Longitudinal Stability of the CBCL-Juvenile Bipolar Disorder Phenotype: A Study in Dutch Twins

Dorret I. Boomsma, Irene Rebollo, Eske M. Derks, Toos C.E.M. van Beijsterveldt, Robert R. Althoff, David C. Rettew, and James J. Hudziak

Background: *The Child Behavior Checklist–juvenile bipolar disorder phenotype (CBCL-JBD) is a quantitative phenotype that is based on parental ratings of the behavior of the child. The phenotype is predictive of* DSM-IV *characterizations of BD and has been shown to be sensitive and specific. Its genetic architecture differs from that for inattentive, aggressive, or anxious–depressed syndromes. The purpose of this study is to assess the developmental stability of the CBCL-JBD phenotype across ages 7, 10, and 12 years in a large population-based twin sample and to examine its genetic architecture.*

Methods: *Longitudinal data on Dutch mono- and dizygotic twin pairs (N = 8013 pairs) are analyzed to decompose the stability of the CBCL-JBD phenotype into genetic and environmental contributions.*

Results: *Heritability of the CBCL-JBD increases with age (from 63% to 75%), whereas the effects of shared environment decrease (from 20% to 8%). The stability of the CBCL-JBD phenotype is high, with correlations between .66 and .77 across ages 7, 10, and 12 years. Genetic factors account for the majority of the stability of this phenotype. There were no sex differences in genetic architecture.*

Conclusions: *Roughly 80% of the stability in childhood CBCL-JBD is a result of additive genetic effects.*

Key Words: Childhood bipolar affective disorder, genetics, twins

The existence, prevalence, and taxonomy of juvenile bipolar disorder (JBD) have been the focus of considerable debate. Although the pediatric and adolescent forms of the illness increasingly are being recognized as valid diagnoses, how best to characterize children continues to be a focus of extensive investigation (Biederman et al 1998; Faedda et al 1995; Geller and Luby 1997; Weller et al 1995). One area of discussion surrounds the degree to which the diagnosis of JBD in children is associated with a more classic DSM-IV profile in adulthood (Carlson et al 2000; Geller and Luby 1997; Leibenluft et al 2003). For example, the adult-onset form of BD is associated with discrete episodes of mania or hypomania and depression, whereas this is not reported as common among children with JBD, in whom the episodes are more often of long duration, with rapid cycling and mixed mania (National Institute of Mental Health 2001). Given the lack of prospective data (Faedda et al 2004), it is not surprising that little is known about the developmental stability and change of JBD using standard DSM approaches.

Several groups have described a JBD profile on the Child Behavior Checklist (CBCL; Achenbach 1991) that differs from the CBCL profiles of children with other DSM disorders (Biederman et al 1995; Geller et al 1998; Wals et al 2001). Children with JBD have a CBCL profile that includes high levels on the Aggressive Behavior (AGG), Anxious/Depressed Behavior (A/D), and Attention Problems (AP) syndrome scales. The extent to which the CBCL-JBD phenotype predicts DSM-III-R and DSM-IV diagnoses of BD and delineates JBD from other childhood psychiatric diagnoses such as attention-deficit/hyperactivity disorder (ADHD; Biederman et al 1995), depression, and other disruptive behavior disorders (Kahana et al 2003) has been examined across samples, countries, and across methodologies (Althoff et al 2005). Results have been replicated across age groups (Biederman et al 1995; Carlson and Kelly 1998; Dienes et al 2002; Geller et al 1998; Hazell et al 1999; Wals et al 2001), across treatment settings (inpatient, outpatient), and across cultures (American, Dutch, Brazilian, Australian). In a recent study, Faraone and colleagues (2005) used a receiver-operating characteristic curve (ROC) analysis on the profile of elevated AP, AGG, and A/D to predict DSM-III-R BD in children enrolled in a family study of ADHD as well as in their siblings. The area under the curve statistic for prediction of DSM-III-R BD using a summed score of the three scales was as high as .97 for children with a current diagnosis of BD.

Recently, we described the prevalence of CBCL-JBD in a general-population twin sample and contrasted the genetic architecture of this phenotype (Hudziak et al, in press) to that of AP. These data indicated that the prevalence of CBCL-JBD in children is ~1% in boys and girls at ages 7, 10, and 12 years. At each age, variation in CBCL-JBD was influenced by additive genetic, shared, and unique environmental factors, with additive genetic influences accounting for the largest part of the variance. This is in contrast to the modeling of AP, which was influenced by additive and nonadditive (dominance) genetic effects without shared environmental influences. Liability threshold models of CBCL-JBD versus CBCL-AP also showed that the CBCL-JBD phenotype is unlikely to be an extreme version of CBCL-AP.

Latent class analyses of AGG, A/D, and AP symptoms in these twins showed that a seven-class model fit best for girls, and an eight-class model, best for boys. The most common class for both boys and girls was one without symptoms. The CBCL-JBD severe latent class was the least common—and was the only one that had significant elevations on the suicidal items of the CBCL. High heritability of the CBCL-JBD was demonstrated, with higher odds ratios between monozygotic twins than between dizygotic twins who fell into this latent class (Althoff et al 2006).

In this article, we explore the developmental stability of the CBCL-JBD phenotype and determine the genetic and environmental contributions to stability. One of the challenges to understanding the prevalence and implication of child psychopathology is the confound of development (Hudziak et al 2000). Children, their brains and their behaviors, change over the course of development. For instance, children with ADHD are

From the Department of Biological Psychology (DIB, IR, ED, TvB), Vrije Universiteit, Amsterdam, The Netherlands; Massachusetts General Hospital, Harvard Medical School (RRA), Boston, Massachusetts; and the Department of Psychiatry (DR, JJH), University of Vermont, Burlington, Vermont.

Address reprint requests to Dr. Dorret Boomsma, Department of Biological Psychology, Vrije Universiteit, Van der Boechorststraat 1, 1081 BT, Amsterdam, The Netherlands; E-mail: dorret@psy.vu.nl.

Received July 21, 2005; revised February 10, 2006; accepted February 10, 2006.

0006-3223/06/$32.00
doi:10.1016/j.biopsych.2006.02.028

often less hyperactive as they grow up, and a certain percentage of individuals no longer will suffer from ADHD as adults (Mannuzza et al 2003). Similarly, aggression diminishes in boys and girls over the course of development (Hudziak et al 2003; Stanger et al 1997). In addition, genetic and environmental influences on behavior may change with development. Changes in genetic and environmental influences have been reported for AP, AGG, and A/D behavior across development. Change across development can be in the type of genetic and environmental influences, as with the study of AP, in which the importance of genetic dominance differs at different ages (Rietveld et al 2003a, 2003b, 2004). Changes can also occur in the magnitude of the genetic and environmental influences, as our group reported elsewhere for both AGG (Hudziak et al 2003) and A/D (Boomsma et al 2005). By studying longitudinal twin data, it is possible to determine the stability and change of behavioral phenotypes across development and to assess the importance of genetic influences to stability. The purpose of this article is to assess the contribution of genetic and environmental factors to the stability of the CBCL-JBD phenotype across ages 7, 10, and 12 years in longitudinal data from a large sample of more than 8000 twin pairs whose behavior was assessed by their mothers.

