

Genetic Influences on Thought Problems in 7-Year-Olds: A Twin-Study of Genetic, Environmental and Rater Effects

Abdel Abdellaoui,¹ Meike Bartels,¹ James J. Hudziak,^{2,1} Patrizia Rizzu,³ Toos CEM van Beijsterveldt,¹ and Dorret I. Boomsma¹

¹ Netherlands Twin Register, Department of Biological Psychology, VU University, Amsterdam, the Netherlands

² Departments of Psychiatry and Medicine (Division of Human Genetics), Center for Children, Youth and Families, University of Vermont, College of Medicine Burlington, Burlington, Vermont, United States of America

³ Department of Human Genetics, Section Medical Genomics, VU University, Amsterdam, the Netherlands

The Thought–Problem scale (TP) of the CBCL assesses symptoms such as hallucinations and strange thoughts/behaviors and has been associated with other behavioral disorders. This study uses parental reports to examine the etiology of variation in TP, about which relatively little is known, in 7-year-old twins. Parental ratings on TP were collected in 8,962 7-year-old twin pairs. Because the distribution of TP scores was highly skewed scores were categorized into 3 classes. The data were analyzed under a threshold liability model with genetic structural equation modeling. Ratings from both parents were simultaneously analyzed to determine the rater agreement phenotype (or common phenotype [TPc]) and the rater specific phenotype [TPs] that represents rater disagreement caused by rater bias, measurement error and/or a unique view of the parents on the child's behavior. Scores on the TP-scale varied as a function of rater (fathers rated fewer problems), sex (boys scored higher) and zygosity (DZ twins scored higher). The TPc explained 67% of the total variance in the parental ratings. Variation in TPc was influenced mainly by the children's genotype (76%). Variance in TPs also showed a contribution of genetic factors (maternal reports: 61%, paternal reports: 65%), indicating that TPs does not only represent rater bias. Shared environmental influences were only found in the TPs. No sex differences in genetic architecture were observed. These results indicate an important contribution of genetic factors to thought problems in children as young as 7 years.

Keywords: thought problems, parental rating, twins, heritability

The Thought–Problems syndrome (TP) of the Child Behavior Checklist (CBCL; Achenbach, 1991) is an empirically derived set of items that cover symptoms such as hallucinations, obsessive–compulsive symptoms and strange thoughts and behaviors. A strong statistical relation between the TP-items was seen in

the factor analysis through which the scale was derived (Achenbach, 1991). The internal consistency of the TP-scale was shown to be sufficient in the CBCL (Kuo et al., 2004; Lin et al., 2006). Compared to the other empirically derived CBCL scales, the TP-scale has received relatively little attention. In a longitudinal study of Ferdinand et al. (2001), thought problems in childhood were associated with substance abuse (alcohol and tobacco) in young adulthood. Other studies showed TP in children to be associated with disorders, such as obsessive–compulsive disorder (OCD; Geller et al., 2004), multiple complex developmental disorder (MCDD; de Bruin et al., 2006) and fragile X syndrome (Hessel et al., 2006). The TP-syndrome also predicted DSM-III-R diagnoses of simple phobia, social phobia, separation anxiety disorder, mood disorders and psychotic disorders (Kasius et al., 1997). There have been a few studies on the heritability of the TP-scale in young twins that found evidence for significant heritability (Edelbrock et al., 1995; Kuo et al., 2004; Lin et al., 2006; Polderman et al., 2006; Schmitz et al., 1995). The aim of this study is to obtain more insight into the genetic and environmental contributions to variation in TP in a large group of 7-year-old twins. The rating of TP in young children might be complicated and we obtained ratings from both parents and modeled the extent to which both parents agreed on the presence or absence of TP.

Parental ratings, such as those collected with the CBCL, provide meaningful information about a child's behavioral and emotional problems. Parents observe their children for long periods and in natural situations, which makes them ideal informants regarding their children's behavior. Studies that collect information

Received 29 August, 2008; accepted 8 September, 2008.

Address for correspondence: A. Abdellaoui, Department of Biological Psychology, VU University Amsterdam, Van der Boechorststraat 1, 1081 BT Amsterdam, The Netherlands. E-mail: a.abdellaoui@psy.vu.nl

from both parents tend to show that parents agree for a substantial part in their assessment of behavioral problems in their children. An analysis by Achenbach et al. (1979) showed a correlation of 0.60 between the paternal and maternal ratings of the same child. Although correlations are high, they are not perfect. To explore the processes underlying disagreements in parental ratings, Hewitt et al. (1992) developed a series of models, in which the disagreement of the parents' reports reflect rater bias and a unique accurately assessed part of the child's behavior.

