

Genetic and environmental influences on Anxious/Depression during childhood: a study from the Netherlands Twin Register

D. I. Boomsma[*,†], C. E. M. van Beijsterveldt[†] and J. J. Hudziak[‡]

[†]Department of Biological Psychology, Vrije Universiteit, Amsterdam, the Netherlands and [‡]Department of Psychiatry and Medicine (Division of Human Genetics), Center for Children, Youth and Families, and University of Vermont, College of Medicine, Burlington, USA

**Corresponding author: D. I. Boomsma, Department of Biological Psychology, Vrije Universiteit, Van der Boechorststraat 1, 1081 BT Amsterdam, the Netherlands. E-mail: dorret@psy.vu.nl*

For a large sample of twin pairs from the Netherlands Twins Register who were recruited at birth and followed through childhood, we obtained parental ratings of Anxious/Depression (A/D). Maternal ratings were obtained at ages 3 years (for 9025 twin pairs), 5 years (9222 pairs), 7 years (7331 pairs), 10 years (4430 pairs) and 12 years (2363 pairs). For 60–90% of the pairs, father ratings were also available. Multivariate genetic models were used to test for rater-independent and rater-specific assessments of A/D and to determine the genetic and environmental influences on individual differences in A/D at different ages. At all ages, monozygotic twins resembled each other more closely for A/D than dizygotic twins, implying genetic influences on variation in A/D. Opposite sex twin pairs resembled each other to same extent as same-sex dizygotic twins, suggesting that the same genes are expressed in boys and girls. Heritability estimates for rater-independent A/D were high in 3-year olds (76%) and decreased in size as children grew up [60% at age 5, 67% at age 7, 53% at age 10 (60% in boys) and 48% at age 12 years]. The decrease in genetic influences was accompanied by an increase in the influence of the shared family environment [absent at ages 3 and 7, 16% at age 5, 20% at age 10 (5% in boys) and 18% at age 12 years]. The agreement between parental A/D ratings was between 0.5 and 0.7, with somewhat higher correlations for the youngest group. Disagreement in ratings between the parents was not merely the result of unreliability or rater bias. Both the parents provided unique information from their own perspective on the behavior of their children. Significant influences of genetic and shared environmental factors were found for the unique parental views. At all ages, the contribution of shared environmental factors to variation in rater-specific views was higher for father ratings. Also, at all ages except age 12, the heritability estimates for the rater-specific phenotype were higher for mother ratings (59% at age 3 and decreasing to 27% at age 12 years) than for father ratings (between 14 and 29%). Differences between children, even as young as 3 years, in A/D are to a large extent due to genetic differences. As children grow up, the variation in A/D is due in equal parts to genetic and environmental influences. Anxious/Depression, unlike many other common childhood psychopathologies, is influenced by the shared family environment. These findings may provide support for why certain family therapeutic approaches are effective in the A/D spectrum of illnesses.

Keywords: Anxious/Depression, childhood, family environment, heritability, rater bias, twins

Received 5 January 2005, revised 10 March 2005, accepted for publication 16 March 2005

The case has been made that studies which estimate the genetic and environmental contributions to individual differences in health and disease and response to treatment are essential for the reduction of illness (Collins 2004). With the remarkable success of the Human Genome Project (Collins *et al.* 2003), there are an increasing number of elegant techniques available to identify genes that influence the variation in complex traits and disorders and their underlying pathophysiology. However, it is likely that the complexity of common disorders, such as those studied in the field of psychiatric genetics, will demand from us to consider contributions of modifying factors when engaging in gene-hunting expeditions. In this article, we look at Anxious/Depression (A/D) in the very young and young and address the influence of three modifying factors, i.e. development, sex and informant, on the genetic architecture of the A/D phenotype.

With complex psychiatric traits such as A/D, characterization of the phenotype remains an issue, especially in subjects who are too young to report on their own behavioral and emotional problems. In the assessment of very young children, the clinician or researcher has to rely on informants who rate the behavior of the child. The importance of genetic and environmental factors may differ according to the informant. When multiple informants such as fathers and

 doi: 10.1111/j.1601-183X.2005.00141.x

mothers rate the behavior of the same child, it is necessary to consider to what extent their information may or not be correlated. Disagreement may be due to a variety of factors, chief among these the fact that informants may be identifying different aspects of the phenotype.

Contributions of genetic factors to individual differences may vary across sex and across age. It is widely accepted that different genetic factors affect core neurodevelopmental processes such as neuronalgenesis, synaptogenesis, myelination and apoptosis. As such, it should be axiomatic that the genetic influences on brain development will have similar and sentinel influences on children's behavioral and emotional problems (such as A/D) and that these influences may vary in their relative importance depending on the age of the child.

The influence of environmental factors may also vary with age and sex of the child. This statement applies not only to the magnitude of these influences (i.e. percentage of variation explained by environmental effects) but also to the type of environmental influence. The influence of the family environment, often referred to as 'shared' or 'common' environment, which includes the effects of parental education, socio-economic status of the family and rearing practices may depend on the age or sex of the child. For example, it is now well established that as children grow older, the large initial influence of shared family environment disappears, whereas the influence of genetic factors on cognitive abilities increases (Posthuma *et al*. 2002; Rietveld *et al*. 2003a). As we review below, for childhood psychopathology, such changes in genetic and environmental influences across development are much less clear. It has been argued that until confounds, or modifiers, such as phenotypic assessment and rater bias and interactions between genes and age and between genes and sex can be addressed, the full impact of molecular genetic studies will not be realized (Rutter & Silberg 2002). All these issues can be addressed in twin studies, but only if the samples are large enough, if the phenotypic information is collected from multiple raters and if the sample is followed across development.

Genetics of A/D behavior in children

The search for genetic influences on symptoms of anxiety and depression in humans has, until recently, focused primarily on adult samples. Relatively, few studies on anxiety-related behaviors or depression in preschool and school children have been published. These studies provide evidence of both genetic and environmental influences on a broad variety of mood disorders but, taken together, do not provide a very clear or consistent picture (e.g. Rice *et al*. 2002a; Rutter 2003; Todd & Botteron 2002; Wamboldt & Wamboldt 2000).

The Child Behavior Checklist (CBCL) is a commonly used epidemiologic screening instrument that queries for a number of syndromes, including A/D. Studies which used the CBCL A/D syndrome scale in very young twins have reported a strong genetic contribution to A/D behavior. Van den Oord *et al*. (1996) and Derks *et al*. (2004) found a high heritability (around 70%) for A/D behavior in large samples of 3-year olds who were rated by both the parents. Schmitz *et al*. (1995) reported a heritability of 32% for maternal ratings of A/D in a sample of 260 3-year-old twin pairs. The shared family environment accounted for 19% of the variation. In the same study, the heritability of A/D in 203 twin pairs aged 7 years and 7 months was 50%. The contribution of shared environment to variation in A/D scores was 15%. This estimate, however, did not reach statistical significance. Eley *et al*. (2003) looked at 16 items on anxiety-related behaviors in a large study of preschool children (4564 twin pairs). They found significant heritability for mother-reported obsessive-compulsive behavior and shyness/inhibition (estimates between 65 and 76%) and lower heritability for separation anxiety (39%) and fears (44%). Variation in these last two anxiety scales was also significantly influenced by shared environmental factors (36 and 18%).

Studies of samples of somewhat older children and adolescents also show evidence for the importance of genetic factors. Hudziak *et al*. (2000) estimated the genetic and environmental contributions to maternally rated A/D in a sample of 492 male and female US twin pairs aged 8–12 years. In this study, the best-fitting model was one that included a moderate contribution of additive genetic factors (65% in males and 61% in females), a moderate contribution of non-shared environmental factors. There was no evidence of shared environmental influences. Sample-size limitations did not allow for testing of sex effects. Edelbrock *et al*. (1995) reported moderate additive genetic influences (34%) on A/D assessed with the CBCL, which was completed by parents of 181 pairs of same-sex twins aged 7–15 years. For shared environment, an estimate of 30% was reported. However, both estimates did not reach statistical significance, probably reflecting that in relatively small samples, it is difficult to distinguish between shared genes and shared environment when trying to explain familial correlations. Eaves *et al*. (1997) obtained extensive data on separation anxiety, over anxiety and depression by interview assessments and on anxiety and depression by questionnaire measures from twins, their fathers and their mothers in a population-based, unselected sample of 1412 Caucasian twin pairs aged 8–16 years. Parent-rated separation anxiety was highly heritable for girls (74%) but not for boys. The sex difference did not replicate for any of the other measures. Self-ratings of separation anxiety showed much lower heritabilities, both in boys (19%) and in girls (31%). The finding of lower heritability for self-assessments than for parental ratings replicated for anxiety and for interview and questionnaire measures of depression; whereas the heritabilities based on parental ratings were around 50% or higher, the twins' self-ratings of anxiety and depression showed heritabilities of around 15%. Significant contributions of shared

environment were found only for variation in twins' self-ratings of anxiety (67% in boys and 55% in girls). In contrast to these findings, Rice *et al*. (2002b), in a study of 1463 families with twins aged 8–17 years, found higher heritabilities for self-ratings of depression (55% in twins over 11 years) than for mother-rated depression (28% in twins over 11 years and 15% in twins younger than 11 years). In accordance with the study of Eaves *et al*., they also found little evidence for sex differences in heritabilities. Rice *et al*. also provide evidence that shared environmental effects are somewhat more important in children (8–11 years) than in adolescents (above 11 years) when analyzing mother-rated depression. Some of the shared environmental variance could be attributed to maternal depression and anxiety symptoms. However, no evidence of shared family environment was seen on the twins' self-ratings of depression. Rice *et al*. concluded that their results call for further research to identify genetic and environmental factors on anxiety depression across development.

