Behavior Genetics, Vol. 34, No. 3, May 2004 (© 2004)

The Genetic Basis of Problem Behavior in 5-Year-Old Dutch Twin Pairs

C. E. M. van Beijsterveldt,[1,4] F. C. Verhulst,[2] P. C. M. Molenaar,[3] and D. I. Boomsma[1]

Received 13 June 2002—Final 20 Oct. 2003

Different instruments can be used in the assessment of psychopathology in young children. In the present study the psychometric properties of a subset of items of the Devereux Child Behavior (DCB) rating scale were evaluated and the contribution of genetic and environmental influences to the variance of derived problem behavior scales was estimated. Maternal and paternal ratings were obtained in about 7600 5-year-old Dutch twin pairs. Six problem scales were derived from exploratory and confirmatory factor analysis and designated as emotional lability, aggressive behavior, attention problems, anxiety problems, physical coordination problems, and dependency. Univariate genetic analyses of the problem scales yielded large additive genetic effects. Heritability estimates ranged from 40% for aggressive behavior to 81% for attention problems. Shared environmental influences were found for aggressive behavior, anxiety problems, dependency, and emotional lability. Rater contrast and/or sibling interaction effects were found for attention problems and physical coordination.

KEY WORDS: Behavior genetics; children; twins; problem behavior.

INTRODUCTION

In clinical and empirical research of child psychopathology the operationalization of childhood behavioral and emotional problems is complex. The lack of independent validating criteria and vague boundaries between different taxonomic constructs complicate the definition of the phenotype (Rutter *et al.,* 1999). In the last decade considerable effort has been devoted to the development of standardized assessment procedures of child psychopathology (Achenbach, 1995; Verhulst and Van der Ende, 2001). The two most widely used approaches are: (i) the clinical diagnostic approach exemplified by DSM (American Psychiatric Association, 1996), in which panels of experts negotiate on the existence of diagnosis and on the criteria defining them, and (ii) the empirical approach, in which multivariate statistical techniques are used to derive syndromes from scores based on rating scales completed for large samples. The first method assumes that disorders are discrete categories defined by *a priori* criteria for the number and type of symptoms. The empirical approach measures problem behaviors on continuous scales by summing the scores of the items that compose the syndromes. In genetic epidemiology, both methods can be used but the empirical approach using rating scales has some advantages. Rating scales assess problem behavior in a standardized way, are easy to administer for large groups in an economic way, and measure problem behavior on a quantitative scale. In addition, the empirical approach may offer a more discriminating etiology of childhood psychopathology and retain more specific information needed for genetic studies (Hudziak, 2001).

An increasing number of studies have used rating scales to examine the importance of genetic and environmental influences on problem behavior in children (Edelbrock *et al.,* 1995; Gjone and Novik, 1995; Hudziak *et al.,* 2000; Kuntsi *et al.,* 2000; Rietveld *et al.,* 2003; Schmitz *et al.,* 1995; Van Den Oord *et al.,* 1996; Van Der Valk *et al.,* 1998, 2001). Although it is evident

[1] Vrije Universiteit, Amsterdam, The Netherlands.
[2] Erasmus Universiteit, Rotterdam, The Netherlands.
[3] Universiteit van Amsterdam, Amsterdam, The Netherlands.
[4] To whom correspondence should be addressed at Vrije Universiteit, Department of Biological Psychology, Van der Boechorststraat 1, 1081 BT Amsterdam, The Netherlands. Tel: +31 (0)20-4448949; Fax: +31 (0)20-4448832. e-mail: toos@psy.vu.nl

0001-8244/04/0500-0229/0 © 2004 Plenum Publishing Corporation

that genetic factors explain a substantial part of the variation in problem behavior in childhood, the estimates of genetic and environmental effects varied among the studies. For example, the heritability estimates for parent-rated anxiety range from 30% to 80% (Rice *et al.,* 2002). In addition to factors such as age, sex, and who rated the behavior, the use of different assessment instruments may contribute to these differences. Although most questionnaires cover a roughly similar content, the scales may vary as function of the item content, the number of items, the statistical technique that was used to compute the syndromes and the samples used to derive the data (Verhulst and Van der Ende, 2001). As a consequence, the content, number of scales, and psychometric qualities of the syndrome scales may differ across the various instruments and may bring about variation in the genetic and environmental estimates. This is illustrated in a study of Nadder *et al.* (2001), in which various assessment instruments were used to assess ADHD in the same sample. A common factor was found that underlies the expression of ADHD across the various instruments. However, about 14% of the genetic variance was due to instrument-specific genetic effects not shared by other assessment instruments.

The contribution of genetic and environmental influences to problem behavior in childhood may vary as function of age. In a cross-sectional study Gjone *et al.* (1996) used the Child Behavior Checklist (CBCL) to assess externalizing and internalizing behavior in a sample of 915 twin pairs ages 5 to 15 years. For internalizing behavior, they found that the heritability estimate decreased from about 70% at age 5 to 6 years to 30% at age 14 to 15 years. Shared environmental influences increased from 0 to 40%. Recent genetic longitudinal studies (Bartels *et al.,* submitted; Rietveld *et al.,* in press; Van der Valk *et al.,* 2003) also reported changes in the etiology of problem behavior during childhood. Using the CBCL as well, these longitudinal studies examined the genetic contribution to problem behavior from ages 3 to 12 in large samples of Dutch twin pairs (partly overlapping each other and partly overlapping with the present sample). For internalizing behavior the influence of genetic factors decreased, whereas the influence of the shared environment increased over time. Across sexes, at age 3 the estimates for genetic and shared environmental influences were 59% and 10%, respectively. Four years later, at age 7, the influence of genetic factors decreased to 31%, while the influence of the shared environment increased to 40%. No changes in heritabilities were observed for externalizing behavior, both at age 3 and age 7 the heritability estimate was about 60%. However, different sets of genes seemed to be expressed at ages 3 and 7. The age related changes in genetic architecture of internalizing and externalizing behaviors might be explained by developmental changes between ages 3 and 7 (Van der Valk *et al.,* 2003). In this period, significant changes in brain structure and functioning are reported (Pfefferbaum *et al.,* 1994; Van Baal *et al.,* 2001). In addition, it is a period in which the transition from preschool to elementary school takes place. School-age children, in comparison with preschool children, develop new cognitive and social skills and experience new environmental demands. These new environmental stressors, together with the accompanying different interactions between these new environmental stressors and the biological make-up of the child, may have an influence on the etiology of children's problem behavior. Van der Valk *et al.* (2003) suggested that the behavior of the preschool child is predominantly influenced by the child's genotype, and thus a relatively high genetic estimate will be found. Older children, with a better developed understanding of others peoples' values and ideas, will be more receptive for parental advice and values. As a consequence, the influence of shared environmental factors will be more important in school-age children.

The period between age 3 and 7 years covers a relatively long period, and little is known about the etiology of the variability in problem behavior in this period. Studying 5-year-old children may fill this gap. As far as we know, there are only two studies that investigated the problem behavior in this critical developmental period (Gjone *et al.,* 1996; Zahn-Waxler *et al.,* 1996). However, both studies yielded different results, used small sample sizes, and investigated the etiology of broad categories of problem behavior. In a sample of 199 twin pairs, Gjone *et al.* (1996) reported a heritability estimate of about 70% for both internalizing and externalizing problems, but found no effect of shared environment influences. Zahn-Waxler *et al.* (1996) examined about 200 twin pairs and found that genetic factors accounted for 50% of the variance of both externalizing and internalizing problems. Shared environment was important only for externalizing problems and explained about 30% of the variance.