Methods and Materials

Subjects and Procedure

Data for this study come from an ongoing longitudinal study that examines environmental and genetic influences on the development of problem behavior in 3- to 12-year-old twins. The families are volunteer members of the Netherlands Twin Register (NTR), kept by the Department of Biological Psychology at the Free University, Amsterdam (Boomsma et al 2002b). Starting in 1987, families with newborn twins were recruited. Currently, 40%–50% of all multiple births in The Netherlands are registered by the NTR. For the present study, we included data of 7-, 10-, and 12-year-old twin pairs (birth cohorts 1986–1996). Parents of twins were asked to fill in questionnaires about problem behaviors of the twins at ages 7, 10, and 12 years. After 2 months, a reminder was sent to nonresponders. If finances permitted, persistent nonresponders were contacted by phone. The total sample consists of 8013 twin families. Appendix 1 shows the distribution of the sample according to their participation across time. There were 2866 families who participated three times, at ages 7, 10, and 12 years; 2322 families participated twice, and 2825 families participated once. Most families who did not participate at ages 10 and 12 years had not reached the proper age for the twins (67% and 49%, respectively), and thus the questionnaires have not been sent to them yet. Among those who returned at least one survey at age 7, 10, or 12 years, 92% returned the survey at age 7 years, 80%, at age 10 years, and 70%, at age 12 years. If a family did not respond at a particular age, they again were approached for the next mailing, so that nonparticipants did not drop from data collection completely.

To examine the possible effect of sample attrition, data from twins whose families participated three times were compared with data from twins whose families participated at age 7 years but did not return the questionnaires at ages 10 years, 12 years, or both. The nonresponse group tended to show larger means in JBD at age 7 ($p < .01$). However, the Cohen's effect size was less than .20, which indicates a small effect. Furthermore, the possible effects of sample attrition on the results of the present study are minimized by inclusion of all available data in the analyses, irrespective of the number of times that a family participated.

The sample includes 1331 MZM (monozygotic male twins), 1338 DZM (dizygotic male twins), 1537 MZF (monozygotic female twins), 1250 DZF (dizygotic female twins), 1310 DOSMF (dizygotic opposite-sex twins, male twin born first, female twin born second), and 1240 DOSFM (dizygotic opposite-sex twins, female twin born first and male twin born second). Zygosity information was missing for seven pairs. For 1089 same-sex twin pairs, zygosity was based on blood group ($n = 370$) or deoxyribonucleic acid (DNA; $n = 719$) typing. For the remaining twins, zygosity was determined by questionnaire items about physical similarity and frequency of confusion of the twins by family and strangers. Classification of zygosity was based on a discriminant analysis of the questionnaire items and on zygosity based on blood or DNA typing in same-sex twin pairs. The questionnaire-based zygosity was correct for nearly 95% of the cases (Rietveld et al 2000). A comparison of parental socioeconomic status (SES) with the SES distribution for the general Dutch population showed a slightly higher frequency of the middle and higher SES groups (Rietveld et al 2003a). Representativeness of the sample at each age is discussed by Van Beijsterveldt et al (2003). A comparison of the twins' emotional and behavioral problems at age 3 year to that of singletons showed no differences between twins and singletons (Van den Oord et al 1996).

Measures

At ages 7, 10, and 12 years, problem behavior was measured with the CBCL/4-18 (Achenbach 1991). The CBCL consists of 118 items developed to assess behavioral and emotional problems. Mothers were asked to rate the behavior of the child in the preceding 6 months on a three-point scale. Children with more than four missing items for the JBD phenotype were excluded from the analyses. This occurred in fewer than 2.5% of the CBCLs. The JBD phenotype was defined as the square-root transformed sum of AP, AGG, and A/D.

Analyses

Analyses were conducted by using structural equation modeling, because it permits the simultaneous analysis of data from multiple groups and allows imposition of parameter constraints across groups. The statistical software packages Mx (Neale et al 2003) and Mplus (Muthen and Muthen 1998) were used. In longitudinal studies such as the current one, not all subjects have yet reached the oldest age, and not all subjects have taken part in the study at all ages. To be able to use all data, full-information maximum likelihood estimation with raw data was used. Twice the negative log-likelihood (−2LL) of the data for each family is calculated, and parameters are estimated so that the overall likelihood of the raw data is maximized.

The fit of the genetic models is evaluated against the fit of a saturated model, in which the covariance matrix and the mean structures are computed without any restriction. Submodels were compared with likelihood-ratio tests that are obtained by subtracting −2LL for a restricted nested model from that for a less restricted model ($\chi^2 = (-2LL_0) - (-2LL_1)$). The resulting test statistic has a χ^2 distribution with degrees of freedom (df) equal to the difference of the df between the two models. The χ^2 statistic is sensitive to large sample sizes. Given large sample sizes, small discrepancies between a model and the observed data can lead to the rejection of the model (Loehlin 2004). The χ^2 difference test applied to nested models has essentially the same weaknesses as does the χ^2 test applied to any single model (Schermelleh-Engle et al 2003). Thus, given the large sample size, a confidence level of 99% ($p = .01$) was chosen. In addition, we

provide alternative goodness-of-fit measures such as the root mean squared mean error of approximation (RMSEA) and Akaike's information criterion (AIC). The RMSEA is a measure of closeness of fit and provides a measure of discrepancy per degree of freedom. A value of .05 indicates a close fit, and values up to .08 represent reasonable errors of approximation in the population (Jöreskog 1993). The AIC compares models on the basis of parsimony, taking jointly into account the χ^2 and the df (Jöreskog 1993). The lower the AIC, the better the fit of the model to the data, and the more parsimonious the model is.

The saturated model was used as a reference to test for the homogeneity of means and variances. Homogeneity of means and variances was tested, constraining them to be equal across birth order, zygosity, sex, and time points (ages 7, 10, and 12 y). For these tests, we also report the standardized root mean square residual (SRMR) and the comparative fit index (CFI). The SRMR is a badness-of-fit measure that is based on the fitted standardized residuals; a value of zero indicates perfect fit, and values of less than .10 may be interpreted as acceptable (Hu and Bentler 1995). The saturated model has a SRMR of zero. Thus an increase in SRMR is entirely a result of the specific homogeneity test and can be seen as an indicator of the amount of variance that is explained by heterogeneity. The CFI is a comparison index in which the model of interest is compared with a baseline or independence model. It is one of the fit indices that is less affected by sample size and can take values from zero to one, for which 97 or higher indicates a good fit, whereas values greater than .95 may be interpreted as acceptable (Schermelleh-Engle et al 2003).