Two adoption studies find little evidence for heritability of A/D. Van den Oord *et al*. (1994) looked at mother-rated A/D in a sample of international adoptees (mean age 12.4 years) which comprised a group of biological siblings (111 pairs), a group of non-biological siblings (221 pairs) and a group of singletons (94). Genetic and shared environmental influences together explained around 50% of the variation in A/D, but genetic influences were very small. Eley *et al*. (1998) confirmed these results in middle childhood using a sample from the Colorado Adoption Project, which included 77 unrelated, adoptive sibling pairs and 93 non-adopted biological sibling pairs. Children's depressive symptoms were reported by the parents at ages 7, 9, 10, 11 and 12 years and by the children themselves at ages 9, 10, 11 and 12 years (child self-report data were available for 56 adoptive pairs and 72 biological sibling pairs). Average ratings across parents and across time of the CBCL A/D scale were analyzed and showed negligible heritability estimates for both parent and child ratings. Shared environment effects were seen for the parental ratings, based on similar correlations for adoptive and biological siblings ($r = 0.51$). The pattern of correlations for child self-report was puzzling: there was no correlation in biological siblings and a correlation of 0.21 in adoptive siblings. These two adoption studies suggest that genetic effects on depressive symptoms are small and that shared environment may play an important role during early adolescence.

The studies presented above in general find evidence for familial clustering of A/D in children. Estimates of heritability from twin and adoption studies differ, and evidence for the importance of shared family environment is found in some studies but not in others. However, most studies of A/D in children suffer from a lack of statistical power to detect shared environmental influences. Some studies, although not all, are characterized by a wide age range which may obscure any genotype by age interactions and by the use of a single informant. We have reported previously on the causes of individual differences in aggression (Bartels *et al*. 2003a; Hudziak *et al*. 2003; Van Beijsterveldt *et al*. 2003) and attention problems (Rietveld *et al*. 2003b; 2004) in large groups of twins, followed from birth through age of 12 years. These studies have allowed the evaluation of the impact of age, sex and informant on estimates of genetic and environmental influences on aggression and attention problems. In each case, we found evidence for changing genetic and environmental influences across development. Further, we have found differences in how mothers and fathers report on their children's behavior and have reported the effects of those differences on genetic and environmental estimates. It is the aim of this article to test rater models for A/D to allow comparison of our findings on this internalizing syndrome with prior reports as well as to compare our findings to the results obtained on externalizing behaviors such as aggression and attention problems.

Variance by informants

Parental ratings of behavioral and emotional problems in their children are predictive of childhood psychopathology (e.g. Hay *et al*. 1999), even though fathers and mothers disagree to a certain extent about the degree of the problems. The disagreement between the parents may reflect unreliability or rater bias but can also reflect the unique perspective fathers and mothers have on the behavior of their children. Agreement and disagreement on the behavioral and emotional problems of offspring reflect a problem which is inherent to twin and family studies in which parents are asked to rate the behavior of their (young) children. If they rate more than one child-rating bias may lead to overestimation of the resemblance between siblings or between twins. If parental bias affects the behavioral ratings of their children, this will systematically influence the ratings of both the children. Such bias thus results in more similarity in the offspring's behavior. A genetic analysis of the resemblance of siblings or twins will thus suggest the presence of shared environmental factors on their behavior. To disentangle the effects due to real shared environment and the effect due to children being rated by the same observer, studies are needed in which sibling or twin pairs are rated by more than one person. This allows the distinction between variance that is common to both the raters (and thus not confounded by rater bias or measurement error) and variance that is unique to each rater (and which may include rater bias). With twin data, the variance common to raters can be decomposed into genetic, shared and non-shared (or unique) environmental parts. The estimate of the proportion of common rater variance explained by shared environment can be regarded as 'real' shared environmental effects unaffected by rater bias. We use these methods (e.g. Bartels *et al*. 2004; Neale & Stevenson 1989) to examine rater effects in twin data. Note that these models have broader applications than in

twin or sibling studies only. For example, researchers studying the genetics of behavior and personality in animals are faced with the same issues of assessing a particular phenotype in subjects who cannot report on their own behaviors and feelings.

Contribution of shared environment

In the genetic analysis of twin data, the influence of the shared environment is inferred through its effect on twin correlations (Boomsma *et al*. 2002a). Shared environment affects the similarity for monozygotic (MZ) and dizygotic (DZ) twins to the same degree. The power to detect such shared environmental effects in the classical twin study is low. For example, if heritability is 50% and shared environmental effects explain 10% of the phenotypic variance, over 5000 twin pairs are needed to obtain a statistical power of 80% (Posthuma & Boomsma 2000). If shared environmental effects explain 20% of the phenotypic variance, over 1000 twin pairs are still needed. The influence of genetic factors on behavioral differences between children is inferred from the difference between MZ and DZ twin correlations. Twice the difference between these correlations gives a first estimate of heritability (Falconer & Mackay 1996). In this article, we analyze maternal and paternal ratings for A/D which were obtained at ages 3, 5, 7, 10 and 12 years. Especially at the younger ages, the data set is large (at ages 3 and 5 years, maternal ratings are available for over 18 000 children). We thus are in a position to detect the presence of shared environmental influences on variation in A/D, even if the effects are small. This study is the first to look at A/D with enough statistical power to test for the contributions of shared environment across development, to assess the effects of rater bias by using multiple informants (mother and father), to assess the phenotype and to test whether heritability differs by sex and whether the same genes and environmental influences are expressed in boys and girls. With these data, we aim to contribute to strategies on phenotypic identification for future studies of the molecular genetic and environmental factors that influence A/D in children.

Materials and methods

Sample

The data presented in this article come from an ongoing longitudinal study, which examines the genetic and environmental influences on the development of behavioral and emotional problems in families with 3- to 12-year-old twins. The families are volunteer members of the Netherlands Twin Register (NTR), established by the Department of Biological Psychology at the Free University in Amsterdam (Boomsma *et al*. 2002b). From 1987 onwards, the NTR recruits families with twins a few weeks or months after birth. Currently 40–50% of all multiple births are registered by the NTR. For this study, we included data of 3- and 5-year-old twin pairs from birth cohorts 1986–1997, of 7-year-old twin pairs from cohorts 1986–1996, of 10-year-old twin pairs from cohorts 1986–1993 and of 12-year-old twin pairs from birth cohort 1986–1990. Both the parents were asked to complete questionnaires about problem behaviors for the first- and second-born (oldest and youngest) twin at ages 3, 5, 7, 10 and 12 years. Due to funding problems, the survey for 3-year olds was not sent to fathers of twins born between May 1989 and November 1991. Two months after mailing out the questionnaire, a reminder was sent to the non-responders. After 4 months, those who still had not responded were telephoned, if resources for telephone follow-up were available at the time. This procedure resulted in a response rate (at least one parental questionnaire returned) of 72% at age 3 years. At ages 5, 7, 10 and 12 years, the participation rates were 67, 66, 64 and 64%, respectively. Note that if a family did not participate at a particular age, they were again approached for the next mailing. Hence, a response rate of 66% at age 7 means that 66% of all the registered families with a twin pair that had reached this particular age returned the questionnaire. Thus, families who did not participate at one age of the twins could enter the study again at subsequent ages. Non-responders also include twin families who changed addresses.

For 854 same-sex twin pairs, zygosity was based on blood group ($n = 436$) or DNA polymorphisms ($n = 418$). For the remaining twins, zygosity was determined by questionnaire items about physical similarity and frequency of confusion of the twins by family and strangers (Goldsmith 1991; Rietveld *et al*. 2000) obtained at age 3, 5, 7, 10 and 12 years. The classification of zygosity was based on a discriminant analysis, relating the questionnaire items to zygosity based on blood/DNA typing in a group of same-sex twin pairs. According to this analysis, the zygosity was correctly classified by questionnaire in nearly 95% of the cases. If a discrepancy in zygosity status appeared across ages, then the most frequent zygosity status was used.