In the present study behavioral/emotional problems were measured in a large sample of Dutch twin pairs with a mean age of 5 years using a subset of 42 items of the Devereux Child Behavior rating scale (DCB, Spivack and Spotts, 1966). The psychometric properties of this shortened version of the DCB are unknown, and therefore in the present study the psychometric properties of the DCB were evaluated using exploratory and confirmatory factor analysis. In the second part of the study, we determined the influence of genetic and environmental factors on the derived problem scales of the DCB.

METHODS

Sample

The data of the present study were derived from a large, ongoing longitudinal twin study that examines the genetic and environmental influences on the development of problem behavior in families with 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 in Amsterdam (Boomsma *et al.,* 2002). Since 1986 the NTR recruits families with twins a few months after birth. Currently 40% to 50% of all twin births in the Netherlands are registered by the NTR (Boomsma, 1998; Boomsma *et al.,* 2002). For the present study, we included data of 5-year-old twin pairs from birth cohorts 1989 to 1997. Questionnaires were mailed to families within 3 months of the twins' fifth birthday. After 2 or 3 months, reminders were sent, and 4 months after the initial mailing, persistent nonresponders were contacted by phone. Families whose address was not available were included in the nonresponse group. A response rate of 66% was obtained ($n = 8219$ families). The mean age of the total sample was 5.3 years (SD = 0.22). In the current analyses, twin pairs were excluded if one or both twins had a mental handicap ($n = 164$). This left 8041 twin pairs for which the mother filled in the questionnaire. The questionnaire included a section with the same 42 items on behavioral problems that had to be completed by the father. In the surveys that were used for the present study, 7202 of the fathers completed this section.

Determination of Zygosity

For 351 same-sex twin pairs, zygosity was based on blood group polymorphisms ($n = 29$) or DNA ($n = 317$) analyses or both ($n = 5$). For the remaining same-sex twins, zygosity was determined by questionnaire items, completed by the mother, about physical similarity and frequency of confusion of the twins by family and strangers (Goldsmith, 1991; Rietveld *et al.,* 2000). The classification of zygosity was based on a discriminant analysis, relating the questionnaire items to zygosity based on blood or DNA typing in a group of same-sex twin pairs. The discriminant function was created, based on a data from 766 twin pairs for whom both DNA or blood results and questionnaire items were available. The zygosity was correctly classified by questionnaire in 94.2% of the cases.

For 140 twin pairs, zygosity could not be determined because one or more items were missing on the questionnaire on zygosity. For the genetic analysis, the final sample included only twin pairs with a complete set of data (thus a score for all problem behavior scales). This left data for 7679/6999 (mother ratings/father ratings) twin pairs: 1220/1131 monozygotic males (MZM), 1270/1154 dizygotic males (DZM), 1445/1311 monozygotic females (MZF), 1188/1076 dizygotic females, and 1307/1183 dizygotic males-females (DOSm_f), and 1249/1144 dizygotic females-males (DOSf_m).

Rating Scale

The study started in 1994, and at that time the rationale to use the DCB questionnaire was the need for a questionnaire that assessed problem behavior in a young children that might be related to cognitive development. The DCB is a rating scale for young children to be completed by parents or parent surrogates (Spivack and Spotts, 1966). Parents are asked to rate the behavior of their child in the last 2 months. Items are scored on a 5-point scale, with 1 indicating never and 5 indicating very frequently. The original DCB consists of 121 items, from which 42 items were used in the present study. Items were chosen whose scale was associated with intelligence (Spivack and Spotts, 1966). In addition, the questionnaire included items from scales that are related to emotional problem behavior. According to the manual the following problem scales might be associated with intelligence: distractibility, poor self-care, social isolation, inadequate need for independence, social aggression, and unethical behavior. Other problem behavior dimensions included were: coordination body, messiness/sloppiness, inability to delay, and anxious-fearful ideation. Four items were included that did not belong to one of the behavior dimensions.

Statistical Analysis

Phenotypic Factor Analyses

To examine the factor structure of the 42 DCB items, exploratory and confirmatory factor analyses (EFA and CFA) were performed. The EFA was performed to obtain an initial description of the factor structure. All EFA's were employed in two subgroups, consisting of a sample of first-born twins rated by the mother that was randomly divided into two groups. The first series of EFAs were conducted using the program SPSS, and were done to determine the number of factors with an eigenvalue larger than 1. The next series of EFA's were performed using unweighted least squares and promax rotation and were applied to polychoric correlations using the program LISREL V8.3 (Jöreskog and Sörbom, 1999). Promax rotations were used because problem dimensions in child psychopathology tend to show positive intercorrelations. The EFA was repeated with six to nine factor solutions (all factors with an eigenvalue larger than 1), and the factor loadings of the different solutions were compared across the two samples. The final factor model was chosen according to the following rules: the factor loading coefficients must be consistent across the two samples; solutions that comprised factors with only two items were excluded; and the final rule was that the solution must have meaningful factors. In the next step, the final factor model was applied to the entire sample of mother ratings to determine which items should be included in the factor model to be evaluated with CFA. Items were assigned to the factor on which they loaded the highest and the loading exceeded .30 on a given factor. Finally, confirmatory factor analysis (CFA) was used to replicate the factor model in four different samples: boys and girls rated by the mother and boys and girls rated by the father. These samples consisted of the data of the second-born twin.

The CFA was done using LISREL V8.3. Because the data were ordinal, polychoric correlations were used. With polychoric correlations as input, an accurate chi-square can only be obtained with weighted least squares estimation (Jöreskog and Sörbom, 1989, p. 193). However, the computation of the weight matrix needed in a weighted least squares (WLS) estimation procedure is not stable in models with a large number of variables and therefore unweighted least square (ULS) estimation was used.

Measurement of Fit in Phenotypic Analysis

To assess the goodness of fit we report several indices. The chi-square statistic is often used as a fit index, but less suitable when ULS estimation is used. For the ULS estimation the chi-square statistic does not follow the theoretical chi-square distribution (Bentler and Dudgeon, 1996) and is therefore less useful. However, the fit of a model may be assessed by other fit indices, such as the goodness-of-fit index (GFI), the root means square error of approximation (RMSEA), and the comparative fit index (CFI). The GFI provides a measure of the relative amount of variance and covariance accounted for by the hypothesized model. The CFI provides a fit of a particular model relative to a null model. The value of GFI and CFI can range from 0 to 1; a value of 0.90 or higher indicates an "accepted" fit, and a value of 0.95 indicated a good fit (the larger the value the better the fit). The RMSEA measures the average of the residual correlation when the predicted correlations from the model are subtracted from the observed correlations. This value should be close to 0.08 or lower (the lower the value, the better the fit).

