Arch Sex Behav (2006) 35:647–658
DOI 10.1007/s10508-006-9072-0

Genetic and Environmental Influences on Cross-Gender Behavior and Relation to Behavior Problems: A Study of Dutch Twins at Ages 7 and 10 Years

C. E. M. van Beijsterveldt · James J. Hudziak · Dorret I. Boomsma

Received: 4 May 2005 / Revised: 29 December 2005 / Accepted: 5 April 2006 / Published online: 16 November 2006

Abstract The aim of this study was to investigate the prevalence of cross-gender behavior during childhood, to estimate the influence of genotype and environment on variation in cross-gender behavior, and to explore the association of cross-gender behavior with maternal ratings of behavior problems as indexed by the Internalizing and Externalizing scales of the Child Behavior Checklist (CBCL). Cross-gender behavior was assessed by two items from the CBCL: "behaves like opposite sex" and "wishes to be of opposite sex." As part of an ongoing longitudinal study of the Netherlands Twin Registry, mothers were asked to complete the CBCL for their twins when they were 7 ($n \sim 14{,}000$ twins) and 10 years old ($n \sim 8{,}500$ twins). The prevalence of cross-gender behavior (as measured by maternal report of behaving like or wishing to be the opposite sex) was 3.2% and 5.2% for 7-year-old boys and girls, respectively, and decreased to 2.4% and 3.3% for 10-year-old boys and girls. Surprisingly, the prevalence rate of cross-gender behavior of girls with a male co-twin was lower than of girls with a female co-twin. At both ages, the similarity for cross-gender behavior was greater in monozygotic than in dizygotic twins pairs. Genetic structural equation modeling showed that 70% of the variance in the liability of cross-gender behavior could be explained by genetic factors, at both ages and for both sexes. Cross-gender behavior was associated with higher scores on Internalizing and Externalizing problems, both in boys and in girls.

Keywords Cross-gender Behavior · Heritability · Twins · Child Behavior Checklist · Gender identity disorder

C. E. M. van Beijsterveldt (✉) · D. I. Boomsma
Department of Biological Psychology, Vrije Universiteit,
Van der Boechorststraat 1,
1081 BT Amsterdam, The Netherlands
e-mail: toos@psy.vu.nl

J. J. Hudziak
Department of Psychiatry and Medicine (Division of Human Genetics) and Centre for Children, Youth, and Families,
University of Vermont College of Medicine,
Burlington, Vermont, USA

Introduction

Gender identity refers to the basic feelings of belonging to one sex or the other. Gender role reflects the amount of typical feminine and masculine behaviors that a person expresses. The expression of typical gender behaviors depends on social norms, and varies with the knowledge of gender roles in a society, given a certain culture and historical period (Zucker, 2005; Zucker & Bradley, 1995). The development of gender identity starts in the second year of life and is manifested in activities such as toy interests and activity preferences. By the age of 6 or 7 years, gender-typical behaviors are well established (Bradley & Zucker, 1997; Carver, Yunger, & Perry, 2003). When a child has strong feelings that his or her biological sex does not agree with his/her gender identity and these feelings lead to persistent problems, a child may be diagnosed as having a Gender Identity Disorder (GID) (American Psychiatric Association, 2000). Estimates of the prevalence of GID in children are scarce (Zucker & Bradley, 1995; Zucker, Bradley, & Sanikhani, 1997) and are mostly derived from the number of adults with GID who have been referred to specialty clinics.

The development of gender identity is a complex process, for which the etiology is largely unknown. Twin studies may contribute to an increased understanding of the etiology of variation in gender identity (Boomsma, Busjahn, & Peltonen, 2002). With twin studies, it is possible to estimate the genetic

and environmental contributions to variation in gender behaviors. Thus far, only a few twin studies have investigated genetic influences on GID (Zucker & Bradley, 1995). Four studies examined the heritability of atypical gender behaviors, but with contrasting results. In one study by Coolidge, Thede, and Young (2002), parents reported on six DSM-IV GID-related behaviors in sample of 157 non-referred twin pairs, aged 4–17 years. Coolidge et al. found that familial resemblance accounted for about 60% of the variance in GID, but the statistical power was too low to determine whether familial resemblance was due to genetic or shared environmental factors.

Three other twin studies reported sex differences in heritability of atypical gender behaviors (Bailey, Dunne, & Martin, 2000; Iervolino, Hines, Golombok, Rust, & Plomin, 2005; Knafo, Iervolino, & Plomin, 2005). Using a retrospective design, Bailey et al. examined the heritability of childhood gender nonconformity in 1891 adult Australian twins. A composite measure of gender nonconformity was based on items from various assessment instruments, including items assessing childhood sex-typed behavior (such as playing stereotypic games and activities) and gender identity (e.g., internal feelings of maleness or femaleness). Childhood gender nonconformity was heritable for both men and women, with larger genetic effects in men ($h^2 = 0.50$) than in women ($h^2 = 0.37$).

Knafo et al. (2005) measured feminine and masculine gender role behavior using parents' report on the PreSchool Activities Inventory (PSAI) (Golombok & Rust, 1993) in a large sample of 5733 same-sex 3–4 year-old twin pairs. From this sample, they selected twin pairs in which at least one member of a pair had an extreme score for gender atypical behavior and estimated the heritability for this group. In addition, Knafo et al. examined whether the etiology of gender atypical behaviors differed as a function of severity and estimated the heritability for fully gender atypical boys/girls and partially atypical boys/girls.

Genetic and environmental influences on atypical-gender behaviors differed between boys and girls. For boys, shared environment explained the largest part of the variability in femininity (51–57%), while the role of genetic influences was modest (21–32%). The environmental and genetic influences on the variability of atypical-gender behaviors did not differ for partially and fully gender atypical boys. In girls, genetic influences on the variability of atypical-gender behaviors were more important than for boys. The heritability estimate ranged between 42–50% and shared environment explained 33–43% of the individual differences in girls' atypical behaviors. In addition, the heritabilities were dependent on the definition used for gender atypical behaviors. If the gender atypical behaviors of girls were defined as fully gender atypical, then the variance in gender atypical behavior was mainly explained by genetic factors (65%) with no effect of the shared environment. If the partially gender atypical definition was used, then the results were comparable with the pattern of the atypical gender behaviors in boys. The influence of shared environmental factors was large (67–72%) and heritability was modest (15–20%).

Using an overlapping sample, as in Knafo et al. (2005), Iervolino et al. (2005) examined the environmental and genetic influences on the full range of variation in gender role behaviors instead of on more extreme aspects of gender role behaviors. It is possible that the relative influence of genes and environment differs for extreme forms of atypical gender behaviors vs. normal ranges of variation. Again, sex differences in heritability were observed. For girls, heritability of gender role behaviors was 57% and shared environmental factors were not important. For boys, heritability was 34% but the role of shared environmental factors was also important (29%).

In sum, the results of the twin studies suggest that genetic factors contribute to variation in gender atypical behaviors, and that the genetic influence may be different for boys and girls. However, the estimates of genetic and environmental influences were inconsistent across studies. Possible explanations for these differences may be the use of different definitions of atypical gender behavior and differences in age ranges. It is known that heritability is not a fixed parameter, but that it can vary from age to age.