A family was excluded if one of the twins had a disease or handicap that interfered severely with normal daily functioning (about 2%). Table 1 summarizes an overview of the number of twin pairs with complete maternal rating and zygosity data. The percentage of families with both maternal and paternal ratings of problem behaviors is given between brackets.

Instruments

At age 3, problem behavior was measured with the CBCL/2–3, a questionnaire that includes 100 items that describe specific behavioral, emotional and social problems. Parents were asked to rate the behavior that the child displayed in the past 2 months on a 3-point scale: 0 = problem item not true, 1 = item somewhat or sometimes true and 2 = item

Table 1: Number of maternal reports on the Child Behavior Checklist Anxious/Depression for twin pairs aged 3, 7, 10 and 12 years and on Devereux anxiety at age 5 years. The percentages within brackets give the percentage of families for whom both maternal and paternal reports are available

	Age 3	Age 5	Age 7	Age 10	Age 12
Zygosity					
mzm	1440 (66.3)	1489 (91.7)	1225 (75.8)	756 (75.8)	429 (77.2)
dzm	1544 (65.2)	1549 (90.4)	1240 (75.6)	684 (71.5)	379 (74.7)
mzf	1567 (69.5)	1749 (90.1)	1405 (75.1)	916 (72.3)	495 (76.6)
dzf	1432 (64.8)	1447 (90.7)	1153 (74.2)	659 (73.1)	352 (79.3)
dos_mf	1564 (63.4)	1535 (89.6)	1193 (76.0)	736 (73.0)	372 (76.9)
dos_fm	1478 (63.3)	1444 (90.2)	1115 (73.7)	679 (73.9)	336 (78.3)
Total	9025	9222	7331	4430	2363

dos_fm, opposite-sex twin pairs with female first-born twin and male second-born twin; dos_mf, opposite-sex twin pairs with male first-born twin and female second-born twin; dzf, dizygotic female twin pairs; dzm, dizygotic male twin pairs; mzf, monozygotic female twin pairs; mzm, monozygotic male twin pairs.

very true or often true. The A/D scale is based on factor analyses of data from several Dutch population samples (Koot *et al*. 1997) and is compatible with the syndrome scale developed by Achenbach (1992). The A/D scale derived from the CBCL/2–3 contains nine items. Anxiety at age 5 was measured with a short version of the Devereux Child Behavior (DCB) rating scale (Spivack & Spotts 1966) that consisted of 42 items (Van Beijsterveldt *et al*. 2004). 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 = never and 5 = very frequently. The anxiety scale included six items. At ages 7, 10 and 12 years, A/D was measured with the CBCL/4–18 (Achenbach 1991; Verhulst *et al*. 1996), a 113-item questionnaire developed to measure problem behavior in 4–18-year-old children. Parents were asked to rate the behavior of the children in the preceding 6 months on a 3-point scale. The A/D scale contains 14 items. An overview of items assessed at each age is given in Appendix 1. Good reliability of the A/D scales has been reported in Dutch epidemiological samples (Verhulst *et al*. 1996; Koot *et al*. 1997). For the CBCL/2–3, the 3-week test–retest correlation was 0.85 (Koot *et al*. 1997), and for the CBCL/4–18, it was 0.84 (Verhulst *et al*. 1996). Internal consistency of the A/D scale was 0.76 for the CBCL/2–3, 0.60 for DCB anxiety at age 5 (Van Beijsterveldt *et al*. 2004) and 0.72 for the CBCL/4–18.

Statistical analyses

Sex, zygosity and rater differences in means

Means and standard deviations for A/D were calculated with SPSS/WINDOWS 11.0 (SPSS 2001). The effects of sex, rater and zygosity on mean scores were evaluated with ANOVA. These effects were first examined in first-born twins. If an effect was found, we tested whether it replicated in second-born twins. To correct for multiple testing, the α-level was set at 0.01, and an effect was only assumed to be present if it was significant in first- and second-born twins.

Genetic analyses

Twin correlations, cross-twin–cross-rater and interparent correlations were calculated by the statistical software program MX (Neale *et al*. 1999). The genetic analyses, based on structural equation modeling, were also carried out with MX. The essence of genetic model fitting is the decomposition of the observed variance in a phenotype (here A/D) due to additive genetic effects (A), shared or common environment effects (C) and non-shared environment (E) factors (Boomsma *et al*. 2002a). 'A' represents the additive effects of alleles at multiple loci and 'C' represents common environment effects which are shared by children growing up in the same family such as parental rearing practices, parental income and religion or socio-economic status. 'E' represents all non-shared environmental influences, including measurement error. In model fitting of twin data, the influences of A, C and E are inferred through their effects on the covariances of relatives. Genetic and shared environmental influences predict similarity among relatives, while non-shared environmental effects are a source of phenotypic differences. The upper part in Figs 1 and 2 summarizes the essence of genetic model fitting to twin data. The upper circles represent the latent, unmeasured genetic and environmental factors. The genetic factors are correlated 1 in MZ twins, as they are nearly always genetically identical. For DZ twins, the additive genetic factors (A) are correlated 0.5, because DZ share on the average half of their segregating genes. The environment shared (C) by two members of a pair is correlated unity and is assumed not to depend on the zygosity of the twins. The non-shared environment (E) is by definition, uncorrelated between two members of a pair, either MZ or DZ. The parameters *a*, *c* and *e* are loadings of the phenotype on the

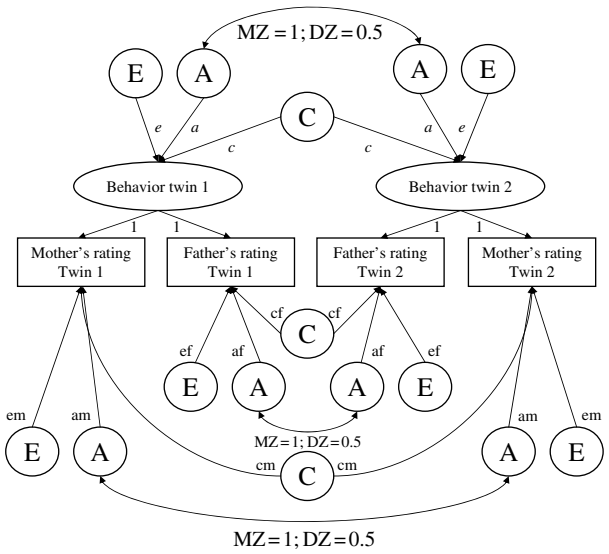

Figure 1: Psychometric model: ratings of the behavior of twin 1 and twin 2 by father and mother are represented by boxes. Latent factors are represented by circles (A stands for additive genetic influences, C for common family environment and E for unique environment). The model combines the mother and father reports into one common latent phenotype, which represents the common parental view of the behavior of the child, i.e. the part of behavior similarly assessed by both the parents. The disagreement between parents is seen as a unique aspect in the child's behavior. The variation of the unique part is decomposed into A, C and E (with path coefficients Am, Cm and Em for mother's unique view and Af, Cf and Ef for father's unique view).

latent factors A, C and E, respectively. The proportion of the variation accounted for by heritability or environmental influences is calculated by squaring the parameters *a*, *c* and *e* and dividing them by the total variance ($a^2 + c^2 + e^2$).

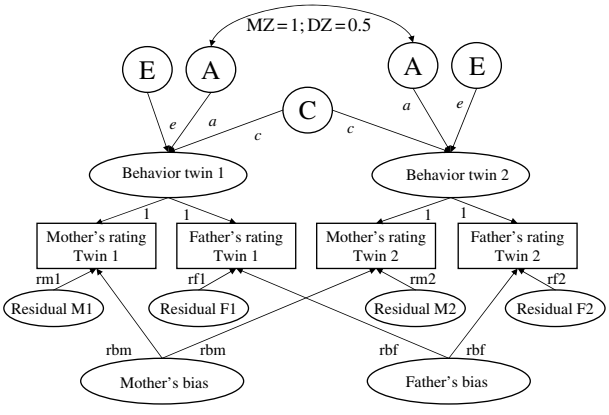

Figure 2: Rater-bias model (see also caption for Fig. 1): no unique specific view of the rater is assumed. The rater-specific part is partitioned into rater bias for mother and father (correlated errors across twins, within each informant) and an unreliability part (residual).

Rater models

Because we collected data from both fathers and mothers, we can evaluate if, and to what extent, rater bias might affect the results. We tested a psychometric model and a rater-bias model (Bartels *et al*. 2003b; 2004; Hewitt *et al*. 1992; Van der Valk *et al*. 2001; 2003). Both models combine the mother and the father reports to one common latent phenotype of A/D but differ with respect to the modeling of disagreement among parents. The common latent phenotype represents the common parental view of the behavior of the child, i.e. the part of behavior similarly assessed by both parents, and can be considered as independent of rater biases and unreliability of the ratings. In the genetic analyses, the variation of this latent factor is decomposed into its genetic and environmental parts.