Genetic Analyses of Problem Behavior

To estimate the influence of genetic and environmental factors on the variance of the problem behavior scales, as identified in the phenotypic factor analyses, genetic covariance structure modeling was employed (Neale and Cardon, 1992). This modeling allows the decomposition of the observed variance resulting from additive genetic factors (A), dominance genetic factors (D), shared environment factors (C), and unique, specific environmental factors (E). The relative contributions of genetic and environmental influences to individual differences in problem behavior were estimated by maximum likelihood, using the computer program Mx (Neale *et al.,* 1999). Model fitting was done on variance-covariance matrices of problem behavior in twin 1 and twin 2. Several models were fitted to the data to test the significance of the A, D, C, and E factors. The basic model fitted to the data was an ACE or an ADE model. Note that estimating D and C at the same time is not possible in a design using only MZ and DZ twins reared together. In that case the number of parameter estimates will be larger than the number of statistics and the model is underidentified. The reduced models consisted of dropping one of variance components A, C, or D or dropping A and C or A and D together. The presence of sex differences in genetic architecture was tested by a common effect sex limitation model (Neale and Cardon, 1992). In this model the relative magnitude of genetic and environmental factors can differ between sexes, but the same genes and/or shared environmental influences are expressed. Before the genetic analyses, it was tested whether the variances of MZ twins and DZ twin pairs could be constrained to be equal, and the fit of this model was compared with a model in which all variances were freely estimated. If the variances between MZ and DZ twin pairs were significantly different then reciprocal paths between the twins' phenotypes were included into the model to test for possible interaction effects. To test whether the interaction effects were the same across sexes, these reciprocal paths were constrained to be equal across boys and girls and across same-sex and opposite-sex twin pairs. To measure the fit of the reduced models, the χ^2 obtained for the different models were compared with χ^2-difference tests. The degrees of freedom (df) for this test are equal to the difference between the df for the full and the reduced model (Neale and Cardon, 1992). The best model was selected by Akaike's Information Criterion (AIC), which is a goodness-of-fit index that takes parsimony (i.e., number of parameters) into account. The AIC is calculated as the χ^2 minus twice the degrees of freedom.

RESULTS

Phenotypic Factor Analyses

Using SPSS, the results of first EFAs indicated nine factors with an eigenvalue larger than 1. Subsequently, the EFAs were repeated with six- to nine-factor solutions in two subsamples. In both samples, the results yielded only small differences in the patterns of factor loadings across the different factor solutions (six to nine factors). In the models with more than seven factors, too many factors with only two items emerged, and these solutions were discarded. This left the choice between a model with six or seven factors. The only difference between these two solutions was that in the six-factor model a general "anxiety" factor appeared, whereas in the seven-factor model this factor split up in an "anxiety" and an "emotional lability" factor. Because these two factors are also present in the original DCB, the model with seven factors was chosen as the best model. To select the items for the model to be evaluated with CFA, an EFA with seven factors was employed on the entire set of mother ratings. An item was assigned to a factor when its loading was greater than .30. When an item loaded on two factors, the item was assigned to the factor with the highest loading. On this ground, six items were excluded from further analyses.

Using CFA, the fit of the final model was evaluated in four different samples (boys and girls rated by mother and boys and girls rated by the father). In the model the factors were allowed to correlate and the variances of the latent factors were fixed to one (needed for statistical identification). To get an impression of the fit of the seven-factor model, we tested first a one-factor model. A summary of the results of the CFA is given in Table I. In each sample the fit of the seven-factor model was better than fit of the one-factor model and the fit was fairly good. Classical test theory assumes that the indicators chosen to reflect the underlying construct are unique, but in practice this assumption is usually violated. Therefore we checked whether the pattern of factor loadings was not too restrictive, and whether correlated errors of measurement or cross-factor loadings had to be specified. This was done using modification indices (MIs) provided by LISREL, which indicate possible misspecifications. To obtain the MI's maximum likelihood, estimation was used. If there was an indication for misspecification of a cross–factor loading or measurement error, the parameter was set free. This resulted in a final model with four cross-factor loadings and five measurement errors. The higher fit indices for this final model can be seen as an indication of the validity for the respecifications of the initial model. Table II displays the factor loadings obtained from the final model in the four samples. Only cross-loadings with an absolute value greater than .30 are shown. Overall, the results of the CFA confirms the fit of the model with seven factors. There were two items with cross-loadings higher than .30. The item

Table I. Goodness-of-Fit Statistics for the Phenotypic Factor Analyses of the DCB Rating Scale

		Girls			Boys		
Model	df	GFI	CFI	RMSEA	GFI	CFI	RMSEA
Mother report							
One-factor model	560	.890	.820	.090	.880	.800	.091
Seven-factor model	573	.950	.920	.063	.940	.910	.063
Final seven-factor model	564	.960	.940	.056	.950	.930	.058
Father report							
One-factor model	560	.900	.840	.086	.900	.820	.084
Initial seven-factor model	573	.950	.920	.062	.950	.920	.059
Final seven-factor model	564	.960	.940	.056	.960	.930	.053

Table II. Factor Structure of the Final Model Obtained by Means of Confirmatory Factor Analysis of the DCB for Four Groups

	Mother report girl	Mother report boy	Father report girl	Father report boy
F1: Emotional lability	$n = 3770$	$n = 3616$	$n = 3472$	$n = 3355$
Markedly impatient	.45	.43	.31 (F4: .32)	.48
Unpredictable in behavior	.74	.75	.77	.68
Changeable in mood	.67	.65	.67	.61
Refuses doing what is asked of him	.41	.47	.27	.46
Nags, demands, repeatedly ask for it	.66	.60	.60	.59
F2: Social Isolation				
Quite timid or shy	.21	.25	.28	.27
Quite rejected by peers	.92	.89	.85	.77
Quite socially isolated	.54	.55	.63	.62
F3: Aggressive behavior				
Hits, bites, scratches other children	.67	.70	.67	.67
Bossy with other children	.55	.58	.52	.57
Annoys or provokes peers into hitting	.74	.74	.71	.71
Intentionally tells lies	.58	.58	.62	.59
Teases or bullies other children	.70	.68	.69	.71
Takes things that do not belong to him	.57	.51	.58	.54
Blames others for his actions	.62	.57	.62	.58
F4: Attention problems				
Jumps from one activity to another	.64	.63	.63	.64
Does not attend to activity	.58	.57	.54	.49
Does not attend to adults	.42 (F5: .36)	.43 (F5: .33)	.36 (F5: .35)	.36 (F5: .37)
Distracted by others	.68	.73	.66	.71
Does not finish activity	.72	.75	.71	.72
F5: Dependency				
Does not want to do things for himself	.44	.41	.38	.39
Not capable of bathing self	.26	.32	.29	.38
Not capable of dressing self	.44	.46	.45	.42
Not doing something new	.39	.33	.38	.41
No meaningful communication	.70	.68	.72	.64
F6: Anxiety problems				
Concern about his physical health	.49	.48	.52	.49
Less tolerance for pain	.54	.49	.55	.51
Expresses fears	.53	.51	.53	.50
Looks unhappy	.62	.61	.62	.63
Wakes up in the night	.34	.37	.34	.27
Overexcitable	.42	.46	.35	.28
F7: Physical coordination				
Gets dirty and untidy	.50	.45	.51	.47
Awkward in gross body movements	.35	.34	.41	.32
Clumsy with fingers	.53	.52	.57	.55
A bit messy in eating habits	.57	.48	.55	.48
Careless about own belongings	.59	.57	.60	.61

Note: F4 denotes a cross-loading on factor 4 (attention problems). F5 denotes a cross-loading on factor 5 (dependency).