Biological explanations of gender-typed behaviors have often focused on the effects of prenatal sex hormones (Collaer & Hines, 1995). Evidence for the influence of prenatal hormones is derived mainly from two lines of research. One research line has studied girls with congenital adrenal hyperplasia. Due to an enzymatic defect, these girls have been exposed to high levels of testosterone during fetal development and show, in comparison to unaffected girls, more masculine behaviors (Berenbaum & Hines, 1992; for review, see Cohen-Bendahan, van de Beek, & Berenbaum, 2005). Other evidence stems from animal research. During fetal development, fetuses may be exposed to different levels of hormones depending on the sex of neighboring fetuses. In mice, it was found that females lying between two males in utero tended to develop more masculine morphological and behavioral traits than females lying between two females (Miller, 1994; for review, see Ryan & Vandenbergh, 2002).

In humans, the intrauterine environment effect on cross-gender behavior can be tested by comparing traits of same-sex DZ twins vs. opposite-sex twins. Analogue to animal research, it is proposed that during the gestation hormones transfer from male to female fetuses. Therefore, DZ girls with a male co-twin may be exposed to higher testosterone levels than DZ girls with a female co-twin, and will have more masculine traits. The finding of a more male-typical finger-length ratio in females with a male co-twin relative to females with a female co-twin, gives evidence

for the hormonal transfer theory during the gestation (Anders, Vernon, & Wilbur, 2006). Studies that have examined the gender-typed behaviors in opposite-sex female twins have revealed more mixed results (Cohen-Bendahan, Buitelaar, van Goozen, & Cohen-Kettenis, 2004; Cohen-Bendahan, Buitelaar, van Goozen, Orlebeke, & Cohen-Kettenis, 2005; Henderson & Berenbaum, 1997; Resnick, Gottesman, & McGue, 1993; Rodgers, Fagot, & Winebarger, 1998). Resnick et al. found that females with a male co-twin showed more sensation-seeking behaviors, which are more typical of males, than females with a female co-twin. In the study by Cohen-Bendahan, Buitelaar et al. (2005), opposite-sex girls showed a more masculine pattern of aggression. However, in a much larger sample, Koopmans, Boomsma, Heath, and van Doornen (1995) did not find any differences in sensation-seeking behaviors between adolescent and young adult girls from opposite-sex and same-sex twin pairs. Two other studies provided no evidence for more masculine behaviors in girls with a male co-twin (Henderson & Berenbaum, 1997; Rodgers et al., 1998). Both studies found that girls with a male co-twin did not spend more time playing with boys' toys than girls with a girl co-twin. In the study by Rodgers et al. the opposite was found: girls with a twin sister played more with boys' toys than girls with brothers. Thus, the direction of the results of some of these studies may support the theory of hormonal transfer during gestation, but other studies do not. Furthermore, postnatal influences on these traits can not be excluded. The present study included opposite-sex twin pairs and offers the opportunity to examine whether DZ girls with a male co-twin show more atypical-gender behaviors than DZ girls with a female co-twin.

A pervasive feature of many psychiatric disorders is the co-occurrence of two or more disorders in the same individual. Gender atypical behaviors are often found to be associated with increased levels of anxiety (Cohen-Kettenis, Owen, Kaijser, Bradley, & Zucker, 2003; Zucker & Bradley, 1995). In a clinical sample of boys with GID, the prevalence of separation anxiety disorder (SAD) was higher in boys with GID than in boys with a subthreshold GID diagnosis (Zucker, Bradley, & Lowry Sullivan, 1996). However, the results depended on the definition of SAD. A diagnosis of SAD was based on a structured interview with the mother and consisted of items related to SAD according to DSM-III criteria. If the diagnosis of SAD was judged to meet the DSM-III criteria, then the prevalence of SAD did not differ between boys with GID and boys with subthreshold GID diagnosis. If the criteria for SAD were less stringent, then the boys with GID showed SAD more often than the boys without the GID diagnosis. In the twin study of Coolidge et al. (2002), a significant correlation was found between GID and depression for non-referred children ($r = .20$), but the correlation between GID and separation anxiety was non-significant ($r = .11$). Yunger, Carver, and Perry (2004) also reported a negative association between gender typicality and Internalizing problems. These results suggest that gender atypical behavior is associated with increased levels of problem behaviors. In our study, we examined whether children with cross-gender behavior, as measured by two CBCL gender items in a non-clinical population, have an increased risk for emotional and behavioral problems.

The present study had three aims. First, we examined the prevalence of cross-gender behavior in a large sample of male and female twins at age 7 and 10 years. We assessed prevalence as a function of age and sex, and compared boys and girls from same- and opposite-sex twin pairs. The second aim was to determine the genetic and environmental contributions to variation in cross-gender behavior and to test for sex differences in the magnitude of genetic and environmental influences. Third, we examined if children with cross-gender behaviors had higher levels of Internalizing (INT) or Externalizing (EXT) problems than children without cross-gender behaviors. The cross-gender behaviors were obtained in longitudinal samples of twin pairs. Cross-gender behavior was derived from 2 items of the maternal reported Child Behavior Checklist (CBCL): "behaves like opposite sex" and "wishes to be of opposite sex."

Method

Participants

The data for this study came from a longitudinal study using the Netherlands Twin Register, which is maintained by the Department of Biological Psychology at the Free University in Amsterdam (see Boomsma et al., 2002; Boomsma, van Beijsterveldt, & Hudziak, 2005). In the present study, we included data of maternal CBCL reports for 7526 twin pairs obtained at age 7 (covering birth years 1987 to 1996) and for 4538 twin pairs at age 10 (covering birth years 1987 to 1993). The number of twin pairs with data at both ages was 3894. The lower sample size at age 10 reflects the fact that the study is an ongoing longitudinal study to which we add new cohorts annually and that questionnaires have not been sent to twins who do not have the appropriate age yet. Twin zygosity was based on blood group or DNA polymorphisms for 822 same-sex pairs (blood, $n = 424$; DNA, $n = 398$). For the remaining twins, the classification of zygosity was based on a discriminant analysis, relating questionnaire items to zygosity based on blood/DNA typing in a group of same-sex twin pairs (Goldsmith, 1991). The zygosity was correctly classified by questionnaire in nearly 95% of the cases (Rietveld et al., 2000).