Rater bias may be caused by the parents' own characteristics (a projection bias), or by parents' response biases (e.g., stereotyping, employing different normative standards or having certain response styles, i.e., judging certain types of behavior more or less severely). The unique views on the phenotype may arise firstly because parents observe their child in distinct situations and environments. For instance, the mother could have the task of bringing the child to school, while the father is the one who accompanies the child to soccer games. Second, the way the mother interacts with the child may differ from the way the father interacts with the child (Achenbach et al., 1979). Another cause for different views in parent reports could be that men are more sensitive to different kinds of input from their children than women (Baron-Cohen & Hammer, 1997; Connelan et al., 2000; Rosenthal et al., 1979).

A number of quantitative genetic studies have examined problem behavior in children with the CBCL, with the main focus on internalizing and externalizing behavior problems (Bartels et al., 2004a; Bartels et al., 2004b; Bartels et al., 2007; Edelbrock et al., 1995; Gjone & Stevenson, 1997; Hudziak et al., 2000; Leve et al., 1998; Schmitz & Mrazek, 2001; Silberg et al., 1994; Van den Oord et al., 1995; Van der Valk et al., 1998; Zahn-Waxler et al., 1996) and a number of these studies have modeled the agreement and disagreements among parental reports. For example, in a study by Rowe and Kandel (1997), parents rated the internalizing and externalizing behavior of their two oldest children (between 9 and 17 years old). 'Individual view' and 'shared view' models were used. The parents assessed similar aspects of the child's behavior, and in addition the mother and father ratings contained a significant individual view component.

Twin studies provide the opportunity to analyze whether the variance of the unique view of the parental assessments can be explained by genetic factors. If genetic factors are found, this implies that the unique view of the parents partly represents real behavior of the child. This rules out the possibility of the rater disagreements containing only rater bias. Hewitt et al. (1992) studied internalizing behavior in prepubertal (8–11 years) and pubertal twins (12–16 years), and found evidence for such genetic effects. Dutch twin studies on internalizing and externalizing

behaviors also found that the unique viewpoint of parents does not solely reflect rater bias (Bartels et al., 2003; Bartels et al., 2004a; Bartels et al., 2007; Van der Valk et al., 2001; Van der Valk et al., 2003).

When data are derived from questionnaires developed to measure the degree of dysfunctional behavior such as the CBCL, a large degree of skewness is often observed. These non-normal distributions can be explained by the fact that in symptom data a majority of the subjects displays few or no symptoms (Van den Oord et al., 2003). This is especially the case for the TP subscale, which is one of the subscales with the lowest mean scores in general population samples (Achenbach, 1991). Logarithmic and square root transformations are often not enough to correct for this non-normality. Categorizing the observations and analyzing the data with a threshold model has shown to be a successful way to decrease bias in parameter estimates (Derks et al., 2004). A disadvantage of categorizing the data may be that it reduces the statistical power of the analysis, and therefore large sample sizes are required.

This study estimates the genetic and environmental influences on variation in thought problems in a large group of 7-year-old Dutch twin pairs, while taking the agreement and disagreements between the parent reports into account. To overcome biased estimates due to skewness of the phenotype, the continuous TP scale was categorized into three classes (low, middle, high) and analyzed with threshold models.

Methods

Subjects

The participants were all registered with the Netherlands Twin Registry (NTR), which was established by the Department of Biological Psychology at the Vrije Universiteit in Amsterdam (Bartels et al., 2007; Boomsma et al., 2006). Parents of young twins receive questionnaires when their twins are 1, 2, 3, 5, 7, 10 and 12 years old. For this study, data of 7-year-old twin pairs from birth cohorts 1986–1997 were used. The questionnaires were mailed to the parents within 3 months of the twins' 7th birthday. Reminders were sent after 2 to 3 months. The response rate was 62% ($N = 8962$). Reasons for families not participating at a particular wave of data collection vary; families may request to not take part in the research at a particular age due to various reasons. Also, families may move to new addresses without notifying the NTR staff. Bartels et al. (2007) showed that the drop-out was largely random according to the definition of Little & Rubin (1987), making generalizations of the results more valid.

Zygoty was determined for 1492 same-sex twin-pairs by blood group ($n = 389$) or DNA polymorphisms ($n = 1103$). The zygoty of the other same-sex twins was determined using a discriminant analysis of questionnaire items answered by the parents. The

questionnaire led to correct classification of the zygosity in about 93 % of the cases (Rietveld et al., 2000).