Genetic Modeling

The path diagram in Figure 1 represents the general genetic model that was tested on the longitudinal JBD data. The diagram represents the model for an opposite-sex twin pair. The first-born twin is male, and the second-born twin is female. Different parameters are estimated for male and female twins. The rectangles represent the phenotypic measures at 7, 10, and 12 years for both twins. A so-called ACE model was fitted in which the variance of the JBD phenotype was explained by additive genetic effects (A), environmental factors shared by the members of the same family (C), and environmental factors specific to the individual (E; the E component is omitted in Figure 1 for clarity). The sources of variance are represented as latent, unmeasured factors within circles. Genetic and environmental effects on stability and change are investigated through a Cholesky or triangular decomposition (Neale and Cardon 1992). Genetic (A) and environmental (C and E) sources of variance-covariance across time are represented by three latent factors, so that the first factors are the stable sources of variance present at 7, 10, and 12 years of age; the second factors represent the sources of variance common to 10 and 12 years of age that were not present at 7 years; and the third factors represent the sources of variance specific to 12 years of age. That is, additive genetic effects are represented by a triangular matrix of factor loadings, as follows:

$$\begin{pmatrix} a_{11} & & \\ a_{21} & a_{22} & \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

(factors in columns and variables in rows); multiplying this matrix by its transpose results in the genetic variance-covariance matrix, as follows:

$$\begin{pmatrix} a_{11}^2 & a_{21}a_{11} & a_{31}a_{11} \\ a_{21}a_{11} & a_{22}^2 + a_{21}^2 & a_{32}a_{22} + a_{31}a_{21} \\ a_{31}a_{11} & a_{32}a_{22} + a_{31}a_{21} & a_{33}^2 + a_{32}^2 + a_{31}^2 \end{pmatrix};$$

dividing this matrix by the implied phenotypic variance-covariance matrix provides the proportion of variances and covariances explained by additive genetic effects; and standardizing it provides the genetic correlation matrix, in which the correlations indicate the overlap of genetic effects across time.

Twins may resemble each other because they share their pre- and postnatal rearing environment, often referred to as shared or common environment (C). In addition, DZ twins may resemble each other because they share 50% of their additive genetic variance (A). MZ twins share all the additive genetic variance, because they always, or nearly always, have identical genotypes. Thus, A factors are correlated 1 across MZ twin pairs and .5 across DZ pairs. The correlation between genetic factors of OS twins can be estimated (r_{gos}), allowing for the possibility that different genes influence the phenotype in male and female twins. C factors are correlated 1 for MZ and DZ twins, and E factors are uncorrelated between pairs by definition. Estimates of the unique environmental effects (E) also include measurement error (Boomsma et al 2002a). First, parameters in the full ACE model were estimated. Next, equality constraints were imposed across the sexes to test for sex differences in variance components and were imposed across time to test for differences in variance components across age.

Results

Sample Characteristics and Descriptive Statistics

Table 1 shows the means and variances estimated in the saturated model across zygosity and sex. Table 2 shows the results of the tests for the homogeneity of means and variances. For all tests, the SRMR and CFI indicate that only the mean and variance differences across sexes can be considered relevant (SRMR $>$ .05, CFI $<$ .97). There is a tendency for male twins to show larger means and variability in CBCL-JBD than female twins.

The summary of twin correlations at each age and of the cross-twin–cross-age correlations is shown in Table 3. The twin correlations within age show that at each age, the DZ correlations appear to be somewhat larger than half the MZ correlations. This suggests that genes and shared family environment both explain familial resemblances in CBCL-JBD. The cross-twin–cross-age correlations represent JBD at one age (e.g., 7 y) in one twin, with CBCL-JBD at another age (e.g., 10 y) in the other twin (correlations constrained to be equal for first- with second-born twin and for second-born with first-born twin). As can be seen, the past behavior of the co-twin is more predictive for the current behavior of his or her twin in MZ pairs than it is in DZ pairs. In fact, for MZ twins, the cross-correlations are almost as high as the within-person correlations across time. These within-person correlations, or stability-coefficient correlations across time, were .72 from 7 to 10 years, .66 from 7 to 12 years, and .77 from 10 to 12 years. On the basis of this pattern of cross-twin–cross-age correlations for MZ and DZ twins, it may be expected that longitudinal stability in JBD is explained by genetic factors and by common environment.

Genetic Analyses

Table 4 shows the standardized parameter estimates from the full ACE model and from the reduced model without sex differences. In these models, the estimate of the genetic correla-