A twin pair was excluded when one or both of the twins had a disease or handicap that interfered severely with normal

Table 1 Number (and percentage) of children with maternal ratings of 1 or 2 on the two Child Behavior Checklist items pertaining to cross-gender behavior

	Age 7				Age 10			
	N	Item 5	Item 110	Both items	N	Item 5	Item 110	Both items
Boys								
MZ	2430	78 (3.2)	17 (0.7)	82 (3.4)	1496	31 (2.0)	11 (0.7)	36 (2.4)
DZ	2477	93 (3.7)	33 (1.5)	102 (4.1)	1367	46 (3.4)	21 (1.5)	51 (3.7)
OS	2295	72 (3.1)	25 (1.0)	83 (3.5)	1403	25 (1.8)	10 (0.7)	28 (2.0)
All	7202	243 (3.4)	75 (1.0)	267 (3.7)	4266	102 (2.4)	42 (1.0)	115 (2.7)
Girls								
MZ	2790	144 (5.2)	44 (1.6)	156 (5.6)	1811	66 (3.6)	19 (1.0)	70 (3.9)
DZ	2303	167 (7.3)	53 (2.3)	179 (7.8)	1308	52 (4.0)	12 (0.9)	56 (4.3)
OS	2302	77 (3.3)	32 (1.4)	85 (3.7)	1411	34 (2.4)	9 (0.6)	35 (2.5)
All	7395	388 (5.2)	129 (1.7)	420 (5.7)	4530	152 (3.4)	40 (0.9)	161 (3.6)

Note. Item 5 = "behaves like opposite sex"; item 110 = "wishes to be of opposite-sex"; both items = summation of Items 5 and 110; MZ = monozygotic; DZ = dizygotic; OS = opposite-sex DZ twins.

daily functioning (about 2%), if zygosity was unknown, or if data were missing for the CBCL gender items. The first column of Table 4 provides an overview of the number of twin pairs with complete data.

Measures

At ages 7 and 10 years, mothers completed the Dutch version of the CBCL/4-18 (Achenbach, 1991), a questionnaire developed to measure problem behavior in 4 to 18 year-old children. The CBCL consists of 118 items that are rated on a 3-point scale (0 = "not true," 1 = "somewhat or sometimes true," and 2 = "very true or often true"). Cross-gender behavior was defined using two items on the CBCL: "behaves like opposite sex" (Item 5) and "wishes to be of opposite sex" (Item 110). The scores of these two items were summed and a child was defined as affected when a least one of the items was rated with "1." To index childhood psychopathology, the Internalizing and Externalizing scales of the CBCL were used. The Internalizing scale consists of the Anxious/Depressed, Somatic Complaints, and Withdrawn subscales and the Externalizing scale consists of Aggressive and Rule Breaking Behavior subscales.

Statistical analyses

Prevalence rates

Prevalence rates were estimated with Mx (Neale, Boker, Xie, & Maes, 1999), a structural equation modeling package, in order to obtain unbiased estimates for data assessed in family members (as the data of member from the same family may not be independent). Likelihood-ratio tests were used to test for prevalence differences between MZ and DZ twins, between boys and girls, between twins from opposite-sex pairs and same-sex pairs, and between ages 7 and 10 years.

Genetic analysis

To summarize twin similarity, we used two indices: the probandwise concordance and the tetrachoric correlation. The probandwise concordance is an estimate of the probability that a twin is affected given that his or her co-twin is affected (Bailey et al., 2000). The proband concordance was computed as: [2(N concordant pairs)]/[2(N concordant pairs) + (N discordant pairs)].

To estimate heritability, the liability threshold model was used, which assumes that many genetic and environmental factors contribute to the liability of a disorder. The (small) effects of all these factors add up and form an underlying continuous distribution of liability. The underlying liability distribution has a mean of 0 and variance of 1. If a critical value of liability (the threshold) is passed, a person is affected; otherwise, a person is unaffected (Falconer, 1989; Neale & Cardon, 1992). The thresholds, expressed as a z-value of the normal distribution, are inferred from the prevalence.

Genetic analyses were also carried out with Mx, using maximum likelihood estimation. The basic idea of twin studies is summarized in Fig. 1. The variation in liability may be partitioned into variance due to additive genetic factors (A), non-additive genetic (dominance) effects (D), and unique or non-shared environmental factors (E). E refers to environmental factors that are not shared by twins and which make them less similar. The contribution of genetic and environmental factors to variance in liability for cross-gender behavior can be inferred from the different levels of genetic relatedness of MZ and DZ twins. Genetic influences predict similarity among relatives, while non-shared environmental effects are a source of phenotypic differences. Genetic

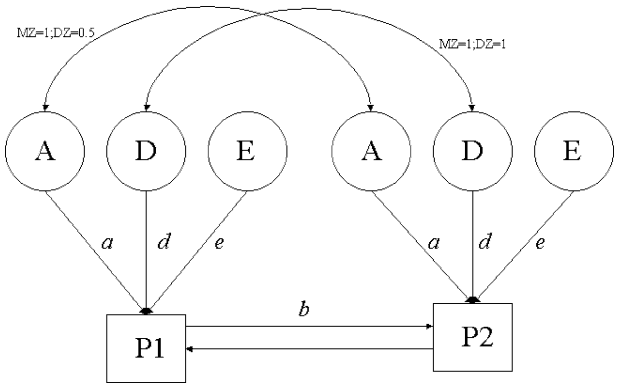

Fig. 1 Example of genetic model. *Note*. A: additive genetic factor; D: genetic dominance factor; E: unique environmental factor, P1: observed phenotype twin 1; P2: observed phenotype twin 2; a, d, and e: factor loadings on respectively additive genetic factor, genetic dominance factor, and environmental factor; b: contrast parameter

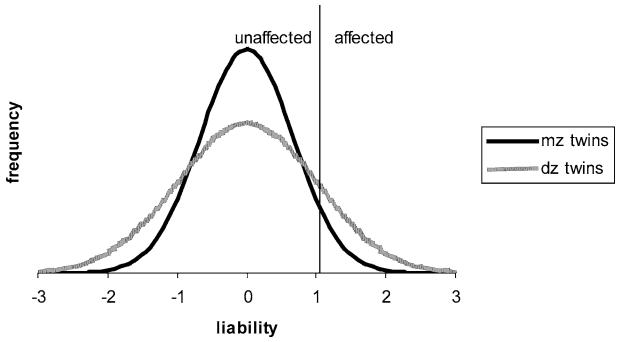

Fig. 2 Liability distribution for MZ and DZ twin pairs, who differ in prevalence but have the same threshold

(additive and dominance) factors are correlated 1.0 in MZ twins, as they are genetically identical. DZ twins share on average half of their segregating genes, giving a correlation of 0.50 for additive genetic factors and 0.25 for dominance genetic effects. The unique environment is, by definition, uncorrelated between two members of a pair. The parameters a, d, and e are factor loadings of the observed phenotype on the latent factors A, D, and E and indicate the strength of the relations between the latent factors and the observed phenotype. The proportion of the variance accounted for by genetic and environmental influences is calculated by squaring the parameters a, d, and e. The ADE model was chosen as the baseline model based on the pattern of MZ and DZ correlations for cross-gender behavior. This model was indicated because the MZ correlation was larger than twice the DZ correlation. If the MZ correlation was smaller than twice the DZ correlation, a model with additive genetic effects and shared environmental influences would have been more appropriate (Neale & Cardon, 1992).

The characteristic pattern of twin correlations of an ADE model (MZ correlations which are larger than twice the DZ correlations) may also the result of a rater contrast effect (Carey, 1986; Eaves, 1976). Rater contrast effects may be due to bias in the report of an informant when the informant rates the behavior of the child in comparison to the behavior to behavior of the co-twin. Therefore, in the model we allowed for a rater contrast effect for same-sex twin pairs. In Fig. 1, this contrast effect is represented by two arrows from P1 to P2 and from P2 to P1 and labeled with b. If a contrast effect is present, then the prevalence of a trait will be different in MZ and DZ twins. The presence of contrast effects will lead to lower variance in liability of cross-gender behavior, and this effect is stronger in MZ than in DZ twins. As shown in Fig. 2, a smaller liability variance, but a similar threshold, leads to a lower proportion of affected persons (Carey, 1986, 1992). The presence of a contrast effect will thus result in a lower prevalence of the disorder in MZ twins than in DZ twins.