In addition to the common part, the psychometric model (Fig. 1) modeled the disagreement among parents as a unique aspect of his or her child's behaviors. In the psychometric model, the variation of the unique part is decomposed due to genetic, shared and non-shared environmental factors. By testing the significance of genetic effects on the unique part, it can be established that the raters must have been assessing a 'real' but unique aspect of the child's behavior. Error and/or unreliability cannot cause the systematic effects necessary to estimate these genetic influences. The shared environmental effects on the unique part may be confounded by rater bias, because rater bias mimics shared environmental effects. As the parents assess the behavior of both twins, parental bias will influence the ratings of both the twins in a similar way and make them more alike.

In contrast to the psychometric model, the rater-bias model assumes no unique specific view of the rater. As shown in Fig. 2, in the rater-bias model, the rater-specific part is partitioned into rater bias (correlated errors across twins, within each informant) and an unreliability part.

Model fitting

We first tested whether the psychometric or the rater-bias model gave the most acceptable fit to the data, in order to establish which of these two models best describes the agreement/disagreement between the mother and the father reports. Secondly, we tested whether the model could be simplified by dropping one or more of the variance components (A, C or E) for the common latent phenotype or by dropping A or C for the rater-specific part. The rater-specific E factor was never dropped from the model, because in addition to non-shared environmental experiences, this factor includes measurement error. Thirdly, the significance of sex differences in one or more of the variance components was tested by constraining the parameter

estimates to be equal across sex. Finally, it was tested whether the parameters describing the specific maternal and paternal views could be constrained to be equal.

The different models were fitted to the data with MX (Neale *et al.* 1999), by the method of raw maximum likelihood estimation. MX provides the possibility to analyze raw data with any pattern of missing data. This allowed the use of all twin data, whether or not there were missing data for one of the parents. In this procedure, the likelihood is calculated separately for each pedigree, and the product of these likelihoods (i.e. the sum of the log-likelihood) is maximized. The use of maximum likelihood estimation requires that the data are approximately normally distributed. The A/D scales of the CBCL and DCB showed a positively skewed distribution. Therefore, prior to the genetic analyses, the normalizing procedure of PRELIS (Jöreskog & Sörbom 1996) was applied to the data to approximate a normal distribution.

Goodness-of-fit was assessed by likelihood-ratio chi-square tests. These tests compare the differences between −2-log likelihood of a full model with that of a restricted nested model. This difference is distributed as a X^2, and the degrees of freedom (df) for this test are equal to the difference between the number of estimated parameters in the full model and that in a restricted model. A large X^2 value in comparison to the number of df suggest that the simpler model does not fit to the data as well as the more complex models. To select the best model, the Akaike's information criterion (AIC) was used. The AIC is the X^2 minus twice the df (Akaike 1987). The model with the lowest AIC value is considered the most parsimonious model.

Results

Descriptive statistics

Table 1 summarizes the number of twin pairs by sex and zygosity at each age. The sample size decreases by age by virtue of the fact that this is an ongoing longitudinal study, in which we add new-born twins annually. The sample size was always smaller for father data with some deviant participation percentages at ages 3 and 5 years. The lower percentage for father ratings at age 3 is explained by the fact that paternal questionnaires were not collected for a period of about 2 years. The higher percentage at age 5 may be due to the fact that the number of problem behavior items is smaller and that the items for both the parents were part of one survey booklet.

Table 2 summarizes the means and variances for the untransformed measures of A/D by age and sex for mother and father reports. For all CBCL ratings, mothers reported more A/D symptoms than fathers for both boys and girls. The differences in mean scores between mother and father reports across all ages were significant (age 3: $F_{1,5672} = 35.28$, $P < 0.001$; age 7: $F_{1,5412} = 338.86$, $P < 0.001$; age 10: $F_{1,3171} = 187.62$, $P < 0.001$ and age 12: $F_{1,1776} = 79.51$, $P < 0.001$). At age 5, fathers reported more anxiety problems than the mothers did ($F_{1,8284} = 34.76$). At age 5, girls had higher anxiety scores than boys ($F_{1,8284} = 39.51$, $P < 0.001$), but no significant sex differences were seen at the other ages. No differences between zygosity groups were found with exception of A/D at 7 years. The means were higher for DZ than for MZ twins ($F_{1,5412} = 6.49$, $P = 0.011$). All results are given for the oldest twin, but the same results were obtained for the youngest twin.

Phenotypic and twin correlations

Twin correlations are provided in the first and second column of Table 3. For each age, sex and rater, MZ twins showed higher correlations than DZ twins. The MZ correlations varied between 0.60 and 0.73 across raters and sex, and the DZ correlations were between 0.31 and 0.55. At most ages, DZ correlations were larger than expected on the basis of additive genetic influences only (more than half of MZ correlations) and suggest shared environmental influences on individual differences in A/D. At all ages, the correlations for DZ twins of opposite sex are strikingly similar in magnitude to the correlations in same-sex dizygotic pairs, indicating that the same genes and shared environmental factors operate in boys and girls. In the third and fourth column of Table 3, the cross-twin–cross-rater correlations are given. These are the correlations between A/D in one twin as rated by the father and A/D in the other twin as rated by the mother (and vice versa) and indicate the similarity between the twins' problem behaviors upon which both the parents agree (common parental view). The cross-twin–cross-rater correlations give a first indication of the involvement of genes and environmental factors on the common parental view. At all ages, the MZ cross-rater correlations were larger than the DZ cross-rater correlations and suggested genetic influences on the common parental view. At older ages, the cross-rater DZ twin correlations seemed to be larger than that expected on the basis of genetic influences alone, and therefore, shared environmental influences seemed also to contribute to the common parental view. For each age, the cross-twin–cross-rater correlations were smaller than the intrarater (mother or father rating both twins) twin correlations. The differences denote the part which is unique to a particular rater, also called the unique rater-specific view. These differences seemed to increase with age, which may point to an increased specific parental view. The last two columns of correlations in Table 3 give the interparent agreement for the oldest and the youngest twins. These correlations vary between 0.65 and 0.70 at age 3 and between 0.55 and 0.65 at all other ages, also pointing to the possibility of an increased importance of specific parental views.

Table 2: Means and standard deviations (SD) for untransformed Child Behavior Checklist Anxious/Depression (age 3, 7, 10 and 12 years) and for Devereux anxiety (age 5 years) across ages, zygosity and raters (mother and father)

		mzm Mean (SD)	dzm Mean (SD)	mzf Mean (SD)	dzf Mean (SD)	dos_boys Mean (SD)	dos_girls Mean (SD)	Boys Mean (SD)	Girls Mean (SD)
Age 3	Mother	3.49 (2.9)	3.60 (3.1)	3.86 (3.1)	3.87 (3.3)	3.73 (3.1)	3.15 (2.9)	3.63 (3.0)	3.64 (3.1)
	Father	3.23 (2.8)	3.41 (3.0)	3.52 (3.0)	3.67 (3.1)	3.52 (2.9)	2.94 (2.8)	3.41 (2.9)	3.39 (3.0)
Age 5	Mother	10.61 (3.3)	10.87 (3.4)	11.14 (3.4)	11.27 (3.5)	10.58 (3.4)	10.81 (3.3)	10.69 (3.4)	11.08 (3.4)
	Father	10.68 (3.2)	11.10 (3.4)	11.33 (3.4)	11.43 (3.5)	10.98 (3.4)	11.08 (3.3)	10.92 (3.3)	11.28 (3.4)
Age 7	Mother	2.12 (2.6)	2.46 (2.9)	2.29 (2.7)	2.77 (3.3)	2.17 (2.8)	2.27 (2.8)	2.25 (2.8)	2.43 (2.9)
	Father	1.54 (2.1)	1.80 (2.3)	1.66 (2.3)	1.83 (2.5)	1.55 (2.1)	1.63 (2.2)	1.63 (2.2)	1.70 (2.3)
Age 10	Mother	2.44 (2.9)	2.74 (3.5)	2.73 (3.3)	2.88 (3.6)	2.58 (3.3)	2.55 (3.3)	2.58 (3.2)	2.72 (3.4)
	Father	1.72 (2.3)	1.99 (2.9)	1.89 (2.6)	2.18 (3.0)	1.85 (2.5)	1.92 (2.6)	1.84 (2.6)	1.98 (2.7)
Age 12	Mother	2.24 (3.0)	2.52 (3.6)	2.49 (3.0)	2.38 (3.1)	2.39 (3.2)	2.12 (2.7)	2.38 (3.3)	2.35 (3.0)
	Father	1.56 (2.4)	1.96 (3.0)	1.71 (2.3)	2.09 (3.0)	1.66 (2.4)	1.62 (2.5)	1.72 (2.6)	1.80 (2.6)

boys, all same-sex boys; girls, all same-sex girls; dos_boys, boys with opposite-sex twin; dos_girls, girls with opposite-sex twin; dzf, dizygotic female twin pairs; dzm, dizygotic male twin pairs; mzm, monozygotic male twin pairs; mzf, monozygotic female twin pairs.