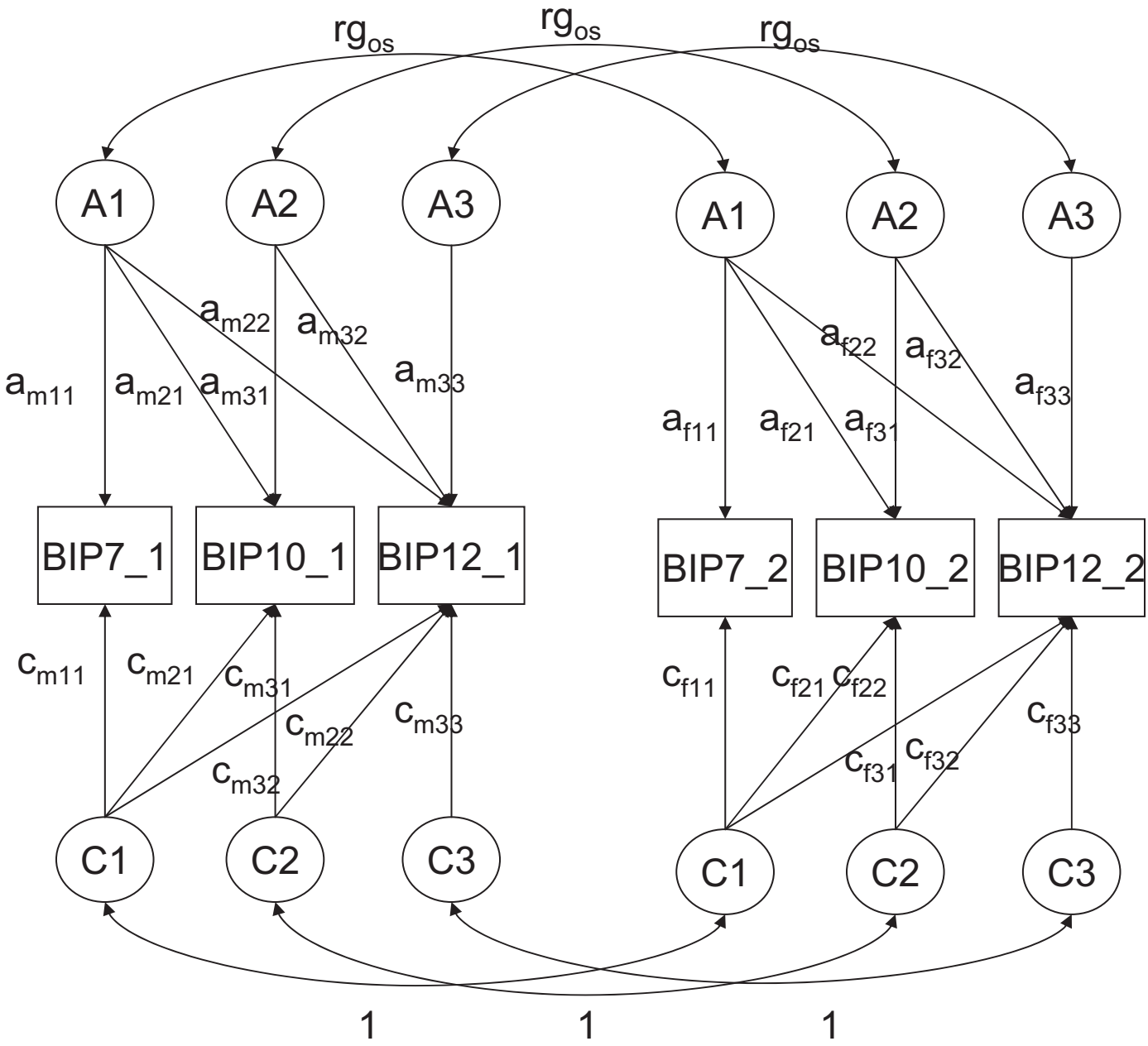

Figure 1. Analysis of longitudinal data on juvenile bipolar disorder (JBD) at ages 7, 10, and 12 years (nonshared environment is omitted in the figure for simplicity but is modeled in a similar way). The figure shows data from a pair of dizygotic opposite-sex twins; the rectangles represent the phenotypic measures at ages 7, 10, and 12 years for first-born twin (male) and for second-born twin (female). In parameter subscripts, m stands for male and f for female. Resemblance in JBD phenotype is explained by additive genetic effects (A) and by common environmental factors that are shared by the members of the same family (C). These are represented as latent, unmeasured factors within circles. Genes (A) and environment (C) across time are represented by three latent factors so that the first factors are the stable sources of variance present at ages 7, 10, and 12 years; the second set of factors represents the sources of variance at ages 10 and 12 years that were not present at age 7 years; and the third set of factors represents the sources of variance that are specific to 12 years of age. Arching arrows between latent factors represent correlations. Common environmental factors are correlated unity within pairs; genetic factors are correlated 0.5 in same-sex dizygotic twins and may be estimated in opposite-sex dizygotic twins (rg_{os}).

tion in opposite sex twins (r_{gos}) was equal to .5, indicating that the same genes are expressed in boys and girls. The fit of the model was −2LL = 187,397.335, df = 27849, $\Delta\chi^2$ = 219.37, Δdf = 89, p = .000, RMSEA = .043 (.035–.049), AIC = 41.37. According to the RMSEA, the longitudinal model provides an acceptable fit to the data. The estimates of the standardized variance components (diagonals in Table 4) suggest that additive genetic effects increase with age (e.g., heritability in male twins is 61% at age 7 y and 75% at age 12 y). The effects of the shared environment tend to decrease with age (e.g., 17% in girls at age 7 y and 7% at age 12 y).

The off-diagonal estimates in Table 4 summarize the results regarding the decomposition of the phenotypic stability across time. The proportions above the diagonal give covariance components, and the estimates below the diagonal give genetic and environmental correlations across time. Genetic and environmental covariance components sum to 100% and give the

Table 1. Means and Variances of CBCL-JBD Across Zygosity by Sex, at Ages 7, 10, and 12 Years

	Mean			Variance			*n* (Twin Pairs)		
	7 y	10 y	12 y	7 y	10 y	12 y	7 y	10 y	12 y
MZM	12.937	12.897	11.170	91.427	109.237	96.606	1,215	746	424
DZM	12.972	12.819	11.509	99.481	118.342	111.695	1,230	681	379
MZF	10.274	9.635	8.698	76.263	78.486	66.492	1,396	910	494
DZF	10.831	10.284	8.752	81.540	91.568	74.478	1,157	657	353
DOS-M	12.308	11.880	10.667	97.666	108.259	103.791	1,186	732	369
DOS-F	9.244	8.794	7.987	68.793	73.434	75.356	1,103	674	330

MZM, monozygotic males; DZM, dizygotic males; MZF, monozygotic females; DZF, dizygotic females; DOS, dizygotic opposite-sex pairs.

proportion of the total covariance across time that is explained by genetic and environmental stable influences. For CBCL-JBD, results suggest that roughly 80% of the stability in childhood is a result of additive genetic effects, and about 10% of stability from 10 to 12 years of age is explained by shared environmental influences. Genetic and environmental correlations can be interpreted as indicators of the extent to which the same genes and environmental factors influence the trait at different ages. As may be seen, genetic correlations are high (.7 or above), whereas environmental correlations are lower.

Table 5 summarizes the model-fitting tests. Model 1 is the ACE model, with sex differences in parameter estimates. In model 2, the factor loadings of the A, C, and E latent factors are constrained to be equal for male and female twins. Model two fits the data significantly worse than does model 1. When the absolute amount of variance explained by each component was constrained independently, only the C component could be constrained to be equal for male and female twins ($\Delta\chi^2 = 11.239$, $\Delta df = 6$, $p = .081$), whereas the amount of variance explained by A and E differed significantly across sexes ($p < .01$).