The sample contained 1466 monozygotic male (MZM), 1516 dizygotic male (DZM), 1675 monozygotic female (MZF), 1445 dizygotic female (DZF), and 2860 dizygotic opposite-sex (DOS) twin pairs. Response rates for mothers were higher than for fathers and the group could be further divided into twin pairs for which both mother and father had replied (1053 MZM, 1105 DZM, 1226 MZF, 999 DZF, 2002 DOS), pairs with only mother-reports (395 MZM, 392 DZM, 423 MZF, 424 DZF, 805 DOS) and pairs with only father-reports ($n = 113$).

Measures

The Child Behavior Checklist (CBCL 4-18; Achenbach, 1991) was used to assess emotional and behavioral problems including thought problems. The questionnaire consists of 113 items, and measures behavior during the preceding 6 months. There are seven TP items: 'Can't get his/her mind off certain thoughts', 'Repeats certain acts over and over', 'Strange behavior', 'Strange ideas', 'Hears sounds or voices that aren't there', 'Sees things that aren't there', 'Stares blankly'. The reliability of the CBCL has been confirmed by in Dutch epidemiological samples. The 2-week test-retest correlation for TP was 0.74 (Verhulst et al., 1996). The TP scores were transformed from continuous to categorical data with two thresholds to limit the number of categories to three (low, middle and high levels of TP). The thresholds were chosen in such a way that there were no empty cells in the contingency tables of twin 1 versus twin 2 scores.

Genetic Analyses

Data from monozygotic (MZ) and dizygotic (DZ) twins were used to decompose the variation in the liability of thought problems into a contribution of additive genetic, shared environmental and non-shared environmental components. The categorical trait thought problems were modeled to have an unobserved, underlying continuous distribution with two thresholds that divide the distribution into three categories (low, middle, high). Such underlying distributions have been termed the liability or vulnerability (Falconer, 1989). The continuous variation in liability may be genetic or environmental in origin. The mean and variance of the liability distribution were standardized with mean zero and unit variance.

The additive genetic influence (A) on the variation in liability represents the sum of allelic influences at each locus in the genome contributing to the phenotype. The environmental influences can be shared and nonshared. Shared environment is common (C) to both twins growing up in the same family. Nonshared environmental influences (E) are unique to each twin and do not lead to twin resemblance.

Model Fitting

With multiple raters (two parents in this study), the variance of the liability distribution can be distinguished

into two parts that represent agreement (commonly rated phenotype by both parents) and disagreement (unique parental views). Both parts of the variance can be influenced by genetic (A), shared (C) and nonshared (E) factors and not necessarily to the same degree. We use the psychometric model (Hewitt et al., 1992) as presented in Figure 1. All variables enclosed in circles are latent (unobserved) that influence observed traits (enclosed in rectangles). The influence of latent variables on other latent variables or on observed traits is given by factor loadings (a), (c), and (e). These factor loadings come with (f) or (m) if the latent variable influences a unique view of the father or the mother.

Mx (Neale et al., 2006) was used to obtain parameter estimates for thresholds and factor loadings. Genetic models were fitted to the raw data with maximum likelihood estimation procedures. In a saturated model thresholds and polychoric twin correlations were estimated separately for MZM, DZM, MZF, DZF, DOSMF and DOSFM groups. This model was also used to test whether the thresholds could be constrained to be equal for mother and father ratings, for MZ and DZ-twins, for boys and girls and for the youngest and the oldest twin. Next, the fit of the psychometric model was compared with the fit of the saturated model.

The psychometric model was used to test for sex differences in factor loadings. Next, the significant contribution of the common and unique A and C variance components was tested. The significance of the unique 'A' component can show whether the rater specific parts represent not only measurement error and rater bias, but also reflects meaningful variation. Significance of the estimates was established by comparing the full model with a simplified model. The more parsimonious nested model is chosen over the

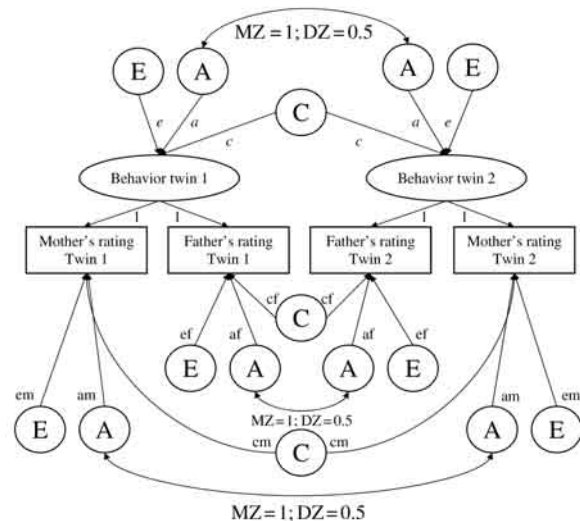


Figure 1
The psychometric-model.

full model when the analysis shows a low nonsignificant χ^2 test statistic ($p > .05$). In addition to the χ^2 test statistic, Akaike's Information Criterion ($AIC = \chi^2 - 2 \times \text{degrees of freedom}$) was computed. The lower the AIC, the better the fit of the model to the observed data.