Genetic analyses of maternal and paternal ratings

Table 4 summarizes the estimates of genetic and environmental influences by fitting separate models to the ratings of A/D by mother and by father. For each age, a model with A, C and E factors and with sex differences in parameter estimates was fitted to the data. The parameter estimates for a^2, c^2 and e^2 were comparable for mother and father ratings of A/D. With increasing age, the amount of variance explained by genetic influences seemed to decrease, while the amount of variance explained by shared environmental influences increased slightly. With exception of age 10, sex differences in parameter estimates were not significant. At age 10, the influences of shared environment seemed to be larger for girls than for boys. In these univariate analyses, the estimates of the shared environmental effects may be confounded by parental bias. Because the parent assesses the behavior of both the twins, parental bias will influence the ratings of both the twins, which may result in more similarity in reported behavior. To disentangle the effects due to 'real' shared environment and the effect due to rater bias, rater models were applied simultaneously to the paternal and maternal ratings of twins to distinguish these two sources of variance.

Rater models

The model fitting results of the simultaneous analysis of mother and father ratings are summarized in Appendix 2. Firstly, we tested if the rater-bias model or the psychometric model best described the data. The goodness-of-fit of these models was compared to that of a fully saturated model. For all ages, the psychometric model provided the best description of the data, as indicated by the larger increase in chi-square for the rater bias model. This implies that the rater-specific part of the maternal and paternal ratings includes meaningful variation which can be attributed to A, C and E factors. Thus, parents have, in addition to a common view of the behavior of their children, unique views that are specific to father and mother.

For the psychometric model, we tested at each age whether (1) the A, C and E variance components contributed significantly to the variance of A/D agreed upon by both the parents and specific by parents, (2) the A, C and E variance components could be constrained across sex and (3) the A, C and E variance components of the unique parental view could be constrained across raters. For all ages, dropping A or E for the common latent phenotype of A/D gave a serious deterioration in fit. This means that genetic and non-shared environmental factors contributed significantly to the variance of the common latent phenotype. Results for dropping the common C factor differed across ages. At ages 3 and 7, the estimates for the common C factor were nearly zero and could be dropped from the model without a significant change of the fit. At ages 5 and 12, leaving out the common C led to a significant drop in fit. At age 10, dropping the common C factor did not lead to a significant drop in fit when tested against two df. However, inspection of the parameter estimates of common C revealed large differences between boys and girls. Shared environmental influences were important for girls but not for boys.

Table 5 summarizes the standardized parameter estimates of genetic and environmental contributions to the variance of the common latent phenotype and the rater-specific part (mother and father). The sum of these contributions gives the total heritability and environmental contribution to the variance in maternal and paternal ratings of A/D. The last column in Table 5 also summarizes the percentage of variance in these ratings explained by the commonly agreed upon phenotype. For both maternal and paternal ratings, the total heritabilities decrease with increasing age. At age

Table 3: Twin correlations, cross-twin–cross-rater correlations and interparent correlations for the Child Behavior Checklist Anxious/Depression (age 3, 7, 10 and 12 years) and Devereux anxiety (age 5 years)

	Twin correlations		Cross-twin correlations		Inter-parent correlations	
	Mother	Father	mo/fa	fa/mo	Twin 1	Twin 2
Age 3						
mzm	0.69	0.67	0.46	0.49	0.65	0.67
dzm	0.32	0.36	0.21	0.22	0.67	0.67
mzf	0.72	0.72	0.54	0.53	0.68	0.67
dzf	0.33	0.43	0.24	0.25	0.67	0.68
dos_mf	0.37	0.38	0.22	0.23	0.65	0.60
dos_fm	0.31	0.36	0.22	0.21	0.69	0.70
Age 5						
mzm	0.73	0.72	0.47	0.42	0.60	0.57
dzm	0.45	0.43	0.27	0.24	0.61	0.60
mzf	0.73	0.72	0.43	0.43	0.56	0.58
dzf	0.42	0.50	0.27	0.27	0.59	0.59
dos_mf	0.51	0.55	0.29	0.29	0.58	0.55
dos_fm	0.47	0.52	0.29	0.26	0.56	0.60
Age 7						
mzm	0.61	0.62	0.37	0.38	0.61	0.55
dzm	0.37	0.39	0.20	0.23	0.60	0.61
mzf	0.63	0.64	0.42	0.37	0.60	0.58
dzf	0.40	0.39	0.17	0.21	0.57	0.58
dos_mf	0.41	0.41	0.20	0.23	0.57	0.53
dos_fm	0.39	0.37	0.21	0.23	0.58	0.58
Age 10						
mzm	0.60	0.61	0.33	0.37	0.55	0.59
dzm	0.38	0.42	0.19	0.28	0.60	0.63
mzf	0.66	0.68	0.43	0.37	0.57	0.54
dzf	0.43	0.50	0.33	0.33	0.67	0.64
dos_mf	0.37	0.31	0.18	0.14	0.63	0.50
dos_fm	0.48	0.45	0.30	0.26	0.63	0.61
Age 12						
mzm	0.60	0.61	0.35	0.40	0.60	0.57
dzm	0.38	0.41	0.21	0.27	0.65	0.65
mzf	0.60	0.63	0.36	0.35	0.53	0.60
dzf	0.39	0.46	0.24	0.29	0.62	0.62
dos_mf	0.45	0.39	0.23	0.24	0.54	0.58
dos_fm	0.52	0.47	0.24	0.37	0.61	0.56

Cross-twin–cross-rater correlations, oldest rated by mother (mo) and youngest by father (fa) and vice versa; dos_fm, opposite-sex twin pairs with female first-born twin and male second-born twin; dos_mf, opposite-sex twin pairs with male first-born twin and female second-born twin; dzf, dizygotic female twin pairs; dzm, dizygotic male twin pairs; interparent correlations, within subject for first-born (twin 1) and second-born twin (twin 2); mzf, monozygotic female twin pairs; mzm, monozygotic male twin pairs; twin correlations, oldest with youngest twin.

3, the total heritability is 69% for maternal ratings of A/D, and this decreases to 37% at age 12 years. A similar trend is present for father ratings. Table 6 summarizes the 'absolute' heritabilities for the common phenotype (proportion of the variance accounted for genetic and environmental factors is calculated by dividing them by the variance of the common parental part instead of dividing them by the total variance) and for the rater-specific phenotype (proportion of variance calculated by dividing by the variance of the unique parental part). For the common phenotype, the genetic influences decreased with increasing age. The variance accounted for by genetic factors decreased from 76% at age 3 to 48% at age 12. With increasing age, a slight increase of the shared environment was observed. At age 3, shared environment was not important, but at age 12, 18% of the variance was accounted by shared environmental influences. The

Table 4: Standardized estimates of additive genetic (a^2), common (c^2) and unique (e^2) environmental effects for Anxious/Depression based on separate analyses of maternal and paternal ratings

		Mother			Father		
		a^2	c^2	e^2	a^2	c^2	e^2
Age 3	Boys	69	0	31	67	1	31
	Girls	71	0	29	58	14	28
Age 5	Boys	52	21	26	50	24	26
	Girls	53	19	28	43	30	27
Age 7	Boys	51	12	37	50	12	38
	Girls	48	15	37	47	16	37
Age 10	Boys	47	16	37	52	11	37
	Girls	45	21	34	41	27	32
Age 12	Boys	40	22	38	45	19	36
	Girls	37	24	39	39	25	36

contribution of non-shared environment to the variance of anxiety ranged from 24 to 34%.