The results concerning sex differences may appear surprising given that the estimates of the proportion of variance explained by A, C, and E in the full ACE model (Table 4) do not look so different. The explanation might rest in the constraint that the absolute factor loadings are equal across sexes, for example, as in A: $a_{m11} = a_{f11}$, $a_{m21} = a_{f21}$, $a_{m31} = a_{f31}$, $a_{m22} = a_{f22}$, $a_{m32} = a_{f32}$, and $a_{m33} = a_{f33}$. However, the proportion of variance that is explained by each component also depends on the total variance. Thus, although the proportion of variance explained by A, C, and E was equal for male and female twins, the absolute amounts of variance explained was larger for male twins who have larger total variances. To allow for this possibility, nonlinear constraints were used to test whether the relative proportion of variance accounted for by A, C, and E was equal across sexes, for example, as in A:

$$\frac{a_{m11}^2}{a_{m11}^2 + c_{m11}^2 + e_{m11}^2} = \frac{a_{f11}^2}{a_{f11}^2 + c_{f11}^2 + e_{f11}^2}$$

$$\frac{\left(a_{m21}^2 + a_{m22}^2\right)}{\left(a_{m21}^2 + a_{m22}^2\right) + \left(c_{m21}^2 + c_{m22}^2\right) + \left(e_{m21}^2 + e_{m22}^2\right)} = \frac{\left(a_{f21}^2 + a_{f22}^2\right)}{\left(a_{f21}^2 + a_{f22}^2\right) + \left(c_{f21}^2 + c_{f22}^2\right) + \left(e_{f21}^2 + e_{f22}^2\right)}$$

$$\frac{\left(a_{m31}^2 + a_{m32}^2 + a_{m33}^2\right)}{\left(a_{m31}^2 + a_{m32}^2 + a_{m33}^2\right) + \left(c_{m31}^2 + c_{m32}^2 + c_{m33}^2\right) + \left(e_{m31}^2 + e_{m32}^2 + e_{m33}^2\right)} = \frac{\left(a_{f31}^2 + a_{f32}^2 + a_{f33}^2\right)}{\left(a_{f31}^2 + a_{f32}^2 + a_{f33}^2\right) + \left(c_{f31}^2 + c_{f32}^2 + c_{f33}^2\right) + \left(e_{f31}^2 + e_{f32}^2 + e_{f33}^2\right)}$$

In model 3, only the cross-sectional variance components are constrained across sexes as shown in the previous equations. Then, in model 4, the same constraint is extended to the cross-time A, C, and E covariance components. Model 4, in which both relative variance and covariance components are constrained to be equal across sexes, fits the data as well as the full ACE model, and when compared with the saturated model, it presents an RMSEA of .013 (.011–.015), indicative of an excellent fit. Thus, it can be concluded that the same relative amount of

Table 2. Tests for Homogeneity of Means and Variances

	$\Delta\chi^2$	Δdf	SRMR	CFI
Homogeneity of means				
Across birth order	89.95[a]	18	.021	.996
Across zygosity	56.14[a]	24	.026	.993
Across sex	411.98[a]	18	.074	.976
Age 7 = age 10	29.83[a]	12	.010	.999
Age 10 = age 12	128.61[a]	12	.027	.993
Age 7 = age 10 = age 12	201.83[a]	24	.038	.989
Homogeneity of variances				
Across birth order	78.678[a]	18	.043	.996
Across zygosity	72.90[a]	24	.050	.997
Across sex	236.99[a]	18	.102	.910
Across age 7 = 10 = 12	102.186[a]	24	.047	.995

$\Delta\chi^2$, Change in chi-squared statistic and degrees of freedom (df) compared with to a fully saturated model; SRMR, standardized root mean squared residual index; CFI, comparative fit index.

[a] $P < .01$.

Table 3. Twin Correlations at Ages 7, 10, and 12 Years and Cross-Twin-Cross-Time Correlations for JBD

	Cross Twin–Within Time			Cross Twin–Cross Time		
	7	10	12	7–10	7–12	10–12
MZM	.84	.84	.81	.66	.60	.67
DZM	.55	.48	.44	.39	.33	.34
MZF	.82	.81	.83	.64	.59	.68
DZF	.52	.48	.47	.39	.39	.39
DOS	.50	.47	.47	.37	.32	.36

MZM, monozygotic males; DZM, dizygotic males; MZF, monozygotic females; DZF, dizygotic females; DOS, dizygotic opposite sex pairs.

Confidence intervals for correlations are given in Appendix 2.

Table 4. Standardized Parameter Estimates from the ACE Model with/without Sex Differences

	Parameter Estimates from Full ACE Model						Parameter Estimates from Reduced Model Without Sex Differences		
	Males			Females					
	7	10	12	7	10	12	7	10	12
Additive genetic architecture (heritability on diagonal, genetic covariance components above and genetic correlations below diagonal)									
7	.61	.76	.82	.65	.75	.79	.63	.75	.78
10	.83	.73	.84	.70	.67	.83	.81	.71	.84
12	.80	.90	.75	.75	.86	.75	.75	.89	.75
Shared environment architecture (% of variance explained by shared environment on diagonal, shared environmental covariance components above and correlations below diagonal)									
7	.24	.16	.10	.17	.13	.11	.20	.15	.11
10	.71	.11	.03	.69	.11	.05	.70	.11	.04
12	.47	.29	.08	.65	.43	.07	.62	.36	.08
Unique environment architecture (% of variance explained by unique environment on diagonal, unique environmental covariance components above and correlations below diagonal)									
7	.15	.08	.08	.18	.12	.10	.17	.10	.10
10	.37	.15	.12	.46	.20	.12	.42	.18	.12
12	.35	.57	.17	.39	.48	.18	.38	.54	.18

Confidence intervals are shown in Appendix 3.

variance within time and covariance across time is explained by A, C, and E for male and female twins.

Table 4 shows the parameter estimates of model 4, in which standardized estimates are the same for girls and boys. Between ages 7 and 12 years, the heritability of JBD increases from 63% to 75%, and the contribution of shared environment decreases from 20% to 8%. The remaining 18% of the variance is explained by unique environment. Covariance components show that the largest part of the stability (between 75% and 84%) between 7 and 12 years is a result of additive genetic effects. Only 4% and 15% is a result of shared environment, and around 10% is a result of stable unique environmental influences.

Longitudinal Trends

Finally, models 5, 6, and 7 tested differences in variance explained by A, C, and E across time. Nonlinear constraints were used to test whether the variance explained by A, C, and E was proportional at ages 7, 10, and 12 years; that is, for A,

$$\frac{a_{11}^2}{a_{11}^2 + c_{11}^2 + e_{11}^2} = \frac{\left(a_{21}^2 + a_{22}^2\right)}{\left(a_{21}^2 + a_{22}^2\right) + \left(c_{21}^2 + c_{22}^2\right) + \left(e_{21}^2 + e_{22}^2\right)} = \frac{\left(a_{31}^2 + a_{32}^2 + a_{33}^2\right)}{\left(a_{31}^2 + a_{32}^2 + a_{33}^2\right) + \left(c_{31}^2 + c_{32}^2 + c_{33}^2\right) + \left(e_{31}^2 + e_{32}^2 + e_{33}^2\right)}$$

Models 5 and 6 fit significantly worse than model 4, whereas model 7 fit as well as model 4. According to these results, additive genetic effects increase significantly with age, whereas the effects of the shared environment decrease. The proportion of variance explained by the unique environment remains the same.