Results

Description of the Data and Threshold Differences

For descriptive purposes, Table 1 summarizes the means and variances for the untransformed measures of Thought Problems by sex for mother and father reports.

Table 2 shows the thresholds from the saturated model in Mx. The thresholds are higher for all father ratings, indicating a tendency for mothers to give higher ratings (i.e., lower thresholds) for both boys and girls ($\chi^2 (24) = 84.19, p < .01$, see Table 4). The thresholds also indicate more TP (i.e., lower thresholds) for the DZ-

twins than for the MZ-twins ($\chi^2 (16) = 29.75, p = .02$). Finally, Table 2 shows higher thresholds for the liability for TP in girls than in boys (i.e. fewer thought problems for girls) ($\chi^2 (24) = 42.13, p = .01$). The differences between the thresholds of the youngest and oldest twin were not significant ($\chi^2 (24) = 32.35, p = .11$).

Correlations

From the saturated model polychoric correlations between twins, between raters and cross-twin cross-rater were obtained. Table 3 shows correlations between twins rated by the same parent in the first and second columns. The last four columns show the cross-correlations between twins each rated by a different parent. The interparent correlations were comparable for both first- and second-born twins. On average the interparent correlation is .66, which is in the same range as the parental agreement found in previous studies of about .60 (Achenbach, 1991).

Model Fitting Results for the Psychometric Model

The model fitting results of the simultaneous analysis of maternal and paternal ratings are summarized in Table 4. Sex differences in factor loadings were not detected ($\chi^2 (7) = 7.35, p = .60$). The common environmental effects on the common view of parents were not significant ($\chi^2 (1) = .972, p = .32$), but the contribution of C to the variance of the unique views of father and mother was significant ($\chi^2 (2) = 19.54, p < .01$). Dropping the additive genetic component on the common view ($\chi^2 (1) = 97.06, p < .01$) as well as on the unique views ($\chi^2 (2) = 18.09, p < .01$) of the parents also gave significant deteriorations of fit, indicating a significant contribution of A to the variance of the common and rater specific parts of the phenotype. Table 5 summarizes the estimates of genetic and environmental influences. The largest part of the variance

Table 1
Means and Standard Deviations (SD) for Untransformed Maternal and Paternal Ratings of Child Behavior Checklist Thought Problems (age 7) for First- and Second-Born Twins

	Mother ratings		Father ratings	
	Boys Mean (SD)	Girls Mean (SD)	Boys Mean (SD)	Girls Mean (SD)
Oldest twin	.39 (.97)	.33 (.86)	.31 (.80)	.24 (.70)
Youngest twin	.40 (.95)	.31 (.80)	.34 (.82)	.25 (.74)

Table 2
Maximum Likelihood Estimates of Thresholds for the Three Levels of Liability For Thought Problems (low, middle, high), with the Scores Ranging From 0 to 2

	Mother		Father	
	Oldest twin	Youngest twin	Oldest twin	Youngest twin
MZM				
Threshold 1	.89	.81	.94	.90
Threshold 2	1.48	1.40	1.55	1.48
DZM				
Threshold 1	.72	.67	.81	.70
Threshold 2	1.31	1.27	1.46	1.37
MZF				
Threshold 1	.90	.86	1.04	1.00
Threshold 2	1.42	1.52	1.67	1.78
DZF				
Threshold 1	.77	.82	.92	.90
Threshold 2	1.34	1.35	1.52	1.61
DOSMF				
Threshold 1	.70	.94	.77	.99
Threshold 2	1.24	1.56	1.41	1.62
DOSFM				
Threshold 1	.92	.71	1.04	.82
Threshold 2	1.49	1.34	1.61	1.44

Note: MZM/DZM = monozygotic/dizygotic males, MZF/DZF = monozygotic/dizygotic females. DOS = dizygotic opposite sex twins

Table 3
Correlations (Ratings by the Same Parent) and Cross-Correlations (Ratings Given by Different Parents) Between the Twins and Between the Parents

