. *Journal of Clinical Endocrinology and Metabolism, 67*, 104-109.

Mustanski, B. S., Viken, R. J., Kaprio, J., Pulkkinen, L. & Rose, R. J. (2004). Genetic and environmental influences on pubertal development: longitudinal data from Finnish twins at ages 11 and 14. *Developmental Psychology, 40*, 1188-1198.

Muthén, L. K. & Muthén, B. O. (1998-2006). *Mplus User's guide* (4th ed.). Los Angeles, CA: Muthén & Muthén.

Neale, M. C., Boker, S. M., Xie, G. & Maes, H. H. (2003). *Mx: Statistical modeling* (6<sup>th</sup> ed.). Richmond, VA: Department of Psychiatry, Virginia Commonwealth University.

Palmert, M. R. & Boepple, P. A. (2001). Variation in the timing of puberty: clinical spectrum and genetic investigation. *Journal of Clinical Endocrinology and Metabolism*, 86, 2364-2368.

Popa, S. M., Clifton, D. K. & Steiner, R. A. (2005). A KiSS to remember. *Trends in Endocrinology and Metabolism*, 16, 249-250.

Prescott, C. A. (2004). Using the Mplus computer program to estimate models for continuous and categorical data from twins. *Behavior Genetics*, 34, 17-40.

Raine, A. (2002). Biosocial studies of antisocial and violent behavior in children and adults: a review. *Journal of Abnormal Child Psychology*, 30, 311-326.

Rieder, J. & Coupey, S. M. (1999). Update on pubertal development. *Current Opinions in Obstetrics and Gynecology*, 11, 457-462.

Ring, H. Z., Lessov, C. N., Reed, T., Marcus, R., Holloway, L., Swan, G. E. & Carmelli, D. (2005). Heritability of plasma sex hormones and hormone binding globulin in adult male twins. *Journal of Clinical Endocrinology and Metabolism*, 90, 3653-3658.

Rogol, A. D., Clark, P. A. & Roemmich, J. N. (2000). Growth and pubertal development in children and adolescents: effects of diet and physical activity. *American Journal of Clinical Nutrition*, 72, 521S-528S.

Satorra, A. & Bentler, P. M. (1999). *A scaled difference chi-square test statistic for moment structure analysis*. Technical report, University of California, Los Angeles.

Sisk, C. L. & Foster, D. L. (2004). The neural basis of puberty and adolescence. *Nature Neuroscience*, 7, 1040-1047.

Sluyter, F., Keijser, J. N., Boomsma, D. I., van Doornen, L. J., Van den Oord, E. J. C. G. & Snieder, H. (2000). Genetics of testosterone and the aggression-hostility-anger (AHA) syndrome: a study of middle-aged male twins. *Twin Research*, 3, 266-276.

Terasawa, E. & Fernandez, D. L. (2001). Neurobiological mechanisms of the onset of puberty in primates. *Endocrine Reviews*, 22, 111-151.

Van den Berg, S. M., Setiawan, A., Bartels, M., Polderman, T. J. C., van der Vaart, A. W. & Boomsma, D. I. (2006). Individual differences in puberty onset in girls: Bayesian estimation of heritabilities and genetic correlations. *Behavior Genetics*, 36, 261-270.

Wierson, M., Long, P. J. & Forehand, R. L. (1993). Toward a new understanding of early menarche: The role of environmental stress in pubertal timing. *Adolescence*, 28, 913-924.

TABLE 1. Frequency distributions of pubertal Tanner stages for breast development and pubic hair growth in girls and genital development and pubic hair growth in boys.

| <b>Tanner stage</b> | <b>♀ (n = 174)<br/>Breast development</b> | <b>♀ (n = 172)<br/>pubic hair growth</b> | <b>♂ (n = 160)*<br/>genital development</b> | <b>♂ (n = 163)*<br/>pubic hair growth</b> |
|---------------------|---|--|---|---|
| 1                   | 38 (21.8%)                                | 47 (27.3%)                               | 24 (15%)                                    | 85 (52.1%)                                |
| 2                   | 66 (37.9%)                                | 41 (23.9%)                               | 27 (16.9%)                                  | 71 (43.6%)                                |
| 3                   | 50 (28.8%)                                | 47 (27.3%)                               | 80 (50%)                                    | 4 (2.5%)                                  |
| 4                   | 20 (11.5%)                                | 33 (19.2%)                               | 27 (16.9%)                                  | 3 (1.8%)                                  |
| 5                   | 0   | 4 (2.3%)                                 | 2 (1.2%)                                    | N/A                                       |

\* Number of boys for which information on either one of the pubertal stages is available is

165.