"markedly impatient" loaded on the factor emotional lability and attention problems. The item was finally included in the scale emotional lability because the cross-loading was larger than .30 only in one sample. The item "does not attend to adults" loaded on the factor attention problems and dependency. This item was finally included in the attention scale because the loadings were larger for this scale in three of the four samples. The factor "social isolation" had only two items with factor loadings higher than .30 in all four samples, and therefore it was decided to exclude this factor from further analysis. The six remaining factors were designated as (i) emotional lability, (ii) aggressive behavior, (iii) attention problems, (iv) anxiety problems, (v) physical coordination problems, and (vi) dependency. These six factors agreed with the original

Table III. Correlations Between DCB Problem Scales as Reported by Mother and Father (between the brackets)

	1. Agg	2. Anx	3. Att	4. Dep	5. Coo	6. Emo
1. Aggressive (Agg)		.50 (.57)	.48 (.48)	.04 (.12)	.42 (.47)	.70 (.69)
2. Anxiety (Anx)	.53 (.62)		.36 (.42)	.23 (.29)	.45 (.50)	.64 (.68)
3. Attention (Att)	.50 (.53)	.39 (.46)		.38 (.40)	.59 (.58)	.59 (.58)
4. Dependency (Dep)	.06 (.17)	.21 (.28)	.40 (.45)		.45 (.43)	.20 (.23)
5. Coordination (Coo)	.49 (.52)	.45 (.52)	.66 (.63)	.39 (.50)		.60 (.60)
6. Emotional (Emo)	.68 (.73)	.66 (.73)	.65 (.48)	.17 (.25)	.56 (.61)	
Cronbachs alpha						
Mother report, girl	0.74	0.62	0.74	0.50	0.57	0.72
Mother report, boy	0.76	0.60	0.75	0.51	0.58	0.71
Father report, girl	0.74	0.63	0.72	0.50	0.60	0.68
Father report, boy	0.76	0.57	0.72	0.52	0.57	0.68

Note: Below the diagonal the factor correlations are given for girls and above the diagonal the correlations are given for boys. In the lower part of the table the Cronbachs alpha is given for each of the problem scales in the 4 samples.

DCB dimension to a large extent. There were two exceptions: the original DCB dimensions coordination body and messiness/sloppiness grouped into one factor that we called physical coordination and the original DCB dimensions social aggression and unethical behavior grouped into one aggressive behavior factor.

As shown in Table III, the pattern of correlations among the six factors was consistent across the four groups. The correlations among the factors aggressive behavior, anxiety problems, attention problems, and emotional lability were higher than between the factors aggressive behavior and dependency. None of the correlations approximate 1; thus all factors can be considered as separate dimensions of problem behavior. To assess the internal consistency, Cronbach's alpha coefficients were calculated for the six factors. As shown in Table III, the internal consistency of the scales dependency and physical coordination problems was not very high. Both the low number of items and the heterogeneity of item content may have caused the low internal consistency. The last point especially applies to the dependency scale, which includes items such as "does not want to do things for himself," but also "not capable of dressing and bathing self." The internal consistency coefficients of the remaining scales were not very high but were acceptable. Finally, the problem scales were constructed by summing up the scores of the items. The scales included the items that correspond with the items of the factors given in Table II.

To examine whether DCB problem scales show an overlap with the problems scales of the CBCL, we calculated the correlations between DCB scales with CBCL scales obtained at ages 3 and 7 years. For about 6500 twin pairs, information was available for both DCB scales and the CBCL scales obtained at age 3. The following correlations (mother/father) were found: for aggressive behavior, .41/.37; for attention problems, .47/.44; and for emotional lability (with CBCL oppositional behavior), .49/.45. Data of parents that filled in both the DCB at age 5 and the CBCL at age 7 were available for about 3000 twin pairs. The correlations (mother/father) of the DCB scales with the CBCL scales were: for aggressive behavior, .48/.44; and for attention problems, .45/.41. The DCB scale emotional lability showed a correlation of .49/.48 with the CBCL aggressive behavior.

Descriptive Statistics of the Problem Scales

Means and standard deviations of the problem scales are shown in Table IV. Both fathers and mothers reported more aggressive behavior [$F(1,7856) = 69.58$; $p < 0.01$], more attention problems [$F(1,7856) = 28.67$; $p < 0.01$], more emotional lability [$F(1,7856) = 42.62$, $p < 0.01$], more physical coordination problems [$F(1,7856) = 518.68$, $p < 0.01$], and more dependency problems [$F(1,7850) = 128.03$] for boys than girls. Girls obtained a higher score than boys for anxiety problems [$F(1,7856) = 28.67$; $p < 0.01$]. Results of paired t tests showed significant mean differences between ratings of mother and father for all problem scales (aggressive behavior: $t = 9.34$, $df = 7208$; $p < 0.01$; attention problems: $t = 7.875$; $df = 7180$, $p < 0.01$; anxiety problems: $t = 7.096$, $df = 7209$; physical coordination problems: $t = 16.444$, $n = 7189$; $p < 0.01$; dependency: $t = 7.991$, $df = 7187$, $p < 0.01$; emotional lability: $t = 4.852$; $df = 7083$, $p < 0.01$). Although scores of the father ratings were significantly higher than scores of ratings given by the mother, these differences were very small. In Table IV the correlations between

Table IV. Estimated Means and Standard Deviation (between the brackets) of Untransformed Scores of the DCB Problem Scales as Reported by Mother (M) and Father (F) Across Sex and Zygosity

			Mother report	Father report	r(M,F)[†]
Aggressive behavior	Boys	MZ	12.55 (3.65)*	12.89 (3.80)	0.61
		DZ	12.99 (3.98)	13.19 (3.99)	0.66
		DOS	12.22 (3.79)	12.63 (3.93)	0.61
	Girls	MZ	11.70 (3.45)	12.01 (3.52)	0.60
		DZ	11.96 (3.50)	11.72 (3.45)	0.60
		DOS	11.97 (3.55)	11.34 (3.27)	0.58
Anxiety problems	Boys	MZ	10.71 (3.25)	10.79 (3.11)	0.60
		DZ	10.92 (3.37)	11.10 (3.25)	0.60
		DOS	10.72 (3.28)	11.02 (3.29)	0.59
	Girls	MZ	11.35 (3.49)	11.45 (3.38)	0.55
		DZ	11.28 (3.47)	11.56 (3.44)	0.59
		DOS	10.90 (3.30)	11.17 (3.27)	0.56
Emotional lability	Boys	MZ	11.78 (3.35)	11.85 (3.22)	0.63
		DZ	11.97 (3.52)	12.02 (3.30)	0.62
		DOS	11.87 (3.55)	12.02 (3.36)	0.62
	Girls	MZ	11.42 (3.23)	11.48 (3.19)	0.57
		DZ	11.53 (3.47)	11.73 (3.24)	0.60
		DOS	11.26 (3.31)	11.55 (3.21)	0.59
Attention problems	Boys	MZ	11.99 (3.40)*	12.09 (3.18)*	0.60
		DZ	11.88 (3.77)	11.98 (3.53)	0.67
		DOS	11.81 (3.65)	12.09 (3.47)	0.63
	Girls	MZ	11.50 (3.33)*	11.80 (3.12)*	0.58
		DZ	11.56 (3.61)	11.72 (3.45)	0.66
		DOS	11.12 (3.43)	11.34 (3.28)	0.63
Dependency	Boys	MZ	11.86 (2.97)	11.95 (2.89)*	0.55
		DZ	11.63 (3.10)	11.84 (3.10)	0.56
		DOS	11.97 (3.09)	12.19 (3.05)	0.53
	Girls	MZ	11.15 (2.94)	11.45 (2.83)	0.51
		DZ	11.23 (2.92)	11.40 (2.99)	0.50
		DOS	10.75 (2.90)	11.14 (2.90)	0.49
Physical coordination	Boys	MZ	9.77 (2.93)*	10.06 (2.92)*	0.59
		DZ	9.91 (3.18)	10.26 (3.16)	0.60
		DOS	10.05 (3.22)	10.49 (3.29)	0.63
	Girls	MZ	8.45 (2.63)*	8.95 (2.76)*	0.57
		DZ	8.71 (2.92)	9.37 (3.11)	0.62
		DOS	8.24 (2.58)	8.82 (2.76)	0.60

Note: In the last column the interparent correlations are given.
* = Significant differences ($p < 0.05$) between MZ and DZ; [†]r(M,F) = interparent correlation.

mothers and fathers ratings are given for the oldest twin. The interparent correlations ranged from to .49 to .67 and resembled the interparent correlations obtained for the CBCL in a Dutch sample (Bartels *et al.,* 2003; Van der Valk *et al.,* 2001). The asterisks in Table IV indicate whether the MZ and DZ variances were different. Across sexes and raters, for the scales attention problems and physical coordination problems the DZ variances were larger than the MZ variances.