	Same Rater		Different Raters			
	Twins		Twins		Interparent	
	Mo/Mo	Fa/Fa	Mo/Fa	Fa/Mo	O	Y
MZM	.72	.78	.49	.46	.60	.65
DZM	.47	.49	.32	.29	.73	.67
MZF	.74	.76	.46	.51	.65	.67
DZF	.46	.42	.22	.23	.67	.59
DOSMF	.41	.44	.26	.23	.72	.60
DOSFM	.44	.53	.35	.33	.71	.69

Note: MZM/DZM = monozygotic/ dizygotic males, MZF/DZF = monozygotic/ dizygotic females. DOS = dizygotic opposite sex twins. Same rater twins = correlation between the oldest and the youngest twin, rated by Mo/Mo = mothers or Fa/Fa = fathers. Different raters twins = cross-correlation: either oldest twin rated by mothers and youngest by fathers (Mo/Fa) or the other way around (Fa/Mo). Different raters interparent: O = correlation between mother and father ratings for the oldest child; Y = idem for the youngest child.

Table 4

Summary of Model Fitting Results of Simultaneous Analysis of Paternal and Maternal Ratings of Thought Problems

	-2 LL	Nr. of parameters	df	$\Delta\chi^2$	Δdf	<i>P</i>	AIC	Compared to:
Saturated model								
Saturated	33966.377	84	30470				-26973.623	
Equal thresholds for:								
Mother and father	34050.563	60	30494	84.186	24	<.01	-26937.437	Saturated
MZ and DZ	33996.124	68	30486	29.746	16	.02	-26975.876	Saturated
Boys and Girls	34008.503	60	30494	42.125	24	.01	-26979.497	Saturated
Oldest/youngest	33998.724	60	30494	32.347	24	.11	-26989.276	Saturated
Psychometric model								
Full psychometric	33985.752	66	30492	19.375	22	.62	-26998.248	Saturated
Simplification psychometric model								
No sex diff.	33993.106	57	30499	7.354	7*	.60	-27008.894	Full psychom.
Factor estimates:								
No common A	34090.166	56	30500	97.06	1	<.01	-26909.834	No sex diff.
No unique A	34011.198	55	30501	18.092	2	<.01	-26994.802	No sex diff.
No common C	33994.078	56	30500	.972	1	.32	-27009.992	No sex diff.
No unique C	34015.649	55	30501	19.543	2	<.01	-26990.351	No sex diff.

Note: * The Δdf for this test is 7 because 2 parameters are constrained, as the total phenotypic variance for mother and father ratings is equal to 1.

Table 5

Variation Explained by Genetic, Shared Environmental and Non-Shared Environmental Factors

	Mother	Father
Genetic factors (A)		
Common	51%	51%
Unique	10%	14%
Shared environmental factors (C)		
Common	—	—
Unique	13%	13%
Nonshared environmental factors (E)		
Common	16%	16%
Unique	10%	6%

in liability for thought problems could be explained by genetic factors.

The commonly assessed part of the phenotype explains 67% (51% [= A] + 16% [= E]; see Table 5) of the total variance in maternal and paternal ratings (total variance = sum of all maternal/paternal parameters) and was influenced mainly by the children's genotype. Heritability of the commonly assessed TP phenotype was 76% (51% [= A] / 67% [= A + E]). The remaining variance of the commonly assessed phenotype was explained by non-shared environmental factors (16% [= E] / 67% [= A + E] = 24%).

The rater specific genetic influences seem somewhat higher for the father than the mother ratings (14% vs. 10%). These significant genetic influences on the unique views indicate that fathers and mothers assess reliable and rater specific information regarding TP in their children. The rater specific shared environmental influences explain about 13% of the variance both parent-ratings. This is the part of the variance that could represent rater bias. Higher maternal specific nonshared environmental

influences are observed (10 % vs. 6 %), which could reflect reliable information about the nonshared environmental influences on the variance of the phenotype, but may also include measurement error.

Discussion

This study examined the influence of genetic and environmental factors on thought problems in 7-year-old twins, based on parental reports. The analyses modeled genetic and environmental influences on both the commonly agreed upon phenotype and on unique views of the parents. Parents agreed to a large extent on the occurrence of thought problems (TP) in their twin offspring. The commonly assessed phenotype explained 67% (51% [= A] + 16% [= E]) of the total variance in maternal and paternal ratings and is influenced mainly by genetic factors. Heritability of the commonly assessed TP phenotype was 76% (51% [= A] / 67% [= A + E]; see Table 5). There was no evidence that shared family environment contributed to variance of the common phenotype. In contrast, shared environmental influences contributed to the unique views of parents. The rater specific shared environmental influences may be due to rater bias, but explain only a small part of the total variance (13% [= C] / 33% [= A + C + E]; see Table 5). The heritability estimates of this study are slightly higher than previously found heritabilities for the TP-scale (Edelbrock et al., 1995; Schmitz et al., 1995; Lin et al., 2006), which used mostly maternal reports. The large heritability estimates found for boys by Kuo et al. (2004) — using parental reports, which were also mostly maternal — were more in line with our findings. The gender differences in genetic and environmental influences found by Kuo et al. (2004) were not replicated in our study. Kuo et al. (2004) also