In further simplification of the model, it was tested whether the A and C factors of the rater-specific part could be dropped. Across all ages, the fit of the models deteriorated significantly when dropping the rater-specific A or C factor (Appendix 2). This emphasized that the rater-specific part represented not only measurement error but also reflected meaningful variation. The influence of the genetic and environmental factors of the rater-specific part could not be constrained across raters. The proportion of total variance of anxiety accounted by the rater-specific part differed across

Table 5: Standardized parameter estimates of genetic (a^2), common (c^2) and unique environmental (e^2) contributions to the variance of the common phenotype and to the rater-specific part (mother and father)

	Common phenotype			Rater-specific phenotype			Total			Total percentage of variance explained by the common phenotype
	a^2	c^2	e^2	a^2	c^2	e^2	a^2	c^2	e^2	
Age 3										
Mother	0.48	–	0.15	0.21	0.01	0.14	0.69	0.01	0.29	0.64
Father	0.53	–	0.17	0.06	0.11	0.13	0.59	0.11	0.30	0.7
Age 5										
Boys										
Mother	0.35	0.09	0.14	0.16	0.13	0.14	0.51	0.22	0.28	0.58
Father	0.36	0.1	0.14	0.11	0.15	0.14	0.47	0.25	0.28	0.6
Girls										
Mother	0.34	0.09	0.13	0.21	0.09	0.14	0.55	0.18	0.27	0.56
Father	0.35	0.1	0.14	0.09	0.19	0.13	0.44	0.29	0.28	0.58
Age 7										
Mother	0.32	–	0.16	0.21	0.11	0.21	0.53	0.11	0.37	0.47
Father	0.48	–	0.23	0.04	0.12	0.13	0.52	0.12	0.36	0.71
Age 10										
Boys										
Mother	0.29	0.03	0.17	0.18	0.13	0.21	0.47	0.16	0.37	0.48
Father	0.44	0.04	0.25	0.05	0.10	0.12	0.49	0.15	0.36	0.73
Girls										
Mother	0.26	0.09	0.13	0.18	0.12	0.20	0.44	0.21	0.34	0.49
Father	0.39	0.14	0.2	0.05	0.1	0.12	0.44	0.24	0.32	0.73
Age 12										
Mother	0.24	0.09	0.17	0.13	0.15	0.21	0.37	0.24	0.38	0.51
Father	0.33	0.12	0.24	0.09	0.09	0.12	0.42	0.21	0.36	0.7

Table 6: Standardized parameter estimates for rater-specific assessments of Anxious/Depression and for the phenotype agreed upon by both the parents (between the brackets the 95% confidence interval)

	Rater-specific phenotype			Common phenotype		
	a^2	c^2	e^2	a^2	c^2	e^2
Age 3						
Mother	0.59	0.02	0.39	0.76 (0.74–0.78)		0.24 (0.22–0.26)
Father	0.19	0.37	0.44			
Age 5						
Boys						
Mother	0.38	0.30	0.32	0.60 (0.54–0.65)	0.16 (0.14–0.22)	0.24 (0.22–0.26)
Father	0.27	0.38	0.35			
Girls						
Mother	0.47	0.21	0.32	0.60 (0.54–0.65)	0.16 (0.14–0.22)	0.24 (0.22–0.26)
Father	0.22	0.47	0.31			
Age 7						
Mother	0.39	0.22	0.39	0.67 (0.64–0.70)		0.33 (0.30–0.36)
Father	0.14	0.41	0.45			
Age 10						
Boys						
Mother	0.36	0.24	0.40	0.60 (0.45–0.69)	0.05 (0.001–0.18)	0.35 (0.30–0.41)
Father	0.19	0.37	0.44			
Girls						
Mother	0.36	0.24	0.40	0.53 (0.36–0.70)	0.20 (0.04–0.35)	0.27 (0.23–0.32)
Father	0.19	0.37	0.44			
Age 12						
Mother	0.27	0.30	0.43	0.48 (0.33–0.63)	0.18 (0.05–0.30)	0.34 (0.29–0.40)
Father	0.29	0.29	0.42			

fathers and mothers. The rater-specific part of the fathers accounted for between 27 and 42% of the total variance and showed a reasonable stable pattern across age. For the mothers the rater-specific part accounted for 36–53% of the total variance and became larger with increasing age. The rater-specific genetic influences accounted for 4–21% of the variance and appeared to be larger for mother ratings. The rater-specific C factor, represented also rater bias, accounted for 1–19% of the variance. At younger ages, the influence of the rater-specific C factor seemed to be larger for father ratings than for mother ratings. After age 5, the rater-specific C factors seemed to be equal for mothers and fathers. The rater-specific non-shared environmental factors (including measurement error) accounted for 12–21% of the variance of anxiety across ages.

Sex differences in genetic and environmental effects were found for the rater specific part at age 5 and for the common parental view at age 10. At age 5, the mother-specific genetic factor explained 16 and 21% of the variance for boys and girls, respectively. The influence of mother-specific C factor was larger for boys than girls. When father rated the behavior, the specific C factor was larger for girls than boys. At age 10, the contribution of shared environmental factors to the variance of the common latent phenotype was larger for girls (20%) than for boys (5%).

Discussion

This report provides a unique perspective for future studies of A/D in children. First, it is clear that at ages 3 through 12 years, A/D is influenced to a large extent by genetic factors. The size of these genetic influences changes with age. Heritability estimates for A/D as agreed upon by both the parents were high in 3-year olds (76%) and decreased in size as children grew older (60% at age 5, 67% at age 7, 53% at age 10 (60% in boys) and 48% at age 12 years). The decrease in heritability estimates was accompanied by an increase in the amount of variance that could be ascribed to the influence of the common family environment shared by children growing up in the same family (absent at ages 3 and 7, 16% at age 5, 20% at age 10 (5% in boys) and 18% at age 12 years). We observed hardly any sex differences in the size of the heritability estimates. Only at age 10, the heritability of rater-independent A/D was somewhat lower in girls than in boys, while the environment shared by girls had a larger influence on A/D variation that the family environment shared by boys. An important finding was that dizygotic twins of opposite sex resembled each other to same extent as same-sex dizygotic twins, suggesting that the same genes and shared environmental effects are expressed in boys and girls. Interesting informant effects were found,

with agreement between mothers and fathers high at young ages ($r = 0.70$ at age 3 years) and diminishing by age 12 ($r = 0.50$). Disagreement in ratings between the parents was not merely the result of unreliability or rater bias. Both the parents provided unique information from their own perspective on their children. Genetic and shared environmental influences were found for the unique parental views at all ages. At all ages, the contribution of shared environment, to variation in rater-specific views, is higher for father ratings. Also, at all ages except age 12, the heritability estimates for the rater-specific phenotype were higher for mother ratings (59% at age 3 and decreasing to 27% at age 12 years) than for father ratings (between 14 and 29%). These data can perhaps be seen as reflecting the changing relationships between children and their mothers and fathers and as demonstrating the importance of rater-specific information across development.

Our data indicate that at early ages, genetic factors account for the majority of the variation in A/D. Yet as children grow up, shared environment becomes more important. These findings may be compared with the analyses of attention problems and aggressive behavior in these twins (Hudziak *et al*. 2003; Rietveld *et al*. 2003b; 2004; Van Beijsterveldt *et al*. 2003) which showed high heritability estimates at all ages and no evidence of any contribution of shared environment for attention problems. For aggression, we observed an increase in heritability from age 3 to age 10 and a decrease in the percentages of variance explained by shared environment. The finding of the importance of shared environment as a major influence on A/D may provide a reason for why certain family therapeutic approaches are effective in the A/D spectrum of illnesses.

The demonstration of the heritability of A/D during childhood allows for a definitive rejection of the radical environmentalist position, which asserts that the clustering of disorders within families is evidence for the importance of environmental familial risk factors (Kendler 2005). A limitation of the genetic epidemiological approach followed in our study is that we can only estimate the effects of all genetic factors and not identify single gene variants. Another limitation is the use of three different checklists at different ages. The use of different instruments across ages may limit the comparison of the heritability between ages, and it is strictly speaking not clear whether the changes in results are due to developmental changes or due to the use of different instruments. However, the trend of a decreasing heritability is observed not only between ages 3 and 7 (different instruments) but also between ages 7 and 12 years (same instruments), which suggests the decrease is genuine and not merely a reflection of a change in instruments.

The animal literature provides us with some suggestions of genes that may affect A/D across development. Gross *et al*. (2002) demonstrated that by turning on the *5-HT1AR* gene in a mouse knockout, mice that had the gene activated only in the perinatal period developed normal anxiety-like behavior as adults. However, if the gene was inactivated during this critical developmental period (regardless of whether it was reactivated in adulthood), the mouse demonstrated behavioral inhibition. The absence of the serotonin-1A receptor thus has developmental effects on behaviors that are irreversible in adulthood, even after the receptor has been restored to normal levels (Freedman 2002). Common psychiatric illnesses, such as emotional and behavioral disorders, are likely to be influenced by multiple genes, of which the *5-HT1AR* gene is only one candidate. Some significant genetic associations have been reported for anxiety and depressive disorders in children and young adults (e.g. Caspi *et al*. 2003; Eley *et al*. 2004; Kaufman *et al*. 2004), yet confirmation of these findings remains in doubt (e.g. Young *et al*. 2003). Although a great deal of optimism exists that through these findings, an increase in the understanding of genetic mechanisms will lead to new insights into pathophysiology and thus treatment (Moldin 2000), an enduring respect for the complexity of these conditions is also realized. In many studies, the results are of small or modest effect size and are germane to only one sex or to one age group. However, the solid heritability findings as obtained in our study of large samples point to the existence of genes influencing A/D. Taken together, these findings argue for a developmentally sensitive approach to gene-hunting expeditions. In order to realize the full promise of the era of Genomic Medicine, it will be necessary for us to approach complex phenotypes with complex solutions. The study of A/D across the lifespan is one such example.