Finally, given that the amount of variance explained by C decreases to values close to zero at age 12 years, three additional models were fitted in which the C component was constrained to be zero at ages 12 years ($c_{33} = c_{32} = c_{31} = 0$), 10 years ($c_{22} =$

Table 5. Model Fitting Results: Tests for Sex Differences and Longitudinal Trends Based on the ACE Model (Additive Genetic, Common Environmental and Unique Environmental Influences)

Model	−2LL	df	C.T.[a]	χ^2	Δdf	*p*
Tests for sex differences in absolute estimates of variance components						
1 Full ACE with rg OS = .5	187,397.335	27,850				
2 No sex differences in variance components	187,727.390	27,868	1	330.05	18	.000
Tests for sex differences in standardized variance components and in longitudinal covariance components						
3 Proportion of Variance explained by ACE equal for males and females	187,418.059	27,859	1	20.724	9	.014
4 Proportion covariance explained by ACE equal for males and females	187,429.772	27,868	3	11.713	9	.229
Tests of longitudinal changes in the proportion of variance explained by A, C and E						
5 Proportion of variance explained by A equal at 7, 10 and 12 years	187,441.126	27,870	4	11.354	2	.003
6 Proportion of variance explained by C equal at 7, 10 and 12 years	187,445.505	27,870	4	15.733	2	.000
7 Proportion of variance explained by E equal at 7, 10 and 12 years	187,431.511	27,870	4	1.739	2	.419

[a]C.T. Compared to model number #.

$c_{21} = 0$), and 7 ($c_{11} = 0$) years. The results showed that the effects of C are significant at ages 7 and 10 years ($p < .01$) but that they are negligible at age 12 years ($p = .119$).

Discussion

This study examined the stability and genetic architecture across time of the CBCL-JBD phenotype, which has been shown to be consistent with DSM conceptualizations of JBD. CBCL-JBD, defined as the sum of the AP, AGG, and A/D subscales, has been demonstrated to be associated with the DSM JBD phenotype across studies. The use of CBCL-JBD as a measure of DSM JBD has been recommended as one possible method to circumvent the diagnostic confounds that continue to be debated (National Institute of Mental Health 2001). The CBCL-JBD construct has the advantage that it is based on empirically derived dimensions of childhood psychopathology, whose summation leads to a continuous scale. The use of continuous scales, as compared with categorical or dichotomous data, leads to an increase in statistical power in genetic studies (e.g., Derks et al 2004; Neale et al 1994). By using the summed score, children with DSM-III-R BD were identified accurately in a large family study population (Faraone et al, in press).

We have found that the CBCL-JBD measure is stable across ages, and we have quantified the genetic and environmental contributions to the variation of CBCL-JBD and to its stability from ages 7 to 12 years. The influence of additive genetic effects on variation in JBD was found to be relatively high at each age, increasing from 63% at age 7 years to 75% at age 12 years. Simultaneously, the effects of the shared environment tend to decrease. At age 7 years, 20% of the variation in CBCL-JBD is explained by the influence of the common family environment, and this percentage decreases to 8% at age 12 years. The small remaining part of the variance at each age was explained by unique or individual-specific environmental influences. The estimates of common and unique environmental variances may be somewhat biased. The CBCL-JBD scale shows a skewed distribution. When an ACE model is fitted to such data, an unbiased estimate of the additive genetic effect is obtained (Derks et al 2004). However, the common environmental effect may be underestimated at the cost of the unique environmental effect.

The standardized estimates of genetic and environmental parameters were found to be the same for boys and girls, but not across time. The heritability estimates were 63%, 71%, and 75% at ages 7, 10, and 12 years, respectively. Our analyses also suggest that similar genes may underlie CBCL-JBD for both girls and boys. The estimates for the percentage of variance explained by common family environment were 20%, 11%, and 8%. Overall, the high heritability estimates obtained for this sample are in line with those obtained in adults with BD (see Smoller and Finn 2003 for a review).

Another major finding from this study concerned the stability of the CBCL-JBD phenotype across development. Correlations across age groups were .72 from 7 to 10 years, .66 from 7 to 12 years, and .77 from 10 to 12 years. Genetic covariance analysis suggests that roughly 80% of the stability on JBD in childhood is a result of additive genetic effects and that about 10% of stability is a result of shared environmental effects. Should this finding hold into adulthood (and, as important, should the phenotypic association between adult-onset BD and JBD be delineated further), it would suggest that many of the candidate chromosomal locations for adult BD genes, including 6q16–22 and 12q23–24 among several others (Boomsma et al 2006; Craddock et al 2005; Dick et al 2003), also may be important in CBCL-JBD. However, regardless of the association between JBD and adult-onset BD, research into the genetic influences on differences among children in JBD has merit in its own right. The association between JBD and adult-onset BD has yet to be clearly established and is considered an important research topic by the leaders in this field (e.g., National Institute of Mental Health 2001). It is possible that early-onset forms of BD have fundamentally different and developmentally important genetic and gene by environment effects than the adult-onset form. These effects cannot be studied without large, longitudinal samples. This is an important topic for further research and is a future aim of our work.

Interestingly, the influences of common environment appeared to decrease over time, particularly between ages 7 and 10 years. Putting together this result with the overall decrease in influence of shared environment from 7 to 12 years suggests the possibility of an important environmental factor during the formative years of CBCL-JBD onset that may not be present later on. Alternatively, this effect could include rater bias (e.g., stereotyping or having certain response styles). Rater bias in this sense will be a continuous process influencing the ratings at all ages and could mimic stability in the trait. Maternal psychopathology is one example that could affect ratings of problem behavior in their children. Because rater bias affects MZ and DZ twin correlations in the same way, it will appear as shared environmental effects. Also, assortative mating in parents could appear as a shared environmental effect. However, for both phenomena, we probably would not expect that their effects diminish between 7 and 12 years.

These findings for CBCL-JBD are in contrast with those for the separate subscales. Modeling of the AP phenotype across ages 3–12 showed additive and dominance genetic effects, along with unique environmental effects and no common environmental effects (Rietveld et al 2003a, 2004). This was replicated in a study comparing the CBCL-JBD profile with the CBCL-AP profile (Hudziak et al, in press). It appears therefore, that the CBCL-JBD is different in terms of its heritability with AP and is unlikely to be an extreme form of that phenotype.