Models were fitted to 2×2 contingency tables within each zygosity by sex group, yielding a measure of goodness of fit (χ^2). First, the full ADE model with a phenotypic interaction between the siblings (b) was evaluated. Parameter estimates were allowed to differ between boys and girls. The significance of sex differences in a, d, and e was tested by constraining them to be equal across sexes. A significant decrease in goodness-of-fit implies that the constraint was not allowed and that there was a significant sex difference. The same procedure was applied to test the significance of the parameters a, d, and b by constraining them at zero. Goodness-of-fit statistics obtained for the different models were compared with likelihood-ratio tests.

Association of behavior problems and cross-gender behavior

Nonparametric Mann–Whitney tests were used to test for differences in internalizing (INT) and externalizing (EXT) behavior problems between children with and without cross-gender behavior. Paired-sample *t*-tests were used to test whether children with cross-gender behavior had more internalizing than externalizing problems. In all analyses, we used the *T*-score of INT and EXT. T-scores were calculated as a linear transformation of the *z*-score, with a mean of 50 and a SD of 10, and were calculated separately for boys and girls. The analyses were performed on a dataset that randomly included the first or second born twin.

Results

Prevalence and distribution

Table 1 provides a summary of the prevalence rates for boys and girls of cross-gender behavior for MZ and same-sex and

Table 2 Distribution of the total score on the two Child Behavior Checklist items pertaining to cross-gender behavior

	Age 7				Age 10			
	Boys		Girls		Boys		Girls	
Score	*N*	%	*N*	%	*N*	%	*N*	%
0	6937	96.3	6975	94.3	4151	97.3	4369	96.4
1	197	2.7	297	4.0	76	1.8	123	2.7
2	60	0.8	91	1.2	33	0.8	26	0.6
3	9	.1	17	0.2	6	0.1	8	0.2
4	1	.0	15	0.2	0	0	4	0.1

opposite-sex DZ twins. It can be seen that the prevalence of cross-gender behavior was low, that these behaviors decreased with age, that the prevalence was higher in girls than in boys, and that the prevalence was higher for Item 5 ("behaves like opposite sex") than for Item 110 ("wishes to be of opposite sex").

Table 2 shows the distribution of the total score on cross-gender behaviors. Given that there were two CBCL gender items, the range of scores was 0–4. Firstly, maternal report of cross-gender behavior was relatively rare, with 94% to 97% of the children having no cross-gender behavior. Secondly, when present, the most frequent answer was "somewhat or sometimes true" and Item 5 was endorsed more frequently than Item 110. Of the children who were reported to have any cross-gender behavior, 74% and 71% of the 7-year-old boys and girls and 66% and 77% of the 10-year-old boys and girls had a score of 1. Mothers reported a total score of 2 in approximately 20% of the sample who had any cross-gender behavior. Scores of 3 were rare: only 4% to 5% of children with cross-gender behavior scored this high. Finally, scores of "frequently" for both items (score 4) appeared mainly in girls (3.6% of the 7 year olds, and 2.5% of the 10 year old girls who had any cross-gender behavior), but not in boys (only one boy at age 7 and none of the boys at age 10).

We examined whether the prevalence rates differed by sex, birth order, and zygosity. Table 3 presents the results of these tests for the 7-year-old twins (second to fourth columns) and the 10-year-old twins (last three columns). In same-sex twin pairs, the prevalence rate was different between boys and girls. As seen in Table 1, the prevalence was higher in girls than in boys. At age 10, the sex difference was smaller, and was no longer significant at the .05 level. We also tested whether the prevalence differed between same-sex MZ and DZ twins. For boys, there was a significant zygosity difference at age 10, but not at age 7. At age 10, more DZ twin boys were affected than MZ twin boys. For girls, there was a significant zygosity effect at age 7, but not at age 10. The prevalence in the 7-year-old DZ twin girls was higher than the prevalence in MZ twin girls. The differences in MZ and DZ prevalence may point to contrast effects; accordingly, a contrast parameter was included in the genetic models.

The last two rows of Table 3 give the results for the tests of whether the prevalence rate depended on the sex of the co-twin. For boys, the prevalence in 7-year-old same-sex DZ twins was similar to the prevalence in twins with a female co-twin, but 10-year-old DZ male twins with a female co-twin showed less cross-gender behavior than DZ twin boys with a male co-twin. For girls, at both ages the prevalence of cross-gender behavior was lower in opposite-sex pairs than in same-sex pairs.

Twin correlations

Table 4 shows concordances and tetrachoric twin correlations for cross-gender behavior at age 7 and 10 years. At age 10, there were no concordant DZ twin pairs and, therefore, these correlations were inestimable. At both ages, the MZ concordances (both "affected") were higher than the DZ concordances, suggesting at least some genetic influences. However, we also obtained a remarkable pattern of opposite-sex twin concordances. While the same-sex DZ concordances were near zero, the opposite-sex concordances were substantially higher. These contrasting findings for same-sex and opposite-sex twin concordances suggest that different processes influence the liability variance of cross-gender behavior. Because of the unexpected opposite-sex concordance, the genetic analyses started with data from same-sex twins

Table 3 Tests of differences in prevalence in cross-gender behavior for birth order, sex, zygosity, and sex of co-twin

	Age 7			Age 10		
	$\Delta\chi^2$	Δdf	*p*	$\Delta\chi^2$	Δdf	*p*
Boys vs. girls	40.51	2	<.001	5.82	2	ns
MZ boys vs. same-sex DZ boys	1.60	1	ns	4.27	1	.04
MZ girls vs. same-sex DZ girls	8.63	1	<.001	0.30	1	ns
OS DZ boys vs. same-sex DZ boys	0.84	1	ns	6.84	1	.01
OS DZ girls vs. same-sex DZ girls	35.28	1	<.001	7.02	1	.01

Note. MZ: monozygotic; DZ: dizygotic; OS: opposite-sex.

Table 4 Proband concordances and twin correlations for cross-gender behavior[a]

	U-U	A-A	U-A	Proband concordance	*r* (95% CI)[b]
Age 7					
MZM	1141	13	55	.32	.66 (.49–.79)
DZM	1130	2	98	.04	− .01 (− .30–.26)
MZF	1255	19	116	.25	.49 (.33–.62)
DZF	974	8	161	.09	.05 (− .15–.23)
DOS	2137	16	136	.19	.46 (.30–.59)
Age 10					
MZM	711	1	34	.06	.17 (− .27–.55)
DZM	628	0	51	.00	—
MZF	844	8	52	.24	.53 (.30–.71)
DZF	593	0	56	.00	—
DOS	1344	5	53	.16	.47 (.21–.67)

Note. U-U: both twins unaffected; A-A: both twins affected; U-A: one twin affected (A) and one twin unaffected (U); MZM: monozygotic males; DZM: dizygotic males; MZF: monozygotic females; DZF: dizygotic females; DOS: opposite-sex twins.