References

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

Achenbach, T.M. (1992) *Manual for the Child Behavior Checklist/2–3 and 1992 Profile*. University of Vermont, Department of Psychiatry, Burlington, VT.

Akaike, H. (1987) Factor analysis and AIC. *Psychometrika* **52**, 317–332.

Bartels, M., Hudziak, J.J., Van den Oord, E.J., Van Beijsterveldt, C.E., Rietveld, M.J. & Boomsma, D.I. (2003a) Co-occurrence of aggressive behavior and rule-breaking behavior at age 12: multi-rater analyses. *Behav Genet* **33**, 607–621.

Bartels, M., Hudziak, J.J., Boomsma, D.I., Rietveld, M.J., Van Beijsterveldt, T.C. & Van den Oord, E.J. (2003b) 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., Boomsma, D.I., Rietveld, M.J., Van Beijsterveldt, C.E., Hudziak, J.J. & Van den Oord, E.J. (2004) Disentangling genetic environmental, and rater effects on internalizing and externalizing problem behavior in 10-year-old twins. *Twin Res* **7**, 162–175.

Boomsma, D.I., Busjahn, A. & Peltonen, L. (2002a) The classical twin study and beyond. *Nat Rev Genet* **3**, 872–882.

Boomsma, D.I., Vink, J.M., Van Beijsterveldt, T.C., de Geus, E.J., Beem, A.L., Mulder, E.J., Derks, E.M., Riese, H., Willemsen, G.A.,

Bartels, M., Van den Berg, M., Kupper, N.H., Polderman, T.J., Posthuma, D., Rietveld, M.J., Stubbe, J.H., Knol, L.I., Stroet, T. & van Baal, G.C. (2002b) Netherlands Twin Register: a focus on longitudinal research. *Twin Res* **5**, 401–406.

Caspi, A., Sugden, K., Moffitt, T.E., Taylor, A., Craig, I.W., Harrington, H., McClay, J., Mill, J., Martin, J., Braithwaite, A. & Poulton, R. (2003) Influence of life stress on depression: moderation by a polymorphism in the 5-HTT gene. *Science* **301**, 386–389.

Collins, F.S. (2004) The case for a US prospective cohort study of genes and environment. *Nature* **429**, 475–477.

Collins, F.S., Green, E.D., Guttmacher, A.E. & Guyer, M.S.; US National Human Genome Research Institute (2003) A vision for the future of genomics research. *Nature* **422**, 835–847.

Derks, E.M., Hudziak, J.J., Van Beijsterveldt, C.E., Dolan, C.V. & Boomsma, D.I. (2004) A study of genetic and environmental influences on maternal and paternal CBCL syndrome scores in a large sample of 3-year-old Dutch twins. *Behav Genet* **34**, 571–583.

Eaves, L.J., Silberg, J.L., Meyer, J.M., Maes, H.H., Simonoff, E., Pickles, A., Rutter, M., Neale, M.C., Reynolds, C.A., Erikson, M.T., Heath, A.C., Loeber, R., Truett, K.R. & Hewitt, J.K. (1997) Genetics and developmental psychopathology: 2. 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. & Thompson, L.A. (1995) A twin study of competence and problem behavior in childhood and early adolescence. *J Child Psychol Psychiatry* **36**, 775–785.

Eley, T.C., Deater-Deckard, K., Fombonne, E., Fulker, D.W. & Plomin, R. (1998) An adoption study of depressive symptoms in middle childhood. *J Child Psychol Psychiatry* **39**, 337–345.

Eley, T.C., Bolton, D., O'Connor, T.G., Perrin, S., Smith, P. & Plomin, R. (2003) A twin study of anxiety-related behaviours in pre-school children. *J Child Psychol Psychiatry* **44**, 945–960.

Eley, T.C., Sugden, K., Corsico, A., Gregory, A.M., Sham, P., McGuffin, P., Plomin, R. & Craig, I.W. (2004) Gene–environment interaction analysis of serotonin system markers with adolescent depression. *Mol Psychiatry* **9**, 908–915.

Falconer, D.S. & Mackay, T.F.C. (1996) *Introduction to Quantitative Genetics*, 4th edn. Longman, London.

Freedman, R. (2002) Long-term effects of early genetic influences on behavior. *N Engl J Med* **347**, 213–215.

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

Gross, C., Zhuang, X., Stark, K., Ramboz, S., Oosting, R., Kirby, L., Santarelli, L., Beck, S. & Hen, R. (2002) Serotonin1A receptor acts during development to establish normal anxiety-like behaviour in the adult. *Nature* **416**, 396–400.

Hay, D.F., Pawlby, S., Sharp, D., Schmucker, G., Mills, A., Allen, H. & Kumar, R. (1999) Parents' judgements about young children's problems: why mothers and fathers might disagree yet still predict later outcomes. *J Child Psychol Psychiatry* **40**, 1249–1258.

Hewitt, J.K., Silberg, J.L., Neale, M.C., Eaves, L.J. & Erickson, M. (1992) The analysis of parental ratings of children's behavior using LISREL. *Behav Genet* **22**, 293–317.

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

Hudziak, J.J., van Beijsterveldt, C.E., Bartels, M., Rietveld, M.J., Rettew, D.C., Derks, E.M. & 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**, 575–589.

Jöreskog, K.D. & Sörbom, D. (1996) *PRELIS 2 User's Reference Guide*. Scientific Software International, Inc., Chicago.

Kaufman, J., Yang, B.Z., Douglas-Palumberi, H., Houshyar, S., Lipschitz, D., Krystal, J.H. & Gelernter, J. (2004) Social supports and serotonin transporter gene moderate depression in maltreated children. *Proc Natl Acad Sci USA* **101**, 17316–11721.

Kendler, K.S. (2005) Psychiatric genetics: a methodologic critique. *Am J Psychiatry* **162**, 3–11.

Koot, H.M., Van Den Oord, E.J., Verhulst, F.C. & Boomsma, D.I. (1997) Behavioral and emotional problems in young preschoolers: cross-cultural testing of the validity of the Child Behavior Checklist/2–3. *J Abnorm Child Psychol* **25**, 183–196.

Moldin, S.O. (2000) Neurobiology of anxiety and fear: challenges for genomic science of the new millennium. *Biol Psychiatry* **48**, 1144–1146.

Neale, M.C. & Stevenson, J. (1989) Rater bias in the EASI temperament scales: a twin study. *J Pers Soc Psychol* **56**, 446–455.

Neale, M.C., Boker, S.M., Xie, G. & Maes, H.H. (1999) *Mx; Statistical Modeling*, 5th edn. Department of Psychiatry, Richmond, VA.

Posthuma, D. & Boomsma, D.I. (2000) A note on the statistical power in extended twin designs. *Behav Genet* **30**, 147–158.

Posthuma, D., de Geus, E.J.C. & Boomsma, D.I. (2002) Genetic contributions to anatomical, behavioral and neurophysiological indices of cognition. In Plomin, R., DeFries, J.C., Craig, I.W. & McGuffin, P. (eds), *Behavioral Genetics in the Postgenomic Era*. APA Books, Washington DC, pp. 141–161.

Rice, F., Harold, G. & Thapar, A. (2002a) The genetic aetiology of childhood depression: a review. *J Child Psychol Psychiatry* **43**, 65–79.

Rice, F., Harold, G.T. & Thapar, A. (2002b) Assessing the effects of age, sex and shared environment on the genetic aetiology of depression in childhood and adolescence. *J Child Psychol Psychiatry* **43**, 1039–1051.

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

Rietveld, M.J., Dolan, C.V., Van Baal, G.C. & Boomsma, D.I. (2003a) A twin study of differentiation of cognitive abilities in childhood. *Behav Genet* **33**, 367–381.

Rietveld, M.J., Hudziak, J.J., Bartels, M., Van Beijsterveldt, C.E. & Boomsma, D.I. (2003b) 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* **117**, 102–113.

Rietveld, M.J., Hudziak, J.J., Bartels, M., Van Beijsterveldt, C.E. & Boomsma, D.I. (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.

Rutter, M. (2003) Commentary: Nature-nurture interplay in emotional disorders. *J Child Psychol Psychiatry* **44**, 934–944.

Rutter, M. & Silberg, J. (2002) Gene-environment interplay in relation to emotional and behavioral disturbance. *Annu Rev Psychol* **53**, 463–490.

Schmitz, S., Fulker, D.W. & Mrazek, D.A. (1995) Problem behavior in early and middle childhood: an initial behavior genetic analysis. *J Child Psychol Psychiatry* **36**, 1443–1458.