Genetic Analysis of Problem Behavior

As shown in Table V all problem scales have higher correlations in MZ than in DZ twin pairs, indicating genetic influences for these problem behaviors. Most of the MZ and DZ twin correlations did not differ for boys versus girls. Also, most of the MZ and DZ correlations did not differ between mother and father ratings. For attention problems and physical coordination problems, DZ correlations were smaller than half the MZ correlations and may reflect nonadditive genetic influences. Therefore, for these scales, the genetic model fitting procedure started with an ADE model. For anxiety problems, aggressive behavior, dependency, and emotional lability, the genetic model fitting procedure started with an ACE model because the pattern of DZ and MZ correlations suggested possible shared environmental influences. Formal testing

Table V. Twin Correlations for the DCB Problem Scales (Transformed Data)

Twin correlations	MZM	DZM	MZF	DZF	DOS_mf	DOS_fm
	Mother/Father	Mother/Father	Mother/Father	Mother/Father	Mother/Father	Mother/Father
Number of twin pairs	1220/1131	1270/1154	1445/1311	1188/1076	1307/1183	1249/1144
1. Aggressive behavior	.71/.74	.48/.52	.71/.71	.46/.52	.50/.56	.52/.52
2. Anxiety problems	.72/.71	.44/.43	.72/.71	.42/.53	.51/.55	.45/.51
3. Attention problems	.59/.56	.03/.05	.64/.56	.00/.09	.11/.15	.12/.13
4. Dependency	.83/.82	.60/.60	.85/.81	.61/.58	.57/.57	.61/.62
5. Physical coordination problems	.72/.70	.28/.34	.68/.70	.20/.29	.26/.34	.24/.32
6. Emotional lability	.68/.69	.33/.40	.64/.71	.29/.40	.30/.45	.29/.40

Note: Dos_mf = dizygotic opposite sex (male-female); Dos_fm = dizygotic opposite sex (female-male).

confirmed the significance of these initial models. For the scales attention problems and physical coordination problems, the DZ variances were significantly larger than MZ variances and therefore an interaction parameter was included in their initial model.

Using model fitting, significance of sex differences and of A, C, and D factors were tested. In the first model the magnitude of A, D (or C), and E parameters was allowed to differ between boys and girls. In the second model the A, D (or C), and E parameters were constrained to be equal across sex. Depending on the results, the significance of the A, D (or C), and E parameters were tested in a model with or without sex differences. The model-fitting results are summarized in Table VI. Dropping A from the model led to enormous reduction of the fit for all problem scales, and therefore these fits are not reported. Because of the large sample size we have used an α-level of 0.01 in all statistical tests.

Aggressive Behavior

The model fitting results showed that the ACE model was the best model for both father and mother ratings. Dropping C, A or both C and A led to a large and significant deterioration of the fit. Constraining the A, C, and E parameters to be equal across sexes significantly worsened the fit. However, inspection of the standardized parameter estimates and their 95% confidence intervals in Table VII revealed no sex differences. Significant sex differences appeared for the unstandardized genetic and environmental variance components. Genetic factors accounted for 40% to 48% of the variance for aggressive behavior. The contribution of shared environmental influences varied between 25% and 32% for mother and father ratings.

Anxiety Problems

The best model for anxiety problems was an ACE model without sex differences. Dropping A or C resulted in a large decrease in the fit. As shown in Table VII, a large part of the variance (53%/43%) was explained by genetic factors.

Dependency

An ACE model with sex differences provided the best fit for dependency rated by the mother. For the father ratings, no indication for sex differences was obtained. Inspection of the standardized estimates for A, C, and E showed no differences between sexes. The only significant sex difference appeared for the unstandardized unique environmental component. The contribution of genetic factors was almost equal for mother and father ratings and accounted for 50% of the variance. Shared environmental factors accounted for 34% of the variance.

Emotional Lability

For this factors, slightly different results were obtained for the mother and father ratings. An AE model was the best model for the mother ratings, while an ACE model provided the best fit for the father ratings. There was no evidence for sex differences in genetic architecture in both the father and mother ratings.

Attention Problems

For attention problems a near-0 DZ correlation was obtained for both father and mother ratings. This low correlation may indicate the presence of genetic dominance, contrast effects, or both. If there is a contrast effect, variance differences between MZ and DZ twins are expected. Indeed both for males and females, larger variances were obtained for DZ twins compared to MZ twins. Because the contrast effects could be different among different groups (Eaves *et al.,* 2000), it was first tested whether the contrast effect could be constrained to be equal across boys and girls and across

Table VI. Model-Fitting Results for the DCB Problem Scales

	Mother report				Father report			
	df	χ^2	AIC	*p*	df	χ^2	AIC	*p*
Aggressive behavior								
ACE, sex differences	**12**	**30.138**	**6.138**	**<0.01**	**12**	**16.047**	**−7.953**	**0.189**
ACE, no sex differences	15	50.155	20.155	<0.01	15	45.888	15.888	<0.01
AE, sex differences	14	146.724	118.724	<0.01	14	197.459	169.459	<0.01
Anxiety problems								
ACE, sex differences	12	28.853	4.853	<0.01	12	41.674	17.674	<0.01
ACE, no sex differences	**15**	**33.962**	**3.962**	**<0.01**	**15**	**49.937**	**19.937**	**<0.01**
AE, no sex differences	16	99.967	67.967	<0.01	16	191.352	159.352	<0.01
Dependency								
ACE, sex differences	**12**	**8.604**	**−15.396**	**0.736**	12	21.147	−2.853	0.05
ACE, no sex differences	15	22.164	−7.836	0.104	**15**	**24.235**	**−5.765**	**0.06**
AE, sex differences	14	298.326	270.326	<0.01	16 no sex*	269.409	237.409	<0.01
Emotional lability								
ACE, sex differences	12	27.267	3.267	<0.01	12	13.622	−10.378	0.326
ACE, no sex differences	15	34.306	4.306	<0.01	**15**	**14.872**	**−15.128**	**0.461**
AE, no sex differences	**16**	**34.306**	**2.306**	**<0.01**	16	36.460	4.460	<0.01
Attention problems								
ADE-I, sex differences	10	18.766	−1.234	0.68	10	12.110	−7.890	0.278
ADE-I, no sex differences	13	32.114	6.114	<0.01	13	14.554	−11.446	0.336
ADE, sex differences	12	119.035	95.035	<0.01	15 no sex*	79.056	49.056	<0.01
AE-I, sex differences	**12**	**22.136**	**−1.864**	**0.036**	**14 no sex***	**14.554**	**−13.446**	**0.409**
AE, sex differences	14	408.893	380.893	<0.01	16 no sex*	256.858	224.858	<0.01
Physical Coordination								
ADE-I, sex differences	11	34.428	12.428	<0.00	11	44.179	22.179	<0.01
ADE-I, no sex differences	14	83.369	55.369	<0.00	14	57.319	29.391	<0.01
ADE, sex differences	12	46.967	22.967	<0.01	**12**	**49.229**	**25.299**	**<0.01**
AE-I, sex differences	**13**	**35.583**	**9.853**	**<0.01**	**13**	**44.231**	**18.231**	**<0.01**
AE, sex differences	14	141.197	113.917	<0.01	14	65.431	37.431	<0.01

Note: Bold typed models were the best-fitting models.
AE-I = AE model with sibling interaction.
* Model without sex differences.

same-sex and opposite-sex twin pairs. Results showed a smaller contrast effect in opposite-sex twin pairs compared to same-sex twin pairs. Subsequently, the significance of D and of the contrast effect was tested. Results showed that an AE model with a contrast effect provided the best fit. Significant sex differences were obtained for mother ratings, but not for father ratings. However, the sex differences of the standardized estimates of the AE parameters were very small. Overall, the contribution of the genetic factors was large (76%–81%).