TABLE 2. Descriptive statistics for testosterone levels (pmol/L) in sample 1 (s1) and sample 2 (s2) in all boys (♂) and girls (♀), and separate for each stage of pubertal development.

|  | <b>N</b> | <b>Mean</b> |               | <b>N</b> | <b>Mean</b> |
|--|----------|-------------|---------------|----------|-------------|
| <b>Total</b>                             |          |             |               |          |             |
| ♂s1                                      | 140*     | 71.6 (23.2) | ♀ s1          | 145*     | 85.4 (29.9) |
| ♂s2                                      | 134*     | 70.7 (24.2) | ♀ s2          | 145*     | 81.3 (28.1) |
| <b>Per stage of pubertal development</b> |          |             |               |          |             |
| ♂s1 Stage 1                              | 14       | 69.9 (21.7) | ♀s1 Stage 1   | 32       | 71.4 (24.6) |
| ♂s2 Stage 1                              | 17       | 68.2 (23.4) | ♀s2 Stage 1   | 35       | 70.6 (21.8) |
| ♂s1 Stage 2                              | 57       | 63.9 (18.9) | ♀s1 Stage 2   | 35       | 77.3 (21.7) |
| ♂s2 Stage 2                              | 54       | 67.2 (16.1) | ♀s2 Stage 2   | 38       | 81.5 (32.4) |
| ♂s1 Stage ≥ 3                            | 58       | 78.7 (25.5) | ♀s1 Stage ≥ 3 | 71       | 94.8 (32.5) |
| ♂s2 Stage ≥ 3                            | 52       | 73.9 (27.3) | ♀s2 Stage ≥ 3 | 63       | 86.9 (26.9) |

\* Tanner stage information missing for 11 boys in both sample 1 and 2. Tanner stage

information missing for 7 girls in sample 1 and for 9 girls in sample 2.

TABLE 3. Twin correlations and cross-correlations per zygosity for testosterone sample 1 and 2 (estimated in Mx using maximum likelihood) and for pubertal development (polychoric correlations in Mplus). 95% confidence interval between parentheses.

|             | <b>Testosterone sample 1</b> | <b>Testosterone sample 2</b> | <b>Twin cross <math>r_{s1-s2}</math></b> | <b>Pubertal development</b> | <b>Twin cross <math>r_{s1s2-pd}</math></b> |
|-------------|------------------------------|------------------------------|--|-----------------------------|--|
| <b>MZM</b>  | .76 (.55-.88)                | .67 (.41-.83)                | .49 (.22-.69)                            | .97 (.93-1.0)               | .03 (-.34-.39)                             |
| <b>DZM</b>  | .12 (.00-.46)                | .00 (.00-.21)                | .02 (-.12-.24)                           | .75 (.52-.99)               | .25 (.02-.48)                              |
| <b>MZF</b>  | .51 (.17-.73)                | .58 (.30-.77)                | .35 (.07-.58)                            | .96 (.89-1.0)               | .21 (-.04-.45)                             |
| <b>DZF</b>  | .27 (.00-.57)                | .23 (.00-.53)                | .25 (-.03-.52)                           | .73 (.48-.97)               | .28 (.05-.51)                              |
| <b>DZMF</b> | .23 (.01-.55)                | .53 (.16-.77)                | .35 (.06-.60)                            | .15 (-.31-.60)              | .15 (-.11-.41)                             |

MZM = monozygotic male twin pairs; DZM = dizygotie male twin pairs; MZF = monozygotic female twin pairs; DZF = dizygotie female twin pairs; DZMF = opposite sex twin pairs. s1 = testosterone sample 1; s2 = testosterone sample 2; pd = pubertal development.

FIGURE 1. Path diagram of the best fitting model with standardized estimates for the different variance components. Observed variables are shown in boxes; latent variables are depicted in circles and ovals; the model for the means is represented by the triangle. Ts1 = testosterone level in sample 1; Ts2 = testosterone level sample 2; PD = Underlying pubertal development, as measured with the ordinal variable pubertal developmental stage. Means and thresholds are estimated for boys (first value) and girls (latter value) separately. At1 = genetic influence on individual differences in testosterone levels; Et1/2 = contribution of non-shared environmental influences on variance in testosterone levels in sample 1 (Et1) and sample 2 (Et2). Apd/Cpd/Epd = Influence of additive genetic (Apd), shared environmental (Cpd) and non-shared environmental (Epd) factors on individual differences unique to pubertal development.