Spivack, G. & Spotts, J. (1966) *The Devereux Child Behavior (DCB) Rating Scale*. The Devereux Foundation, Devon, PA.

SPSS (2001) *SPSS for Windows, Rel 11.0.1*. SPSS, Chicago.

Todd, R.D. & Botteron, K.N. (2002) Etiology and genetics of early-onset mood disorders. *Child Adolesc Psychiatr Clin N Am* **11**, 499–518.

Van Beijsterveldt, C.E.M., Hudziak, J.J., Bartels, M. & Boomsma, D.I. (2003) Causes of stability of aggression from early childhood to adolescence: a longitudinal genetic analysis in Dutch twins. *Behav Genet* **34**, 229–241.

Van Beijsterveldt, C.E., Verhulst, F.C., Molenaar, P.C. & Boomsma, D.I. (2004) The genetic basis of problem behavior in 5-year-old Dutch twin pairs. *Behav Genet* **34**, 229–242.

Van den Oord, E.J., Boomsma, D.I. & Verhulst, F.C. (1994) A study of problem behaviors in 10- to 15-year-old biologically related and unrelated international adoptees. *Behav Genet* **24**, 193–205.

Van den Oord, E.J., Verhulst, F.C. & 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.J., Verhulst, F.C. & 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., Van den Oord, E.J., Verhulst, F.C. & Boomsma, D.I. (2003) Using shared and unique parental views to study the etiology of 7-year-old twins' internalizing and externalizing problems. *Behav Genet* **33**, 409–420.

Verhulst, F.C., Van der Ende, J. & Koot, H.M. (1996) *Handleiding Voor de CBCL/4–18*. Afdeling Kinder-en Jeugdpsychiatrie, Sophia Kinderziekenhuis/Academisch Ziekenhuis Rotterdam/ Erasmus Universiteit Rotterdam, Rotterdam.

Wamboldt, M.Z. & Wamboldt, F.S. (2000) Role of the family in the onset and outcome of childhood disorders: selected research findings. *J Am Acad Child Adolesc Psychiatry* **39**, 1212–1219.

Young, S.E., Smolen, A., Stallings, M.C., Corley, R.P. & Hewitt, J.K. (2003) Sibling-based association analyses of the serotonin transporter polymorphism and internalizing behavior problems in children. *J Child Psychol Psychiatry* **44**, 961–967.

Acknowledgments

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

Appendix 1

A/D items completed at each age by parents of twins

Age 3 (Koot *et al.* 1997)	Age 5 (Van Beijsterveldt *et al.* 2004)	Ages 7, 10 and 12 (Verhulst *et al.* 1996)
Afraids to try new	Concerns about physical health	Lonely
Avoids eye contact	Less tolerance for pain	Cries a lot
Clings to adults (A)	Looks unhappy	Fears doing bad
Disturbed by change	Over-excitable	Must bed perfect
Upset by separation (A)	Expresses fears	Feels unloved
Self-conscious (A)	Wakes-up in the night	Feels persecuted
Shy (A)		Feels worthless
Too fearful/anxious (A)		Nervous, tense
Upset by new		Fearful, anxious
		Feels too guilty
		Self-conscious
		Suspicious
		Unhappy, sad
		Worries

(A), items overlapping with anxious/depression scale from American manual (Achenbach 1991).

Appendix 2

Summary of model fitting results of simultaneous analysis of paternal and maternal ratings of Anxious/Depression at ages 3, 5, 7, 10 and 12 years

	−2 LL	Number of parameters	df	$\triangle X^2$	$\triangle$df	*P*	AIC
Age 3, overall model							
Saturated model	140166.88	84	29931				
Rater bias model	140386.06	30	29985	155.27	54	0.00	47.27
Psychometric model	140235.16	34	29981	68.28	50	0.04	−31.72
Simplification best overall model							
No common genetic factors	140745.95	32	29983	510.79	2	0.00	506.79
No common shared E	140235.42	32	29983	0.26	2	0.88	−3.74
No common non-shared E	141109.89	32	29983	874.73	2	0.00	870.73
No unique shared E	140288.29	30	29985	53.13	4	0.00	45.13
No unique genetic factors	140386.06	30	29985	150.90	4	0.00	142.90
No sex dif							
No sex dif common factors	140242.25	31	29984	7.09	3	0.07	1.09
No sex dif unique factors	140243.55	28	29987	8.39	6	0.21	−3.61
No sex dif common and unique factors	140249.17	25	29990	14.01	9	0.12	−3.99
No rater dif unique views	140319.36	28	29987	84.20	6	0.00	72.20
Age 5, overall model:							
Saturated model	169695.49	84	34996				
Rater bias estimates	169975.17	30	35050	279.84	54	0.00	171.84
Psychometric model	169790.33	34	35046	94.84	50	0.00	−5.32
Simplification best overall model							
No common genetic factors	170122.11	32	35048	331.78	2	0.00	327.78
No common shared E	169821.20	32	35048	30.87	2	0.00	26.87
No common non-shared E	170775.51	32	35048	985.18	2	0.00	981.18
No unique shared E	170015.90	30	35050	225.57	4	0.00	217.57
No unique genetic factors	169975.17	30	35050	185.00	4	0.00	177.00
No sex dif							
No sex dif common factors	169793.64	31	35049	3.31	3	0.34	−2.69
No sex dif unique factors	169805.03	28	35052	14.70	6	0.02	2.70
No sex dif common and unique factors	169807.00	25	35055	16.67	9	0.05	−1.33
No rater dif unique views	169822.29	28	35052	31.96	6	0.00	19.96
Age 7, overall model							
Saturated model	112764.49	84	25558				
Rater bias estimates	112984.50	30	25612	220.01	54	0.00	112.01
Psychometric model	112893.00	34	25608	128.51	50	0.00	28.51
Simplification best overall model:							
No common genetic factors	113056.07	32	25610	163.07	2	0.00	159.07
No common shared E	112895.86	32	25610	2.86	2	0.24	−1.14
No common non-shared E	113818.62	32	25608	925.62	2	0.00	921.62
No unique shared E	112969.92	30	25612	76.92	4	0.00	68.92
No unique genetic factors	112984.50	30	25612	91.50	4	0	83.50
No sex dif							
No sex dif common factors	112893.77	31	35049	0.77	3	0.86	−5.23
No sex dif unique factors	112900.18	28	35052	7.18	6	0.30	−4.82
No sex dif common + unique factors	112900.51	25	35055	7.51	9	0.58	−10.49
No rater dif unique views	113627.38	28	35052	734.38	6	0.00	722.38
Age 10, overall model							
Saturated model	71963.51	84	15244				
Rater bias estimates	72097.31	30	15298	133.80	54	0.00	25.80
Psychometric model	72049.99	34	15294	86.48	50	0.00	−13.52
Simplification best overall model							
No common genetic factors	72150.65	32	15296	100.66	2	0.00	96.66

Appendix 2. Continued.

No common shared E	72055.73	32	15296	5.74	2	0.06	1.74
No common non-shared E	72586.21	32	15296	536.22	2	0.00	532.22
No unique shared E	72098.46	30	15298	48.47	4	0.00	40.47
No unique genetic factors	72097.31	30	15298	47.32	4	0.00	39.32
No sex dif							
No sex dif common factors	72059.41	31	15297	9.42	3	0.02	3.42
No sex dif unique factors	72055.07	28	15300	5.08	6	0.53	−6.92
No sex dif common and unique factors	72063.87	25	15303	13.88	9	0.12	−4.12
No rater dif unique views	72477.12	28	15300	427.13	6	0.00	415.13
Age 12, overall model							
Saturated model	38126.93	84	8269				
Rater bias estimates	38221.13	30	8323	94.20	54	0.00	−13.80
Psychometric model	38202.30	34	8319	75.37	50	0.01	−24.63
Simplification best overall model							
No common genetic factors	38237.81	32	8321	35.51	2	0.00	31.51
No common shared E	38209.40	32	8321	7.10	2	0.03	3.10
No common non-shared E	38519.71	32	8321	317.41	2	0.00	313.41
No unique shared E	38233.44	30	8323	31.14	4	0.00	23.14
No unique genetic factors	38221.13	30	8323	18.83	4	0.00	10.83
No sex dif							
No sex dif common factors	38203.51	31	8322	1.21	3	0.75	−4.79
No sex dif unique factors	38205.17	28	8325	2.87	6	0.82	−9.13
No sex dif common + unique factors	38207.41	25	8328	5.11	9	0.82	−12.89
No rater dif unique views	38364.20	28	8325	161.90	6	0.00	149.90

−2 LL, −2-log likelihood; AIC, Akaike's information criterion; df, degrees of freedom; dif, differences; E, environment.