Physical Coordination

The results showed a higher variance for DZ twins compared to MZ twins for both male and female twin pairs. Therefore an interaction parameter was included into the model. For the mother ratings, an AE-I model with sex differences was indeed the best model. For the father ratings, it was not possible to distinguish between an ADE and AE-I model, but the fit of the AE-I model was slightly better. For both father and mother ratings a model with sex differences was the best model, but the sex differences of standardized estimates were minimal.

DISCUSSION

The present study evaluated the phenotypic factor structure of a subset of items of the DCB and estimated to what extent genetic and environmental influences are involved in the individual differences of the derived problem scales in a large sample of 5-year-old twin pairs. The findings of the exploratory and confirmatory factor analyses of the DCB suggested a solution with six problem scales. These problem scales were interpreted as emotional lability, aggressive behavior,

Table VII. For each DCB Problem Scale the Estimates of the Absolute and Relative Contribution (with 95% CI between the brackets) of Genetic and Environmental Factors Across Sex

	Mother report		Father report	
	Standardized estimates	Unstandardized estimates	Standardized estimates	Unstandardized estimates
Aggressive behavior				
Boys A	48 (41–54)	0.13 (0.11–0.15)	44 (37–51)	0.13 (0.10–0.15)
C	25 (19–31)	0.07 (0.05–0.09)	31 (25–37)	0.09 (0.07–0.11)
E	27 (25–30)	0.08 (0.07–0.08)	25 (23–27)	0.07 (0.07–0.08)
Girls A	47 (40–53)	0.11 (0.09–0.13)	40 (32–47)	0.10 (0.08–0.12)
C	29 (19–31)	0.06 (0.05–0.08)	32 (26–39)	0.08 (0.06–0.10)
E	25 (26–31)	0.07 (0.06–0.07)	28 (26–30)	0.07 (0.06–0.08)
Anxiety				
A	53 (47–58)	0.13 (0.12–0.14)	43 (38–48)	0.10 (0.09–0.11)
C	19 (15–23)	0.05 (0.04–0.06)	28 (24–33)	0.07 (0.06–0.08)
E	28 (27–30)	0.07 (0.07–0.07)	29 (27–30)	0.07 (0.06–0.07)
Dependency				
Boys A	50 (43–55)	0.10 (0.09–0.11)	49 (45–53)	0.09 (0.09–0.10)
C	34 (28–39)	0.07 (0.06–0.08)	34 (30–37)	0.06 (0.06–0.07)
E	16 (27–30)	0.04 (0.03–0.04)	17 (16–19)	0.03 (0.03–0.04)
Girls A	50 (43–56)	0.10 (0.08–0.11)	*	*
C	35 (30–41)	0.07 (0.06–0.08)		
E	15 (14–16)	0.03 (0.03–0.03)		
Emotional lability				
A	66 (65–68)	0.16 (0.16–0.17)	58 (52–63)	0.13 (0.12–0.14)
C	—	—	12 (7–17)	0.03 (0.02–0.04)
E	34 (34–32)	0.08 (0.08–0.09)	30 (29–32)	0.07 (0.06–0.07)
Attention problems				
Boys A	79 (76–81)	0.24 (0.22–0.25)	76 (73–78)	0.20 (0.19–0.21)
E	21 (19–24)	0.06 (0.06–0.07)	24 (22–27)	0.06 (0.06–0.07)
Girls A	81 (79–83)	0.23 (0.22–0.24)	*	*
E	19 (17–21)	0.05 (0.05–0.06)		
Interaction				
Same sex	—	−0.20 (−0.22– −0.18)	—	−0.18 (−0.20– −0.15)
Opposite sex	—	−0.15 (−0.17– −0.13)	—	−0.13 (−0.15– −0.11)
Coordination				
Boys A	80 (78–82)	0.20 (0.19–0.21)	76 (73–79)	0.18 (0.17–0.19)
E	20 (18–22)	0.05 (0.04–0.05)	24 (21–29)	0.06 (0.05–0.06)
Girls A	77 (74–79)	0.16 (0.15–0.17)	75 (72–77)	0.16 (0.15–0.17)
E	23 (21–26)	0.05 (0.04–0.05)	25 (23–28)	0.05 (0.05–0.06)
Interaction, all	—	−0.09 (−0.10– −0.07)	—	−0.04 (−0.06– −0.02)

Note: A = additive genetic, D = dominance genetic, C = shared environment, E = unique environment, I = interaction.
* The estimates for girls are the same as for boys.

attention problems, anxiety problems, physical coordination problems, and dependency. Univariate genetic analyses of the problem scales, reported by mother and father, yielded moderate to large additive genetic effects. Heritability estimates ranged from 40% for aggressive behavior to 81% for attention problems. These high heritability estimates confirm the general finding that in early childhood genetic influence is an important determinant of problem behavior (Edelbrock, *et al.,* 1995; Gjone and Novik, 1995; Gjone *et al.,* 1996; Schmitz *et al.,* 1995; Van den Oord *et al.,* 1996; Van der Valk *et al.,* 1998; Zahn-Waxler *et al.,* 1996). Shared environment influences also accounted for a quarter to one third of the variance of aggressive behavior, anxiety problems, and dependency.

Genetic factors accounted for 43% to 53% of the variance of anxiety problems, and shared environmental factors accounted for 19% to 28% of the variance. Previous studies yielded quite varying results. By reviewing the literature of anxiety/depression in childhood, Rice *et al.* (2002) reported a large variability of heritability estimates for parent-rated anxiety/depression (30% to

80%). Most of these studies used samples with children of 8 years and older. Also, the evidence for shared environmental influences on anxiety in children varied across studies (Eley, 1999). Some studies that investigate anxiety in twin pairs ages 7 to 16 years did not find evidence for shared environment (Eaves, 1997; Hudziak *et al.,* 2000; Kuntsi *et al.,* 2000; Thapar and McCuffin, 1995), whereas others reported that one quarter to one third of the variance in anxiety is explained by environmental factors (Edelbrock *et al.,* 1995; Eley and Stevenson, 1999; Feigon *et al.,* 2001). The only two studies that measured internalizing behavior in twin pairs with the same age as in our study did not find any indication for shared environmental influences. Because of the small number of twins pairs the two studies may not have had enough statistical power to detect shared environment. The findings of shared environmental influences on anxiety agree with the findings of a longitudinal study on CBCL internalizing behavior in Dutch twin pairs of 3 to age 12 years (Van der Valk *et al.,* 2003; Bartels *et al.,* submitted). In these studies, shared environmental influences on internalizing behavior were absent at age 3, but played a role at the older ages. Our shared environmental estimate for anxiety corresponds very well with the finding of increased importance of shared environmental influences with increasing age.