[a]Affected is defined as a score of 1 or more on the summation of two CBCL gender items.

[b]Tetrachoric correlation.

only (four group analysis) and then proceeded with the genetic analyses with five groups (including opposite-sex twin pairs).

Genetic analyses

The model fitting procedure started with the ADE-*b* model in which the parameters were allowed to differ between boys and girls. Because the contrast effects could be different among different groups (Eaves et al., 2000), it was first tested whether the interaction parameter (*b*) could be constrained to be equal across zygosity and sex. The results indicated no significant differences in *b* parameters among zygosity and sex groups and, therefore, in the models reported, the *b* parameter was constrained to be equal across sex and zygosity. Subsequently, it was tested whether the ADE-b model could be constrained to be equal across boys and girls. As shown in Table 5, no deterioration of the fit occurred; thus, there was no evidence for sex differences in genetic influences on the liability variance of cross-gender behavior.

In the next model, the significance of genetic dominance was tested by dropping D from the model. This did not result in a deterioration of the fit. Dropping the interaction parameter (*b*) from the model led to a significant deterioration of fit at both ages. Thus, the best model seemed to be an AE-*b* model, i.e. a model with additive genetic effects, non-shared environmental effects and a contrast effect, without sex differences in estimates for *a*, *e*, and *b*.

Next, data from the opposite-sex group were included in the genetic analyses. Because of the high twin correlations in the opposite-sex twins, we did not specify a contrast parameter (b) for this group. As shown in Table 5, the inclusion of opposite-sex twins led to the same conclusion. Table 6 provides the standardized estimates of the genetic and environmental influences. At both ages, a large part of the variance was for accounted by genetic factors (around 70%) and, in same-sex pairs, there was a negative phenotypic interaction (around − 0.20) at both ages, indicating a contrast effect when mothers rate same-sex offspring.

Cross-gender behavior and behavior problems

In order to examine the relation between cross-gender behavior and childhood psychopathology, we compared the levels of INT and EXT between children with and without cross-gender behavior. Secondly, we tested whether problem behavior depended on the level of cross-gender behavior by comparing INT and EXT between children with a score of

Table 5 Summary of fitting genetic models of cross-gender behavior at age 7 and 10

	Age 7			Age 10		
4 groups	χ^2	df	*p*	χ^2	df	*P*
ADE-b with sex differences	3.53	5	ns	9.55	5	ns
Tests	$\Delta\chi^2$	Δdf		$\Delta\chi^2$	Δdf	
No sex differences	1.42	2	ns	1.37	2	ns
No genetic dominance (drop D)	.03	1	ns	1.42	1	ns
No interaction effect (drop b)	11.80	1	<.001	9.48	1	<.001
5 groups	χ^2	df		χ^2	df	
ADE-b with sex differences and genetic covariance in DOS free	4.43	6	ns	12.86	6	.04
Tests	$\Delta\chi^2$	Δdf		$\Delta\chi^2$	Δdf	
Genetic covariance in DOS fixed to .5	0	1	ns	0	1	ns
No sex differences	1.48	2	ns	1.24	2	ns
No genetic dominance (drop D)	0	1	ns	0	1	ns
No interaction effect (drop b)	23.86	1	<.001	14.42	1	<.001

Note. DOS: opposite-sex twins.

Table 6 Estimates and 95% confidence intervals for genetic influences (a^2), nonshared environmental influences (e^2), and phenotypic interaction (b) of cross-gender behavior at age 7 and 10

	a^2	e^2	b
Age 7, 5 groups	.77 (.69–.83)	.23 (.16–.31)	− .19 (− .24–.12)
Age 10, 5 groups	.71 (.56–.81)	.29 (.19–.44)	− .22 (− .30–.13)

1 and children with a score of 2 or higher on cross-gender behavior. This analysis was limited to two groups because the number of children with a score of 3 or higher was small. Thirdly, we tested whether children with cross-gender behavior had more internalizing than externalizing problems.

Fig. 3 shows the means and the 95% confidence intervals of the INT and EXT T-scores across ages and sex. For each sex, age, and kind of problem behavior (INT and EXT) the tests showed significant differences between children with and without cross-gender behavior (all ps $< .01$). There was one exception: for the 10-year-old girls, the INT score did not differ between the groups with and without cross-gender behavior. In the second run of analyses, we tested whether the INT and EXT scores differed between the "1" and the "≥ 2" group. In Fig. 3, it appears that the "≥ 2" group showed more problem behaviors, but the analyses did not show significant differences between the "1" and the "≥ 2" group for any age or sex group. This result could probably explained by the small number in the "≥ 2" cross-gender group.

In the third series of analyses, the INT and EXT scores were compared to each other within the cross-gender groups (pooled over "1" and "≥ 2"). For boys, the level of INT and EXT did not differ (age 7: $t(139) = 1.13$, ns; age 10: $t(68) = -0.617$, ns), but for girls the EXT score was significantly higher than the INT score at both ages (age 7: $t(212) = 3.53$, $p < .01$; age 10: $t(82) = 3.83$, $p < .01$).

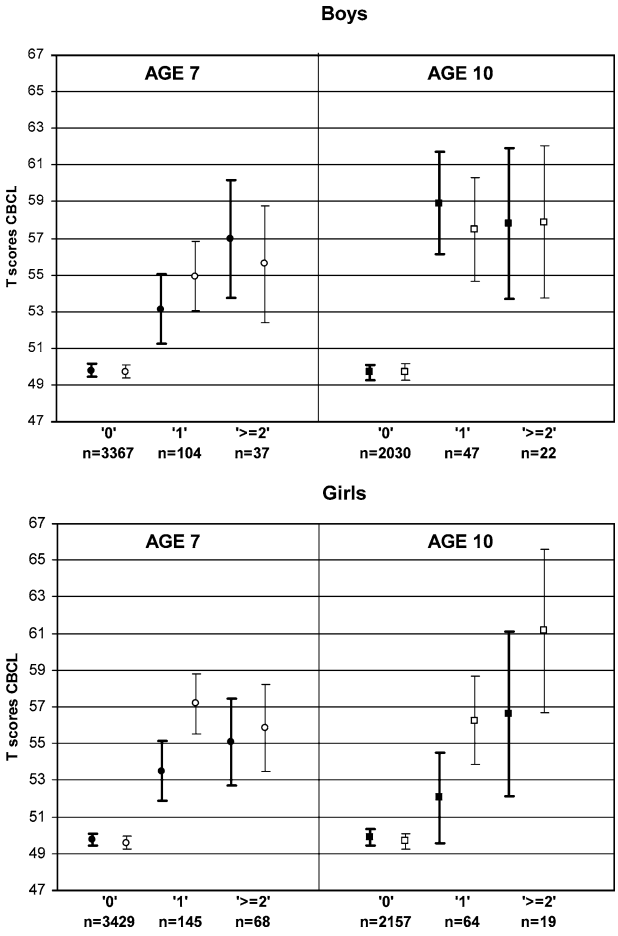

Fig. 3 Internalizing and Externalizing CBCL T-scores and their 95% confidence interval by cross-gender behavior subgroup and age for boys (upper part) and girls (lower part). Bold/thick lines represent Internalizing *T*-scores and thin lines represent Externalizing *T*-scores

Discussion

In a large sample of 7- and 10-years-old twin pairs, we examined the prevalence of cross-gender behavior. The data of genetically related individuals allowed the exploration of the genetic and environmental contributions to the variability in liability of cross-gender behavior. The most important findings can be summarized as following: (1) prevalence of cross-gender behavior was not as uncommon as the GID syndrome; (2) girls showed more cross-gender behavior than boys; (3) the frequency of cross-gender behavior decreased with age; (4) a large part of the variance in liability of cross-gender behavior was accounted for by genetic factors; and (5) cross-gender behavior was associated with higher scores on both the internalizing and externalizing scales of the CBCL.