observed significant small nonadditive genetic influences for boys. Shared environmental influences were not significant in any of the previous studies (Edelbrock et al., 1995; Schmitz et al., 1995). In the study by Polderman et al. (2006), TP was rated by teachers and familial influences were detected, but to explain the familial influences, it was not possible to distinguish between common environment and genetic factors.

The genetic influences on the variance of the rater-specific parts are higher for the father reports in the analyses. Previous studies of internalizing and externalizing disorders in young twins usually found these estimates to be larger in maternal ratings (Bartels et al., 2003; Bartels et al., 2004a; Boomsma et al., 2005; Van der Valk et al., 2001; Van der Valk et al., 2003). This may support the notion that there are sex differences in the perception of parents with respect to this phenotype in their children and that fathers may add important extra information regarding the thought problems phenotype. Sex differences in human perception have been suggested from a very early age onwards (Connelan et al., 2000).

We observed sex differences in TP with boys showing more TP than girls. A similar sex difference is found in OCD, which is associated with TP (Hanna, 1995; Geller et al., 1998; Zohar, 1999). We also observed differences between mono- and dizygotic twins, with MZ twins obtaining lower TP scores than DZ twins. This suggests the presence of negative social interactions in twin pairs, since under interaction the prevalence rates for a categorical variable between MZ and DZ twins are expected to differ if that trait has an underlying continuous distribution (Carey, 1992).

The distribution of the TP scale was highly skewed and we analyzed the data using a threshold model. The means of the untransformed (continuous) data shown in Table 1 are also indeed lower than most other CBCL-subcales (Achenbach, 1991), which indicates that a majority of the subjects display few or no symptoms, which is expected in such a skewed distribution (Van den Oord et al., 2003). A simulation study by Derks et al. (2004) showed that analysis of L-shaped distributed data results in an underestimation of additive genetic, and shared environmental influences and an overestimation of non-shared environmental effects. After conducting the Mx-analyses on the untransformed data (results not shown), these differences were also found when comparing those estimates with the estimates made through analysis of the categorical data. Categorical analysis also has disadvantages. First, the statistical power is reduced, and requires large samples. This study used a large sample of 8962 twin pairs. Another disadvantage of the categorical data analyses is that they are computationally more demanding.

The most important finding from these data is the substantial heritability that is observed in young children. This strengthens the notion that the TP-scale measures a true syndrome. Other ways to support this

finding may include a linkage and genome wide association analyses. Further study is needed to investigate the differences between the parental reports. Further research is also needed to obtain insight into estimates of heritability in other age groups and to obtain information on the relation between thought problems and other disorders such as autism, OCD and psychotic disorders, to examine whether the thought problem phenotype could also have a predictive value for these disorders. It is important that such studies in children include both maternal and paternal reports.

Acknowledgments

We are grateful to the twin families for their participation. This research was supported by NWO grants 463-06-001; 575-25-006; 480-04-004; 'Spinozapremie' (NWO/SPI 56-464-14192); NIMH, RO1 MH58799-03; CNCR (Centre Neurogenetics Cognition Research); Dr Bartels is supported by NWO-VENI grant 451-04-034. We would like to thank David Sondervan for zygosity typing.

References

- Achenbach, T. M. (1991). *Manual for the Child Behavior Checklist/4-18 and 1991 profile*. Burlington, VT: University of Vermont, Department of Psychiatry.
- Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1979). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, *101*, 213–232.
- Baron-Cohen, S., & Hammer, J. (1997). Is autism an extreme form of the male brain? *Advances in Infancy Research*, *11*, 193–217.
- Bartels, M., Boomsma, D. I., Hudziak, J. J., Rietveld, M. J. H., van Beijsterveldt, T. C. E. M., & Van der Oord, E. J. C. G. (2004a). Disentangling genetic, environmental, and rater effects of internalizing and externalizing problem behavior in 10-year-old twins. *Twin Research and Human Genetics*, *7*, 162–175.
- Bartels, M., Hudziak, J. J., Boomsma, D. I., Rietveld, M. J. H., Beijsterveldt van, C. E., & M., Oord van, E. J. C. G. (2003). A study of parent ratings of internalizing and externalizing problem behavior in 12-year-old twins. *Journal of the American Academy of Child & Adolescent Psychiatry*, *42*, 1351–1359.
- Bartels, M., van Beijsterveldt, T. C. E. M., Derks, E. M., Stroet, T. M., Polderman, T. J. C., Hudziak, J. J., & Boomsma, D. I., (2007). Young Netherlands Twin Register (Y-NTR): A longitudinal multiple informant study of problem behavior. *Twin Research and Human Genetics*, *10*, 3–11.
- Bartels, M., Van den Oord, E. J. C. G., Hudziak, J. J., Rietveld, M. J. H., Beijsterveldt van, C. E. M., & Boomsma, D. I. (2004b). Genetic and environmental mechanisms underlying stability and change in problem behaviors at ages 3, 7, 10 and 12. *Developmental Psychology*, *40*, 852–867.