In line with other twin studies we found that the genetic factors contributed to the largest part of the variance of aggressive behavior (Edelbrock *et al.,* 1995; Eley *et al.,* 1999; Hudziak *et al.,* 2003; Leve *et al.,* 1998; Schmitz *et al.,* 1995; Van den Oord *et al.,* 1996; Van der Valk *et al.,* 1998). These studies reported heritability estimates around 60%. In a meta-analysis of aggressive behavior, Miles and Carey (1997) reported an overall genetic effect that accounts for 50% of the variance on aggressive behavior. Regarding the role of shared environment on aggressive behavior the findings in the literature are mixed and seem to depend on the kind of aggressive behavior. The CBCL distinguishes between aggressive behavior and rule-breaking behavior. In a study of 1022 twin pairs ages 7 to 9 years, Eley *et al.* (1999) reported that genetic factors accounted for about 60% of the variance of aggressive behavior, whereas shared environmental effects accounted for 10% of the variance. For rule-breaking behavior, genetic and environmental factors both accounted for 40% of the variance. Our DCB aggressive behavior scale includes items that belong to both dimensions, aggressive behavior and rule-breaking behavior. Our scale includes typical aggressive behavior items such as "hits and bites other children," but also typical rule-breaking behavior items such as "intentionally tells lies." The inclusion of items associated with the domain of rule-breaking behavior may have elevated the role of shared environment for the DCB aggressive behavior scale.

Many studies reported a high heritability for attention problems and for hyperactive behavior. A heritability estimate around 75% is a consistent finding across different ages (Rietveld *et al.,* 2003). A heritability of about 80% in 5-year-olds corresponds very well with the general results in literature. Another consistent finding is that MZ correlations are more than twice the DZ correlations, which may point to possible sibling interaction effects (Eaves *et al.,* 2000; Nadder *et al.,* 1998; Rietveld *et al.,* 2003; Simonoff *et al.,* 1998). We have to note that we have also found lower DZ correlations for the physical coordination scale. The following discussion of the contrast effect accounts for this scale as well. Competition or contrast effects are characterized by the behavior of one twin having an inhibitory effect on the behavior of the other twin. Sibling interaction will produce differences in the pattern of MZ and DZ twin correlations that are difficult to distinguish from effects due to genetic dominance (Neale and Cardon, 1992; Simonoff *et al.,* 1998). If a contrast effect is present, it decreases both MZ and DZ correlations, but DZ correlations to a greater extent. In addition, contrast effects predict a smaller variance in MZ as compared to DZ twins (Carey, 1986; Eaves, 1976; Neale and Cardon, 1992). The lower variances in the MZ twins (see Table IV) and the low DZ correlations are suggestive for a contrast effect instead of genetic dominance.

A contrast effect may also be due to bias in parental reports when parents rate the behavior of the child in comparison to the childs' sibling. Whether the contrast effect is due to rater or due to real sibling interaction cannot be tested with the current data. Results from studies that included both teacher and parent ratings suggested that contrast effects are due to rater bias instead of real sibling interactions (Eaves *et al.,* 2000; Simonoff *et al.,* 1998). Therefore the most likely conclusion is that the found contrast effect is due to rater bias. The finding of a less marked contrast effect in the opposite-sex twin pairs favored this explanation. Because boys and girls behave themselves differently, parents may be better able to evaluate the behavior independently of the behavior of the co-twin and their rating is less affected by rater bias (Eaves *et al.,* 2000).

A limitation of the present study is that the heritability estimates are based on parent ratings only. It is a well-known fact now that different informants, such as parents, teachers, and the children themselves tend to disagree in their ratings of children's problem behavior (Achenbach *et al.,* 1987; Simonoff *et al.,* 1995). Such rater differences may influence the estimate of heritabilities. In a recent twin study by Martin *et al.* (2002), the heritability of attention deficit hyperactivity disorder was investigated as assessed by parents and teachers. The heritability estimates were high both for teacher- and parent-rated data, but bivariate analyses suggested that a substantial part of the variance of teacher and parent-rated data was influenced by specific genetic factors. These results clearly stress the need for the use of multiple informants in genetic studies of problem behavior.

The use of the DCB questionnaire makes the comparison of the heritability estimates across studies more complicated. Differences with other studies may reflect not only real age differences but also differences that are due to the use of different assessment instruments. However, the DCB scales showed reasonable correlations with the CBCL, which suggests that the DCB problem scales measure partly similar underlying constructs as the CBCL at ages 3 and 7. It should be mentioned that comparisons of different assessment instruments across different ages may yield "fuzzy" conclusions. A correlation between two assessment instruments may reflect not only the similarity of the underlying construct but also developmental changes.

SUMMARY

The findings of this study suggest that problem behavior at age 5 is affected by both genetic factors and environmental factors. The extent of these influences differs across problem scales. Genetic factors accounted more than 70% of the variance in the parental report of attention problems and emotional lability, while for aggressive behavior, anxiety problems, and dependency shared environmental factors were important too.

ACKNOWLEDGMENTS

This work was supported by a grant of The Netherlands Organization for Scientific Research (Grant 904-57-094).

REFERENCES

Achenbach, T. M. (1995). Empirically based assessment and taxonomy: Applications to clinical research. *Psychol. Assess.* **7**:261–274.

Achenbach, T. M., McConaughy, S. H., and Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychol. Bull.* **101**:213–232.

American Psychiatric Association. (1996). *Diagnostic and Statistical Manual of Mental Disorders, DSM-IV.* Washington, D.C.: American Psychiatric Association.

Bartels, M., Hudziak, J. J., Boomsma, D. I., Rietveld, M. J. H., Van Beijsterveldt, C. E. M., and Van den Oord, E. J. C. G. (2003). A study of parent ratings of internalizing and externalizing problem behavior in 12-year-old twins. *J. Am. Acad. Child Adolesc. Psychiatry* **42**:1351–1359.

Bartels, M., van den Oord, E. J., Hudziak, J. J., Rietveld, M. J. H., van Beijsterveldt, C. E. M., and Boomsma, D. I. Genetic and environmental mechanisms underlying stability and change in problem behaviors at ages 3, 7, 10, and 12. (submitted).

Bentler, P. M., and Dudgeon, P. (1996). Covariance structure analysis: Statistical practice, theory, and directions. *Ann. Rev. Psychol.* **47**:563–592.

Boomsma, D. I. (1998). Twin registers in Europe: An overview. *Twin Res.* **1**:34–51.

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

Carey, G. (1986). Sibling imitation and contrast effects. *Behav. Genet.* **16**:319–341.

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

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

Eaves, L. J., Silberg, J. L., Meyer, J. M., and Maes, H. H. (1997). Genetics and developmental psychopathology. II. The main effects of genes and environment on behavioral problems in the Virginia Twin Study of adolescent behavioral development. *J. Child. Psychol. Psychiatry* **38**:965–980.

Edelbrock, C., Rende, R., Plomin, R., and Thompson, L. A. (1995). A twin study of competence and problem behavior in childhood and early adolescence. *J. Child Psychol. Psychiat. Allied Discipl.* **36**:775–785.

Eley, T. C. (1999). Behavioral genetics as a tool for developmental psychology: Anxiety and depression in children and adolescents. *Clin. Child Fam. Psychol. Rev.* **2**:21–36.

Eley, T. C., Lichtenstein, P., and Stevenson, J. (1999). Sex differences in the etiology of aggressive and nonaggressive antisocial behavior: Results from two twin studies. *Child Dev.* **70**:155–168.

Eley, T. C., and Stevenson, J. (1999). Using genetic analyses to clarify the distinction between depressive and anxious symptoms in children. *J. Abnor. Child Psychol.* **27**:105–114.