The prevalence of "any cross-gender behavior" differed between boys and girls. At age 7, more girls than boys showed cross-gender behavior. About 5.7% of the 7-year-old girls showed any cross-gender behavior, while 3.7% of the 7-year-old boys showed cross-gender behavior. At age 10, the differences in prevalence between girls and boys were reduced. These results agree with the finding that in non-clinical samples more girls than boys engage in cross-gender behaviors (Bradley & Zucker, 1997; Coolidge et al., 2002; Sandberg & Meyer-Bahlburg, 1994). With the present data, it is difficult to say whether these sex differences are the result of a greater tolerance of cross-gender behavior for girls by the rater or that girls really engage more in cross-gender behaviors.

The fact that more girls than boys engage in cross-gender behavior than boys seems to be in contrast with the consistent finding of more referred boys. In a Canadian sample of 3 to 12-year-old referred children, the sex ratio was 6.6:1 of boys vs. girls (Zucker et al., 1997). Also, in a Dutch sample, more boys than girls were referred and the overall sex ratio was 4.7:1 (Cohen-Kettenis et al., 2003). Zucker et al. (1997) concluded that social factors partly account for the sex differences in referral rates. Feminine behaviors in boys are less accepted than masculine behaviors in girls and the threshold for clinical referral is probably higher for girls than for boys. A greater tolerance of cross-gender behavior for girls was confirmed by a study of Cohen-Kettenis et al. (2003). In their study, girls were later referred than boys, referred girls were more likely to meet the complete DSM criteria, and girls had higher ratings on the two CBCL gender items.

The prevalence rates and the boy-girl ratio were lower than those reported by Zucker et al. (1997) (based on the data from Achenbach & Edelbrock, 1981). Although their approach was the same as ours (using the two CBCL items measured in a non-clinical sample), their estimates for 7-year-old boys and girls were 6% and 12%, respectively. However, our prevalence rates were comparable with those reported in a Dutch sample of 1200 non-referred boys and girls aged 4–11 years (Verhulst, van der Ende, & Koot, 1996). In that sample, parents reported that 5% of the girls and 2.6% of the boys sometimes or frequently behaved like the opposite sex.

A possible explanation for the difference in prevalence rates between our study and that of the study of Zucker et al. (1997) are differences in normative gender roles in both societies. Perhaps the tolerance for cross-gender behavior is greater in Dutch society than it is in North-America. In line with this, a Dutch-Canadian comparison of the characteristics of children with GID revealed that Canadian children were one year younger at time of referral than the Dutch children (Cohen-Kettenis et al., 2003).

Cross-gender behavior was a highly heritable trait. About 70% of the variance in liability of cross-gender behavior was influenced by additive genetic factors. The magnitude of the genetic and environmental influences did not differ between boys and girls. Also, the results did not indicate changing heritabilities with age.

The heritability estimate obtained in this study was larger than heritability reported in other studies that looked at atypical gender behavior (Bailey et al., 2000; Iervolino et al., 2005; Knafo et al., 2005). The different approaches in identifying cross-gender behavior, the differences in ages, and the differences in tolerance toward cross-gender behavior may explain these variations. Knafo et al. measured gender role behaviors based on the PSAI, while in our study cross-gender behavior was derived from two CBCL items. The contents of items in the two ratings were different. In the PSAI, items were related to more specific and concrete kinds of behavior (e.g., whether children show interest in real cars or the extent to which children play with girls), whereas in the CBCL a more general judgment of the parents is required. In addition, in the study of Knafo et al. the twins were 3–4 years old, and younger children may be in a different phase of gender identity development. It is possible that in an earlier developmental phase, the influence of shared environment is a more important factor than at later ages. From genetic developmental research, it is known that heritability is not a fixed parameter, but may change during life. For example, for IQ it is known that the relative importance of genetic factors increases with age (Posthuma, de Geus, & Boomsma, 2002).

The finding of a high heritability does not mean that environmental manipulations can not contribute to changes in the expression of the trait. It is becoming more and more clear that the interplay between genetic and environmental factors (G × E interactions) is important for behavioral traits (Rutter, 2003). This interaction yields the possibility of the modification of genetic risk factors by environmental risk factors, but also that specific genotypes may have different sensitivity to environmental risk factors. For example, Caspi et al. (2002) showed that genotype can moderate the effect of maltreatment in the development of antisocial problems. As G x E findings may be significant for multiple traits, this should be a focus in future research.

The twin analyses revealed evidence for the presence of a rater contrast effect in the evaluation of the gender behaviors of same-sex twin pairs. Two findings accounted for this contrast effect. First, the number of concordant same-sex DZ twins was lower than expected on the basis of additive genetic effects. Secondly, the prevalence rate of cross-gender behavior differed between MZ and DZ twins. Same-sex DZ twins tended to show more cross-gender behaviors than MZ twins. When parents are asked to evaluate and report upon the behavior of their children, they may contrast the twins' behaviors (Carey, 1986; Eaves, 1976). However, it is also possible that a contrast effect represents an interaction effect between the twins. In that case, the behavior of one twin has an inhibitory effect on the behavior of the other twin (Carey, 1986; Eaves, 1976; Simonoff et al., 1995). With the current data, it was not possible to distinguish between sibling interaction effects and rater contrast effects. The higher concordances and the lack of a contrast effect in opposite-sex twins may favor the interpretation of rater contrast effect.

The finding that DZ girls with a male co-twin showed less cross-gender behavior than DZ girls with a female co-twin was somewhat surprising. The hypothesis of hormonal transfer during gestation assumes that a female twin sharing the womb with a male co-twin is exposed to higher levels of masculine hormones than a female twin sharing the womb with a female co-twin. As consequence, female twins with a male

co-twin would show more masculine behaviors. As DZ twins with an opposite-sex co-twin tend to show less atypical gender behaviors, our findings do not fit with the hypothesis of hormonal transfer during gestation. At the same tine, these findings also did not support social learning theories. According these theories, it is expected that children imitate the behavior of their siblings. Boys with older brothers and girls with older sisters, showed more sex-typed behaviors than same sex singletons, who, in turn, were more sex-typed than children with other-sex siblings (Rust et al., 2000). These processes may be different for children of the same age or alternatively parents of opposite twins are less focused on sex-stereotyped behaviors than parents of same-sex DZ girls (Rust et al., 2000).