Geller and colleagues (2001) have shown that there is a high degree of overlap between childhood-onset major depression and adult diagnosis of BD. Could CBCL-JBD be an expression of A/D? When we look at the previous modeling of A/D, this explanation appears unlikely. In our work on A/D, we found that although additive genetic, common, and unique environmental factors are important (similar to CBCL-JBD), the heritability of A/D decreases with increasing age (from ages 3 to 12 y), with the common environmental component increasing—exactly the opposite pattern that is seen with CBCL-JBD (Boomsma et al, 2006). Thus, although the shared family environment becomes more important to the expression of A/D as the child ages, it becomes less important to the expression of CBCL-JBD. We have demonstrated similar increases in the contribution of the shared environmental factor in AGG (Van Beijsterveldt et al 2003) but only in female twins. Male twins showed a relatively consistent contribution of shared environmental contribution for AGG across childhood. Overall, these findings suggest that the CBCL-JBD construct is something different than its component parts.

In summary, this study provides evidence from a large sample that many of the symptoms comprising JBD are stable across time and are strongly influenced by additive genetic factors that tend to increase with time in contrast to shared environmental factors which tend to decrease. Moreover, this stability of the CBCL-JBD phenotype also is due in large part to additive genetic influences.

It is important to note that we observed no sex differences in genetic architecture or in the stability of the CBCL-JBD phenotype, indicating that for gene-finding studies, data may be pooled across boys and girls.

This work was supported by National Institute of Mental Health Grant No. MH58799 and Nederlandse Organisatie voor Wetenschappelijk Onderzoek Grant Nos. 575-25-006, 575-25-012, and 904-57-94.

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

Althoff RR, Faraone SV, Rettew DC, Morley CP, Hudziak JJ (2005): Family, twin, adoption and molecular genetic studies of juvenile bipolar disorder. *Bipolar Disord* 7:598–609.

Althoff RR, Rettew DC, Faraone SV, Boomsma DI, Hudziak JJ (2006): Latent class analysis shows strong heritability of the CBCL-juvenile bipolar phenotype. *Biol Psychiatry* 60:903–911.

Biederman J, Klein RG, Pine DS, Klein DF (1998): Resolved: Mania is mistaken for ADHD in prepubertal children. *J Am Acad Child Adolesc Psychiatry* 37:1091–1096.

Biederman J, Wozniak J, Kiely K, Ablon S, Faraone S, Mick E, et al (1995): CBCL clinical scales discriminate prepubertal children with structured interview-derived diagnosis of mania from those with ADHD. *J Am Acad Child Adolesc Psychiatry* 34:464–471.

Boomsma DI, van Beijsterveldt CEM, Hudziak JJ (2005): Genetic and environmental influences on anxious/depressed behavior during childhood: A large twin study. *Genes, Brains and Behav* 4:466–481.

Boomsma D, Busjahn A, Peltonen L (2002a): Classical twin studies and beyond. *Nat Rev Genet* 3:872–882.

Boomsma DI, Cacioppo JT, Slagboom PE, Posthuma D (2006): Genetic linkage and association analysis for loneliness in Dutch twin and sibling pairs points to a region on chromosome 12q23–24. *Behav Genet* 36:137–146.

Boomsma DI, Vink JM, Van Beijsterveldt TC, de Geus EJC, Beem L, Mulder EJ, et al (2002b): Netherlands twin register: A focus on longitudinal research. *Twin Res* 5:401–406.

Carlson GA, Bromet EJ, Sievers S (2000): Phenomenology and outcome of subjects with early- and adult-onset psychotic mania. *Am J Psychiatry* 157:213–219.

Carlson GA, Kelly KL (1998): Manic symptoms in psychiatrically hospitalized children—What do they mean? *J Affect Disord* 51:123–135.

Craddock N, O'Donovan MC, Owen MJ (2005): The genetics of schizophrenia and bipolar disorder: Dissecting psychosis. *J Med Genet* 42:193–204.

Derks EM, Dolan CV, Boomsma DI (2004): Effects of censoring on parameter estimates and power in genetic modeling. *Twin Res* 7:659–669.

Dick DM, Foroud T, Flury L, Bowman ES, Miller MJ, Rau NL, et al (2003): Genomewide linkage analyses of bipolar disorder: A new sample of 250 pedigrees from the National Institute of Mental Health Genetics Initiative. *Am J Hum Genet* 73:107–114.

Dienes KA, Chang KD, Blasey CM, Adleman NE, Steiner H (2002): Characterization of children of bipolar parents by parent report CBCL. *J Psychiatr Res* 36:337–345.

Faedda GL, Baldessarini RJ, Glovinsky IP, Austin NB (2004): Pediatric bipolar disorder: Phenomenology and course of illness. *Bipolar Disord* 6:305–313.

Faedda GL, Baldessarini RJ, Suppes T, Tondo L, Becker I, Lipschitz DS (1995): Pediatric-onset bipolar disorder: A neglected clinical and public health problem. *Harv Rev Psychiatry* 3:171–195.

Faraone SV, Althoff RR, Hudziak JJ, Monuteaux M, Biederman J (2005): The CBCL predicts DSM bipolar disorder in children: A receiver operating characteristic curve analysis. *Bipolar Disord* 7:518–524.

Geller B, Luby J (1997): Child and adolescent bipolar disorder: A review of the past 10 years. *J Am Acad Child Adolesc Psychiatry* 36:1168–1176.

Geller B, Warner K, Williams M, Zimerman B (1998): Prepubertal and young adolescent bipolarity versus ADHD assessment and validity using the WASH-U-KSADS, CBCL and TRF. *J Affect Disord* 51:93–100.

Geller B, Zimerman B, Williams M, Bolhofner K, Craney JL (2001): Bipolar disorder at prospective follow-up of adults who had prepubertal major depressive disorder. *Am J Psychiatry* 158:125–127.

Hazell PL, Lewin TJ, Carr VJ (1999): Confirmation that Child Behavior Checklist clinical scales discriminate juvenile mania from attention deficit hyperactivity disorder. *J Paediatr Child Health* 35:199–203.

Hu L, Bentler P (1995): Evaluating model fit. In: Hoyle RH, editor. *Structural Equation Modeling. Concepts, Issues and Applications*. London: Sage, 76–99.

Hudziak JJ, Althoff RR, Derks EM, Faraone SV, Boomsma DI (2005): Prevalence and genetic architecture of Child Behavior Checklist-juvenile bipolar disorder. *Biol Psychiatry* 58:562–568.

Hudziak JJ, Rudiger LP, Neale MC, Heath AC, Todd RD (2000): A twin study of inattentive, aggressive, and anxious/depressed behaviors. *J Am Acad Child Adolesc Psychiatry* 39:469–476.

Hudziak JJ, van Beijsterveldt CE, Bartels M, Rietveld MJ, Rettew DC, Derks EM, Boomsma DI (2003): Individual differences in aggression: Genetic analyses by age, gender, and informant in 3-, 7-, and 10-year-old Dutch twins. *Behav Genet* 33:575–589.