Feigon, S. A., Waldman, I. D., Levy, F., and Hay, D. A. (2001). Genetic and environmental influences on separation anxiety disorder symptoms and their moderation by age and sex. *Behav. Genet.* **31**:403–411.

Gjone, H., and Novik, T. S. (1995). Parental ratings of behaviour problems: A twin and general population comparison. *J. Child Psychol. Psychiat. Allied Discipl.* **36**:1213–1224.

Gjone, H., Stevenson, J., Sundet, J. M., and Eilertsen, D. E. (1996). Changes in heritability across increasing levels of behavior problems in young twins. *Behav. Genet.* **26**:419–426.

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

Hudziak, J. J. (2001). The role of phenotypes (diagnoses) in genetic studies of attention-deficit/hyperactivity disorder and related child psychopathology. *Child Adolesc. Psychiatr. Clin. North Am.* **10**:279–297.

Hudziak, J. J., Rudiger, L. P., Neale, M. C., Heath, A. C., and Todd, R. D. (2000). A twin study of inattentive, aggressive, and anxious/depressed behaviors. *J. Am. Acad. Child Adolesc. Psychiatry* **39**:469–476.

Hudziak, J. A., van Beijsterveldt, C. E. M., Bartels, M., Rietveld, M. J. H., Rettew, D. C., Derks, E. M., and Boomsma, D. I. (2003). Individual differences in aggression: Genetic analyses by age, gender, and informant in 3-, 7-, and 10-year-old Dutch twins. *Behav. Genet.* **33**:571–585.

Jöreskog, K. G., and Sörbom, D. I. (1989). *LISREL 7: A guide to the Program and Applications.* Chicago: SPSS Inc.

Jöreskog, K. G., and Sörbom, D. I. (1999). *LISREL 8.3.* Chicago: Scientific Software International.

Kuntsi, J., Gayan, J., and Stevenson, J. (2000). Parents' and teachers' ratings of problem behaviours in children: Genetic and contrast effects. *Twin Res.* **3**:251–258.

Leve, L. D., Winebarger, A. A., Fagot, B. I., Reid, J. B., and Goldsmith, H. H. (1998). Environmental and genetic variance in children's observed and reported maladaptive behavior. *Child Dev.* **69**:1286–1298.

Martin, N., Scourfield, J., and McGuffin, P. (2002). Observer effects and heritability of childhood attention-deficit hyperactivity disorder symptoms. *Br. J. Psychiatry* **180**:260–265.

Miles, D. R., and Carey, G. (1997). Genetic and environmental archictecture of human aggression. *J. Pers. Soc. Psychol.* **72**:207–217.

Nadder, T. S., Silberg, J. L., Eaves, L. J., Maes, H. H., and Meyer, J. M. (1998). Genetic effects on ADHD symptomatology in 7- to 13-year-old twins: Results from a telephone survey. *Behav. Genet.* **28**:83–99.

Nadder, T. S., Silberg, J. L., Rutter, M., Maes, H. H., and Eaves, L. J. (2001). Comparison of multiple measures of ADHD symptomatology: A multivariate genetic analyses. *J. Child Psychol. Psychiatry* **42**:475–486.

Neale, M. C., Boker, S. M., Xie, G., and Maes, H. H. (1999). *Mx: Statistical Modeling,* 5th ed. Richmond, VA: Department of Psychiatry, Virginia Commonwealth University.

Neale, M. C., and Cardon, L. R. (1992). *Methodology for Genetic Studies of Twins and Families*. Dordrecht: Kluwer.

Pfefferbaum, A., Mathalon, D., Sullivan, E., Rawles, J., Zipursky, R., and Lim, K. (1994). A quantitative magnetic resonance imaging study of changes in brain morphology form infancy to late adulthood. *Arch. Neurol.* **51**:874–887.

Rice, F., Gordon, H., and Thapar, A. (2002). The genetic aetiology of childhood depression: A review. *J. Child Psychol. Psychiatry* **43**:65–79.

Rietveld, M. J. H., Hudziak, J. J., Bartels, M., Van Beijsterveldt, C. E. M., and Boomsma, D. I. (2003). Heritability of attention problems in children. I. Cross-sectional results from a study of twins, age 3–12 years. *Am. J. Med. Genet.* **117**:102–113.

Rietveld, M. J. H., Hudziak, J. A., Bartels, M., van Beijsterveldt, C. E. M., and Boomsma, D. I. Heritability of attention problems in children. II. Longitudinal results from a study of twins, age 3 to 12. *J. Child Psychol. Psychiatry.*

Rietveld, M. J. H., Vander Valk, J. C., Bongers, I. L., Stroet, T. M., Slagboom, P. E., and Boomsma, D. I. (2000). Zygosity diagnosis in young twins by parental report. *Twin Res.* **3**:134–141.

Rutter, M., Silberg, J., O'Connor, T., and Simonoff, E. (1999). Genetics and child psychiatry. I advances in quantitative and molecular genetics. *J. Child Psychol. Psychiat. Allied Discipl.* **40**:3–18.

Schmitz, S., Fulker, D. W., and Mrazek, D. A. (1995). Problem behavior in early and middle childhood: An initial behavior genetic analysis. *J. Child Psychol. Psychiat. Allied Discipl.* **36**:1443–4458.

Simonoff, E., Pickles, A., Hervas, A., Silberg, J. L., Rutter, M., and Eaves, L. (1998). Genetic influences on childhood hyperactivity: Contrast effects imply parental rating bias, not sibling interaction. *Psychol. Med.* **28**:825–837.

Simonoff, E., Pickles, A., Hewitt, J., Silberg, J., Rutter, M., Loeber, R., Meyer, J., Neale, M., and Eaves, L. (1995). Multiple raters of disruptive child behavior: Using a genetic strategy to examine shared views and bias. *Behav. Genet.* **25**:311–326.

Spivack, G., and Spotts, J. (1966). *The Devereux Child Behavior (DCB) Rating Scale.* Devon, A: The Devereux Foundation.

Thapar, A., and McGuffin, P. (1995). Are anxiety symptoms in childhood heritable? *J. Child Psychol. Psychiatry* **36**:439–447.

van Baal, G. C., Boomsma, D. I., and de Geus, E. J. (2001). Longitudinal genetic analysis of EEG coherence in young twins. *Behav. Genet.* **31**:637–651.

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

Van der Valk, J. C., Van den Oord, E., Verhulst, F. C., and Boomsma, D. I. (2003). Genetic and environmental contributions to stability and change in children's internalizing and externalizing problems. *J. Am. Acad. Child. Adolesc. Psychiatry* **42**:1–9.

Van der Valk, J. C., Van den Oord, E. J., Verhulst, F. C., and Boomsma, D. (2001). Using parental ratings to study the etiology of 3-year-old twins' problem behaviors: Different views or rater bias? *J. Child Psychol. Psychiatry* **42**:921–931.

Van der Valk, J. C., Verhulst, F. C., Stroet, T. M., and Boomsma, D. I. (1998). Quantitative genetic analysis of internalising and externalising problems in a large sample of 3-year-old twins. *Twin Res.* **1**:25–33.

Verhulst, F., and Van der Ende, J. (2001). Rating scales. In M. Rutter and E. A. Taylor (eds.), *Child and Adolescent Psychiatry,* 4th ed. (pp. 70–86). Oxford: Blackwell Science.

Zahn-Waxler, C., Schmitz, S., Fulker, D., Robinson, J., and Emde, R. (1996). Behavior problems in 5-year-old monozygotic and dizygotic twins: Genetic and environmental influences, patterns of regulation, and internalization of control. *Dev. Psychopathol.* **8**:103–122.

Edited by Wendy Slutske