- Boomsma, D. I., de Geus, E. J., Vink, J. M., Stubbe, J. H., Distel, M. A., Hottenga, J. J., Posthuma, D., Beijsterveldt van, T. C. E. M., Hudziak, J. J., Bartels, M., & Willemsen, G. (2006). Netherlands Twin Register: From twins to twin families. *Twin Research and Human Genetics*, 9, 849–857.
- Boomsma, D. I., Van Beijsterveldt, C. E. M., & Hudziak, J. J. (2005). Genetic and environmental influences on anxious/depression during childhood: A study from the Netherlands Twin Register. *Genes, Brain and Behavior*, 4, 466–481.
- Carey, G. (1992). Twin imitation for antisocial behavior: Implications for genetic and family environment research. *Journal of Abnormal Psychology*, 101, 18–25.
- Connellan, J., Baron-Cohen, S., Wheelwright, S., Batki, A., & Ahluwalia, J. (2000). Sex differences in human neonatal social perception. *Infant Behavior & Development*, 23, 113–118.
- de Bruin, E. I., de Nijs, P. F. A., Verheij, F., Hartman, C. A., & Ferdinand, R. F. (2006). Multiple complex developmental disorder delineated from PDD-NOS. *Journal of Autism and Developmental Disorders*, 37, 1181–1191.
- Derks, E. M., Dolan, C. V., & Boomsma, D. I. (2004). Effects of censoring on parameter estimates and power in genetic modeling. *Twin Research*, 7, 659–669.
- Edelbrock, C., Rende, R., Plomin, R., & Thompson, L. (1995). A twin study of competence and problem behavior in childhood and early adolescence. *Journal of Child Psychology and Psychiatry*, 36, 775–785.
- Falconer, D. S. (1989). *Introduction to quantitative genetics*. Essex, England: Longman Scientific and Technical.
- Ferdinand, R. F., Blüm, M., & Verhulst, F. C. (2001). Psychopathology in adolescence predicts substance use in young adulthood. *Addiction*, 96, 861–870.
- Geller, D. A., Biederman, J., Jones, J., Shapiro, S., Schwartz, S., & Park, K. S. (1998). Obsessive compulsive disorder in children and adolescents: A review. *Harvard Review of Psychiatry*, 5, 260–273.
- Geller, D. A., Biederman, J., Faraone, S., Spencer, T., Doyle, R., Mullin, B., Magovcevic, M., Zaman, N., & Farrell, C. (2004). Re-examining comorbidity of obsessive compulsive and attention-deficit hyperactivity disorder using an empirically derived taxonomy. *European Child & Adolescent Psychiatry*, 13, 83–91.
- Gjone, H., & Stevenson, J. (1997). The association between internalizing and externalizing behavior in childhood and early adolescence: Genetic and environmental common influences? *Journal of Abnormal Child Psychology*, 25, 277–286.
- Hanna, G. L. (1995). Demographics and clinical features of obsessive-compulsive disorder in children and adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry*, 34, 19–27.
- Hessel, D., Dyer-Friedman, J. D., Glaser, B., Wisbeck, J., Barajas, R. G., Taylor, A., & Reiss, A. L. (2006). The influence of environmental and genetic factors on behavior problems and autistic symptoms in boys and girls with fragile X syndrome. *Pediatrics*, 108, 1–9.
- 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. *Behavior Genetics*, 22, 293–317.
- Hudziak, J. J., Rudiger, L. P., Neale, M. C., Heath, A. C., & Todd, R. D. (2000). A twin study of inattentive, aggressive and anxious/depressed behaviors. *Journal of the American Academy of Child and Adolescent Psychiatry*, 39, 469–476.
- Kasius, M. C., Ferdinand, R. F., van den Berg, H., & Verhulst, F. C. (1997). Associations between different diagnostic approaches for child and adolescent psychopathology. *Journal of Child Psychology and Psychiatry*, 38, 625–632.
- Kuo, P., Lin, C., Yang, H., Soong, W., & Chen, W. (2004). A twin study of competence and behavioral/emotional problems among adolescents in Taiwan. *Behavior Genetics*, 34, 63–74.
- Leve, L. D., Winebarger, A. A., Fagot, B. I., Reid, J. B., & Goldsmith, H. H. (1998). Environmental and genetic variance in children's observed and reported maladaptive behavior. *Child Development*, 69, 1286–1298.
- Lin, C., Kuo, P., Su, C., & Chen, W. (2006). The Taipei Adolescent Twin/Sibling Family Study I: Behavioral problems, personality features, and neuropsychological performance. *Twin Research and Human Genetics*, 9, 890–894.
- Little, R. J. A., & Rubin, D. B. (1987). *Statistical analysis with missing data*. New York: Wiley.
- Neale, M. C., Boker, S. M., Xie, G., & Maes, H. H. (2006). *Mx: Statistical modeling* (6th ed.). Richmond, VA: Department of Psychiatry, Medical College of Virginia.
- Polderman, T., Posthuma, D., De Sonneville, L., Verhulst, F., & Boomsma, D. I. (2006). Genetic analyses of teacher ratings of problem behavior in 5-year-old twins. *Twin Research and Human Genetics*, 9, 122–130.
- Rietveld, M. J. H., 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 reports. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, 1, 68–71.
- Rosenthal, R., Hall, J. A., DiMatteo, M. R., Rogers, P. L., & Archer, D. (1979). *Sensitivity to nonverbal communication, the PONS test*. Baltimore, MD: John Hopkins University Press.
- Rowe, D. C., & Kandel, D. (1997). In the eye of the beholder? Parental ratings of externalizing and internalizing symptoms. *Journal of Abnormal Child Psychology*, 25, 265–275.
- Schmitz, S., Fulker, D. W., and Mrazek, D. A. (1995). Problem behavior in early and middle childhood: An initial behavior genetic analysis. *Journal of Child Psychology and Psychiatry*, 36, 1443–1458.