Our findings of more problem behaviors in children with cross-gender behavior agree with general findings of more problem behaviors in children with GID reported (Cohen-Kettenis et al., 2003; Zucker & Bradley, 1995). However, our findings differ in some aspects with that of studies that examined children with a clinical diagnose of GID. In the clinical samples, children with GID had Internalizing and Externalizing scores in the clinical range (Cohen-Kettenis et al., 2003). We found that the affected group had larger Internalizing and Externalizing scores, but the means of the problem behaviors were lower than the means of clinical groups (Verhulst et al., 1996). Another general finding was the predominance of Internalizing problems in children with GID, but we did not find evidence for this. If there were any differences, then it was just in the opposite direction: children with cross-gender behaviors showed more Externalizing than Internalizing problems.

As in the case of the unclear nature of the causal relationship between GID and anxiety, a variety of explanations are possible for the higher level of problem behaviors in children with cross-gender behavior. First, it may be that any cross-gender behavior, as reported by parents at ages 7 and 10, are already associated with higher rates of emotional and behavioral problems. These relations may be at a genetic level (e.g., cross-gender behavior may share common genetic vulnerabilities with Internalizing and Externalizing problems). Equally plausible is that there are already environmental consequences of cross-gender behavior (e.g., that the children endure consequences of their behaviors). In addition, parents may contribute to this relation by contributing their own bias. They may see the cross-gender behavior as a part of a constellation of problem behavior and over-report the prevalence of Internalizing and Externalizing symptoms. These possibilities can be explored by including other raters (teachers and the children themselves when they are a bit older) and will be the focus of future studies by our group. Whatever the etiologic implications, our data support the contention that children's cross-gender behavior, as reported by their mothers, is associated with higher rates of both Internalizing and Externalizing symptoms in boys and girls at ages 7 and 10.

Clinical implications

Although we did not directly test the characteristics of children who receive health care versus those who do not, there are a few points that clinicians can take from general population studies using genetic epidemiologic methods such as ours. First, cross-gender behavior is relatively rare, and high scores on the two items described herein were associated with increased risk for emotional and behavioral problems. Thus, in general settings, reports of cross-gender behavior could be used as a stimulus to screen for emotional behavioral problems. Second, more boys than girls are typically referred for treatment, yet our data indicate that more girls than boys engage in cross-gender behavior and the emotional/behavioral toll of these behaviors on girls is at least as high as they are on boys. Such information is useful to the individual clinician, and to health care planners, who may be informed by being more aware of this bias, and screening more aggressively for cross-gender behavior in girls in general medical settings (e.g., a pediatrician's office), because it is apparent that girls are not being "referred" on at a rate consistent with boys. Finally, screening for emotional behavioral problems with instruments such as the one used in this study will allow clinicians to simultaneously screen for cross-gender behavior and associated emotional and behavioral problems.

Study limitations

A limitation of the study was that it was not a study of GID. We did not complete DSM clinician interviews but measured cross-gender behavior with two items of the CBCL in a large general population sample of twins. Prior research showed that children who met the DSM criteria for GID had higher mean ratings on these two CBCL gender items than children who did not meet the complete DSM criteria (Cohen-Kettenis et al., 2003). However, the association between scores on the two CBCL items and the GID scores were found in study of children drawn from a clinical population, and it is unclear what the meaning of these items is in a non-clinical population.

Further, some caution is in order with our conclusion since we have used only two items to measure cross-gender behavior. Ideally, we would like to obtain more objective measures of cross-gender behavior. Using objective measures may give more reliable results and therefore the estimates of genetic and environmental influences will be less biased. However, in order to have adequate power in epidemiologic studies it is simply not cost effective to obtain objective measures on general population samples. Information from studies such

as ours can set the stage for a well-grounded selection approach (e.g., selection of extremes) for objective measures.

Another limitation was that the heritability estimates were based on mother ratings only. By using only one informant, the rating may be biased. Regarding cross-gender behavior, it is possible that the rating of his/her child may be biased by the relationship with the child or by her/his own personality traits or by response style (e.g., stereotyping, employing different normative standards). If an informant has the tendency to overestimate or underestimate scores consistently, then the rates of pairs of twins will be more similar and will result in a shared environmental effect. This is a pattern of results we did not observe. That the mother rating is not merely a reflection of rater bias, is supported also by the finding of high correlations between mother and father ratings of gender-related behaviors (Johnson et al., 2004; Zucker et al., 1997). Johnson et al. reported a mother-father correlation of .90 using the Gender Identity Questionnaire for Children and Zucker et al. reported a mother-father correlation of .69 for the two CBCL gender items.

In summary, in middle childhood, cross-gender behavior was influenced by genetic factors and there appeared to be no change in heritability from 7 to 10 years. An additional important finding was that shared environmental factors did not contribute to the variance of liability in cross-gender behavior. Only nonshared environmental factors played a significant role in the expression of cross-gender behavior. Further, from this report, it appears that cross-gender behavior, as reported by the mothers, was associated with increases in Internalizing and Externalizing problems. Our data are derived from an on-going longitudinal study, and we aim to study the relation between cross-gender behavior and behavioral problems in adolescence and adulthood using a multi-informant, developmental approach.

Acknowledgments This work was supported by NWO Grant numbers 575-25-006, 575-25-012 and 904-57-94 (Boomsma, P.I.), by the Centre for Neurogenomics and Cognition Research (CNCR) of the Vrije Universiteit, Amsterdam and by NIMH Grant number MH58799 (Hudziak, P.I.).

References

Achenbach, T. M. (1991). *Manual for the Child Behavior Checklist/4–18 and 1991 Profile*. Burlington: University of Vermont, Department of Psychiatry.

Achenbach, T. M., & Edelbrock, C. S. (1981). Behavioral problems and competencies reported by parents of normal and disturbed children aged four through sixteen. *Monographs of the Society for Research in Child Development*, *46*(1, Serial No. 188).

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders* (4th ed., text rev.). Washington, DC: Author.

Anders, S. M., Vernon, P. A., & Wilbur, C. J. (2006). Finger-length ratios show evidence of prenatal hormone-transfer between opposite-sex twins. *Hormones and Behavior*, *49,* 315–319.

Bailey, J. M., Dunne, M. P., & Martin, N. G. (2000). Genetic and environmental influences on sexual orientation and its correlates in an Australian twin sample. *Journal of Personality and Social Psychology*, *78,* 524–536.

Berenbaum, S. A., & Hines, M. (1992). Early androgens are related to childhood sex-typed toy preferences. *Psychological Science*, *3,* 203–206.

Boomsma, D. I., Busjahn, A., & Peltonen, L. (2002). The classical twin study and beyond. *Nature Review Genetics*, *3,* 872–882.

Boomsma, D. I., van Beijsterveldt, C. E., & Hudziak, J. J. (2005). Genetic and environmental influences on anxious/depressed behavior during childhood: A large twin study. *Genes, Brain and Behavior*, *4,* 466–481.

Boomsma, D. I., Vink, J. M., van Beijsterveldt, C. E., de Geus, E. J., Beem, L., & Mulder, E. J. (2002). Netherlands Twin Register: A focus on longitudinal research. *Twin Research*, *5,* 1–6.

Bradley, S. J., & Zucker, K. J. (1997). Gender identity disorder: A review of the past 10 years. *Journal of the American Academic of Child and Adolescent Psychiatry*, *36,* 872–880.

Carey, G. (1986). Sibling imitation and contrast effects. *Behavior Genetics*, *16,* 319–341.