Jöreskog K (1993): Testing structural equation models. In: Bollen KA, Long SJ, editors. *Testing Structural Equation Models*. Newbury Park, CA: Sage, 294–317.

Kahana SY, Youngstrom EA, Findling RL, Calabrese JR (2003): Employing parent, teacher, and youth self-report checklists in identifying pediatric bipolar spectrum disorders: an examination of diagnostic accuracy and clinical utility. *J Child Adolesc Psychopharmacol* 13:471–488.

Leibenluft E, Charney DS, Towbin KE, Bhangoo RK, Pine DS (2003): Defining clinical phenotypes of juvenile mania. *Am J Psychiatry* 160:430–437.

Loehlin J (2004): *Latent Variable Models, 4th ed.* Mahwah, NJ: Erlbaum. Mannuzza S, Klein RG, Moulton JL 3rd (2003): Persistence of attention-deficit/hyperactivity disorder into adulthood: What have we learned from the prospective follow-up studies? *J Atten Disord* 7:93–100.

Muthen LK, Muthen BO (1998): *Mplus: The Comprehensive Modeling Program for Applied Researchers, version 3.13*. Los Angeles, CA: Muthén & Muthén.

National Institute of Mental Health (2001): National Institute of Mental Health research roundtable on prepubertal bipolar disorder. *J Am Acad Child Adolesc Psychiatry* 40:871–878.

Neale M, Boker SM, Xie G, Maes H (2003): *Mx: Statistical Modeling*, 6th ed. Virginia Institute for Psychiatric and Behavioral Genetics, VCU Box 900126, Richmond, VA 23298: Department of Psychiatry.

Neale M, Cardon L (1992): *Methodology for Genetic Studies of Twins and Families* North Atlantic Treaty Organization, Scientific Affairs Division. Dordrecht, The Netherlands: Kluwer Academic.

Neale MC, Eaves LJ, Kendler KS (1994): The power of the classical twin study to resolve variation in threshold traits. *Behav Genet* 24:239–258.

Rietveld MJ, Hudziak JJ, Bartels M, van Beijsterveldt CE, Boomsma DI (2003a): Heritability of attention problems in children: I. Cross-sectional results from a study of twins, age 3−12 years. *Am J Med Genet B Neuropsychiatr Genet* 117B:102–113.

Rietveld MJH, Hudziak JJ, Bartels M, van Beijsterveldt CEM, Boomsma DI (2004): Heritability of attention problems in children: Longitudinal results from a study of twins, age 3 to 12. *J Child Psychol Psychiatry* 45:577–588.

Rietveld MJ, Posthuma D, Dolan CV, Boomsma DL (2003b): ADHD: Sibling interaction or dominance: An evaluation of statistical power. *Behav Genet* 33:247–255.

Rietveld M, van der Valk JC, Bongers IL, Stroet T, Slagboom PE, Boomsma DI (2000): Zygosity diagnosis in young twins by parental report. *Twin Res* 3:134–141.

Schermelleh-Engle K, Moosbrugger H, Muller H (2003): Evaluating the fit of structural equation models: Tests of significance and descriptive goodness of fit measures. *Methods Psychol Res Online* 8:23–74.

Smoller JW, Finn CT (2003): Family, twin, and adoption studies of bipolar disorder. *Am J Med Genet* 123C:48–58.

Stanger C, Achenbach TM, Verhulst FC (1997): Accelerated longitudinal comparisons of aggressive versus delinquent syndromes. *Dev Psychopathol* 9:43–58.

Van Beijsterveldt CEM, Bartels M, Hudziak JJ, Boomsma DI (2003): Causes of stability of aggression from early childhood to adolescence: A longitudinal genetic analysis in Dutch twins. *Behav Genet* 33:591–605.

Van den Oord EJ, Verhulst FC, Boomsma DI (1996): A genetic study of maternal and paternal ratings of problem behaviors in 3-year-old twins. *J Abnorm Psychol* 105:349–357.

Wals M, Hillegers MH, Reichart CG, Ormel J, Nolen WA, Verhulst FC (2001): Prevalence of psychopathology in children of a bipolar parent. *J Am Acad Child Adolesc Psychiatry* 40:1094–1102.

Appendix 1. Participation Rates by Age

		12 y			
7 y	10 y	No Response	Response	No Questionnaire Sent	Total
No response	No response	0	104	2	106
	Response	230	260	37	527
	No questionnaire sent	0	0	1	1
	Total	230	364	40	634
Response	No response	542	213	486	1,239
	Response	719	2,866	1,116	4,701
	No questionnaire sent	0	0	1,408	1,407
	Total	1,261	3,079	3,007	7,347
No questionnaire sent	Response	1	14	2	17
	No questionnaire sent	0	15	0	15
	Total	1	29	2	32

Weller EB, Weller RA, Fristad MA (1995): Bipolar disorder in children: Misdiagnosis, underdiagnosis, and future directions. *J Am Acad Child Adolesc Psychiatry* 34:709–714.

Appendix 2. 99% Confidence Intervals for Twin Correlations

	Cross Twin–Within Time (by Ages in y)			Cross Twin–Cross Time (by Ages in y)		
	7	10	12	7–10	7–12	10–12
MZM	.816–.858	.807–.861	.764–.844	.637–.684	.567–.636	.632–.696
DZM	.498–.600	.408–.546	.337–.525	.334–.392	.267–.392	.263–.401
MZF	.800–.844	.776–.831	.791–.858	.611–.659	.560–.625	.643–.704
DZF	.469–.577	.402–.548	.369–.570	.337–.444	.317–.448	.306–.459
OS	.464–.541	.407–.512	.389–.533	.325–.406	.268–.370	.298–.407

MZM, monozygotic males; DZM, dizygotic males; MZF, monozygotic females; DZF, dizygotic females; OS, opposite sex.

Appendix 3. 99% Confidence Intervals for Model 4 (Reduced Model)

	Parameter Estimates from Model without Sex Differences		
Age in y	7 y	10 y	12 y
Additive genetic architecture (heritability on diagonal, with genetic covariance components above and genetic correlations below diagonal)			
7	**.576–.687**	.678–.754	.786–.893
10	.764–.857	**.641–.788**	.550–.835
12	.683–.826	.845–.944	**.648–.824**
Shared environment architecture (% of variance explained by shared environment on diagonal, with shared environmental covariance components above and correlations below diagonal)			
7	**.202–.253**	.071–.216	.000–.209
10	.487–.705	**.109–.177**	.042–.135
12	.186–.997	−.646–.737	**.003–.168**
Unique environment architecture (% of variance explained by unique environment on diagonal, with unique environmental covariance components above and correlations below diagonal)			
7	**.167–.181**	.089–.118	.075–.123
10	.364–.476	**.084–.118**	.155–.202
12	.292–.456	.541–.606	**.156–.202**