- Schmitz, S., & Mrazek, D. A. (2001). Genetic and environmental influences on the associations between attention problems and other problem behaviors. *Twin Research*, 4, 453–458.
- Silberg, J. L., Erickson, M. T., Meyer, J. M., Eaves, L. J., Rutter, M. L., & Hewitt, J. K. (1994). The application of structural equation modeling to maternal ratings of twins' behavioral and emotional problems. *Journal of Consulting and Clinical Psychology*, 62, 510–521.
- Van den Oord, E. J. C. G., Koot, H. M., Boomsma, D. I., Verhulst, F. C., & Orlebeke, J. F. (1995). A twin-singleton comparison of problem behavior in 2- to 3-year-olds. *Journal of Child Psychology and Psychiatry*, 36, 449–458.
- Van den Oord, E. J. C. G., Pickles, A., & Waldman, I. (2003). Normal variation and abnormality: An empirical study of the liability distributions underlying depression and delinquency. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 44, 180–192.
- Van der Valk, J. C., van der Oord, E. J. C. G., Verhulst, F. C., & Boomsma, D. I. (2001). Using parental ratings to study the etiology of 3-year-old twins' problem behaviors: Different views or rater bias? *Journal of Child Psychology and Psychiatry*, 42, 921–931.
- Van der Valk, J. C., van der Oord, E. J. C. G., 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. *Behavior Genetics*, 33, 409–420.
- Van der Valk, J. C., Verhulst, F. C., Stroet, T. M., & Boomsma, D. I. (1998). Quantitative genetic analysis of internalizing and externalizing problems in a large sample of 3-year-old twins. *Twin Research*, 1, 25–33.
- Verhulst, F. C., Van der Ende, J. & Koot, H.M. (1996). *Handleiding voor de CBCL/4–18 [Manual for the CBCL/4–18]*. Afdeling Kinder- en Jeugdpsychiatrie, Sophia Kinderziekenhuis/Academisch Ziekenhuis Rotterdam/Erasmus Universiteit Rotterdam.
- Zahn-Waxler, C., Schmitz, S., Fulker, D., Robinson, J., & Emde, R. (1996). Behavior problems in 5-year-old monozygotic and dizygotic twins: Genetic and environmental influences, patterns of regulation, and internalization of control. *Development and Psychopathology*, 8, 103–122.
- Zohar, A. H. (1999). The epidemiology of obsessive-compulsive disorder in children and adolescents. *Child and Adolescent Psychiatric Clinics of North America*, 8, 445–460.
-