Carey, G. (1992). Twin imitation for antisocial behavior: Implications for genetic and family environment research. *Journal of Abnormal Psychology*, *101,* 18–25.

Carver, P. R., Yunger, J. L., & Perry, D. G. (2003). Gender identity and adjustment in middle childhood. *Sex Roles*, *49,* 95–109.

Caspi, A., McClay, J., Moffitt, T. E., Mill, J., Martin, J., Craig, I. W., et al. (2002). Role of genotype in the cycle of violence in maltreated children. *Science*, *297,* 851–854.

Cohen-Bendahan, C. C., Buitelaar, J. K., van Goozen, S., & Cohen-Kettenis, P. T. (2004). Prenatal exposure to testosterone and functional cerebral lateralization: a study in same-sex and opposite-sex twin girls. *Psychoneuroendocrinology*, *47,* 230–237.

Cohen-Bendahan, C. C., Buitelaar, J. K., van Goozen, S., Orlebeke, J., & Cohen-Kettenis, P. T. (2005). Is there an effect of prenatal testosterone on aggression and other behavioral traits? A study comparing same-sex and opposite-sex twin girls. *Hormones and Behavior*, *47,* 230–237.

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

Cohen-Kettenis, P. T., Owen, A., Kaijser, V., Bradley, S. J., & Zucker, K. J. (2003). Demographic characteristics, social competence, and behavior problems in children with gender identity disorder: A cross-national, cross-clinic comparative analysis. *Journal of Abnormal Child Psychology*, *31,* 41–53.

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

Coolidge, F. L., Thede, L. L., & Young, S. E. (2002). The heritability of gender identity disorder in a child and adolescent twin sample. *Behavior Genetics*, *32,* 251–257.

Eaves, L. J. (1976). A model of sibling effects in man. *Heredity*, *36,* 206–214.

Eaves, L. J., Rutter, M., Silberg, J. L., Shillady, L., Maes, H., & Pickles, A. (2000). Genetic and environmental causes of covariation in interview assessments of disruptive behavior in child and adolescent twins. *Behavior Genetics*, *30,* 321–334.

Falconer, D. S. (1989). *Introduction to quantitative genetics* (3rd ed.). New York: Longman.

Goldsmith, H. H. (1991). A zygosity questionnaire for young twins: A research note. *Behavior Genetics*, *21,* 257–270.

Golombok. S., & Rust, J. (1993). The Pre-School Activities Inventory: A standardized assessment of gender role in children. *Psychological Assessment*, *5,* 131–136.

Henderson, B. A., & Berenbaum, S. A. (1997). Sex-typed play in opposite-sex twins. *Developmental Psychobiology*, *31,* 115–123.

Iervolino, A. C., Hines, M., Golombok, S. E., Rust, J., & Plomin, R. (2005). Genetic and environmental influences on sex-typed behavior during the preschool years. *Child Development*, *76,* 826–840.

Johnson, L. L., Bradley, S. J., Birkenfeld-Adams, A. S., Kuksis, M. A. R., Maing, D. M., Mitchell, J. N., & Zucker, K. J. (2004). A parent-report gender identity questionnaire for children. *Archives of Sexual Behavior*, *33,* 105–116.

Knafo, A., Iervolino, A. C., & Plomin, R. (2005). Masculine girls and feminine boys: Genetic and environmental contributions to atypical gender development in early childhood. *Journal of Personality and Social Psychology*, *88,* 400–412.

Koopmans, J. R., Boomsma, D. I., Heath, A. C., & van Doornen, L. J. (1995). A multivariate genetic analysis of sensation seeking. *Behavior Genetics*, *25,* 349–356.

Miller, E. (1994). Prenatal sex hormone transfer: A reason to study opposite-sex twins. *Personality and Individual Differences*, *4,* 511–529.

Neale, M. C., Boker, S. M., Xie, G., & Maes, H. H. (1999). *Mx; Statistical modeling* (5th ed.). Virginia: Department of Psychiatry, Richmond, Virginia.

Neale, M. C., & Cardon, L. R. (1992). *Methodology for genetic studies of twins and families*. Dordrecht: Kluwer Academic Publishers.

Posthuma, D., de Geus, E. J., & Boomsma, D. I. (2002). Genetic contributions to anatomical, behavioral, and neurophysiological indices of cognition. In R. Plomin, J. C. Defries, I. W. Craig, & P. McGuffin (Eds.), *Behavioral genetics in the postgenomic era* (pp. 141–161). Washington, DC: American Psychological Association.

Resnick, S. M., Gottesman, I. I., & McGue, M. (1993). Sensation seeking in opposite twins: An effect of prenatal hormones? *Behavior Genetics*, *23,* 323–329.

Rietveld, M. J., VanderValk, J. C., Bongers, I. L., Stroet, T. M., Slagboom, P. E., & Boomsma, D. I. (2000). Zygosity diagnosis in young twins by parental report. *Twin Research*, *3,* 134–141.

Rodgers, C. S., Fagot, B. I., & Winebarger, A. (1998). Gender-typed toy play in dizygotic twin pairs: A test of hormone transfer theory. *Sex Roles*, *39,* 173–184.

Rust, J., Golombok, S., Hines, M., Johnston, K. Golding, J., & The ALSPAC Study Team. (2000). The role of brothers and sisters in the gender development of preschool children. *Journal of Experimental Child Psychology*, *77,* 292–303.

Rutter, M. (2003). Commentary: Nature-nurture interplay in emotional disorders. *Journal of Child Psychology and Psychiatry*, *44,* 934–944.

Ryan, B. C., & Vandenbergh, J. G. (2002). Intrauterine position effects. *Neuroscience and Biobehavioral Reviews*, *26,* 665–678.

Simonoff, E., Pickles, A., Hewitt, J., Silberg, J., Rutter, M., Loeber, R., et al. (1995). Multiple raters of disruptive child behavior: Using a genetic strategy to examine shared views and bias. *Behavior Genetics*, *25,* 311–326.

Verhulst, F. C., van der Ende, J., & Koot, H. (1996). *Handleiding voor de CBCL/4–18 [Manual for the CBCL/4–18]*. Rotterdam: Afdeling Kinder- en Jeugdpsychiatrie Sophia Kinderziekenhuis, Academisch Ziekenhuis Rotterdam, Erasmus Universiteit Rotterdam.

Yunger, J. L., Carver, P. R., & Perry, D. G. (2004). Does gender identity influence children's psychological well-being? *Developmental Psychology*, *40,* 572–582.

Zucker, K. J. (2005). Gender identity disorder in children and adolescents. *Annual Review of Clinical Psychology*, *1,* 467–492.

Zucker, K. J., & Bradley, S. J. (1995). *Gender identity disorder and psychosexual problems in children and adolescents*. New York: Guilford Press.

Zucker, K. J., Bradley, S. J., & Lowry Sullivan, C. B. (1996). Traits of separation anxiety in boys with gender identity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, *35,* 791–798.

Zucker, K. J., Bradley, S. J., & Sanikhani, M. (1997). Sex differences in referral rates of children with gender identity disorder: Some hypotheses. *Journal of Abnormal Child Psychology*, *25,* 217–227.