

# Psychopathology in 7-Year-Old Children: Differences in Maternal and Paternal Ratings and the Genetic Epidemiology

Laura W. Wesseldijk,<sup>1,2\*</sup> Iryna O. Fedko,<sup>1</sup> Meike Bartels,<sup>1,2,3</sup> Michel G. Nivard,<sup>1</sup> Catharina E. M. van Beijsterveldt,<sup>1</sup> Dorret I. Boomsma,<sup>1,2,3</sup> and Christel M. Middeldorp<sup>1,3,4</sup>

<sup>1</sup>Department of Biological Psychology, VU University Amsterdam, Amsterdam, The Netherlands

<sup>2</sup>EMGO+ Institute for Health and Care Research, Amsterdam, The Netherlands

<sup>3</sup>Neuroscience Campus Amsterdam, Amsterdam, The Netherlands

<sup>4</sup>Department of Child and Adolescent Psychiatry, GGZ inGeest/VU University Medical Center, Amsterdam, The Netherlands

Manuscript Received: 27 July 2016; Manuscript Accepted: 20 September 2016

The assessment of children's psychopathology is often based on parental report. Earlier studies have suggested that rater bias can affect the estimates of genetic, shared environmental and unique environmental influences on differences between children. The availability of a large dataset of maternal as well as paternal ratings of psychopathology in 7-year old children enabled (i) the analysis of informant effects on these assessments, and (ii) to obtain more reliable estimates of the genetic and non-genetic effects. DSM-oriented measures of affective, anxiety, somatic, attention-deficit/hyperactivity, oppositional-defiant, conduct, and obsessive-compulsive problems were rated for 12,310 twin pairs from the Netherlands Twin Register by mothers (N = 12,085) and fathers (N = 8,516). The effects of genetic and non-genetic effects were estimated on the common and rater-specific variance. For all scales, mean scores on maternal ratings exceeded paternal ratings. Parents largely agreed on the ranking of their child's problems ( $r$  0.60–0.75). The heritability was estimated over 55% for maternal and paternal ratings for all scales, except for conduct problems (44–46%). Unbiased shared environmental influences, i.e., on the common variance, were significant for affective (13%), oppositional (13%), and conduct problems (37%). In clinical settings, different cutoffs for (sub)clinical scores could be applied to paternal and maternal ratings of their child's psychopathology. Only for conduct problems, shared environmental and genetic influences explain an equal amount in differences between children. For the other scales, genetic factors explain the majority of the variance, especially for the common part that is free of rater bias.

© 2016 The Authors. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics* Published by Wiley Periodicals, Inc.

**Key words:** psychopathology; parental ratings; rater bias; shared environment; twins

## INTRODUCTION

Many childhood psychiatric disorders, including attention deficit hyperactivity disorder (ADHD), behavioral disorders, and anxiety

### How to Cite this Article:

Wesseldijk LW, Fedko IO, Bartels M, Nivard MG, van Beijsterveldt CEM, Boomsma DI, Middeldorp CM. 2017. Psychopathology in 7-Year-Old Children: Differences in Maternal and Paternal Ratings and the Genetic Epidemiology. *Am J Med Genet Part B* 174B:251–260.

disorders, are already prevalent at age 7. Pooled prevalence rates across countries at this age are estimated in a meta-analysis at 3.4% for ADHD, 3.6% for oppositional defiant disorder (ODD), 2.1%

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made. On behalf of all authors, the corresponding author states that there is no conflicts of interest.

Grant sponsor: European Research Council; Grant number: ERC-230374; Grant sponsor: Spinozapremie; Grant number: NWO/SPI 56-464-14192; Grant sponsor: Twin-family database for behavior genetics and genomics studies; Grant number: NWO 480-04-004; Grant sponsor: Genetic influences on stability and change in psychopathology from childhood to young adulthood; Grant number: ZonMW 912-10-020; Grant sponsor: European Union Seventh Framework Program; Grant number: 602768; Grant sponsor: Royal Netherlands Academy of Science Professor Award; Grant number: PAH/6635.

\*Correspondence to:

Laura W. Wesseldijk, Department of Biological Psychology, VU University Amsterdam, Van der Boeorchstraat 1, 1081 BT Amsterdam, The Netherlands.

E-mail: l.w.wesseldijk@vu.nl

Article first published online in Wiley Online Library (wileyonlinelibrary.com): 24 October 2016

DOI 10.1002/ajmg.b.32500

for conduct disorder (CD), and 6.5% for anxiety disorders [Polanczyk et al., 2015]. There are no pooled world-wide prevalences for obsessive-compulsive disorder (OCD), the estimated prevalence rates in children vary between 1% and 7% across countries [Thomsen, 2000; Hudziak et al., 2006]. In the current study, performed in a large sample of Dutch 7-year old twins, we investigated whether there are differences in paternal and maternal ratings of their child's psychopathology. Next, contributions of genetic, shared familial and unshared environmental factors on the differences between children in psychopathology as assessed by mothers and fathers were estimated.

Measures of childhood psychopathology below age 12 are often based on reports from mothers and/or fathers. The outcome of these assessments can depend on the rater. Earlier studies in 7-year-old children have found maternal ratings to be higher than paternal ratings [Duhig et al., 2000; van der Valk et al., 2003; Luoma et al., 2004; Boomsma et al., 2005; Abdellaoui et al., 2008; Langberg et al., 2010; Mascendaro et al., 2012; Sollie et al., 2013]. The ranking of the children's problems has also been found to vary between fathers and mothers resulting in correlations of 0.60 between maternal and paternal ratings [Achenbach et al., 1987; Duhig et al., 2000].

Several twin studies have investigated parental ratings of childhood psychopathology and estimated the influences of additive genetic (A), common environmental (C) factors shared by children growing up in the same household, and non-shared environmental (E) factors [Burt, 2009; Polderman et al., 2015]. As monozygotic (MZ) twin pairs share almost all of their genetic material, while dizygotic (DZ) twin pairs share on average 50%, a higher MZ than DZ twin correlation indicates that A plays a role in differences between children. If the DZ twin correlation is higher than half of the MZ twin correlation, C, representing environmental influences that create resemblance among siblings, may also be of influence. The remaining part of the variance is attributed to E, representing environmental influences that create differences among siblings and measurement error. Genetic factors have been consistently reported to have an influence on differences between children in psychiatric disorders or symptoms. In addition, usually C was found to explain additional variation, except for ADHD [Burt, 2009; Polderman et al., 2015].

Most estimates of A and C were based on studies that analyzed reports of one rater in which the assessments can partly reflect characteristics of the rater. A rater may systematically over or underestimate certain behavior in children and when the same rater assesses behavior in multiple children, the extent to which they resemble each other thus can in part be due to rater characteristics [Hewitt et al., 1992]. In genetic epidemiological analyses, this bias will create an effect that resembles an effect of common environment (C) as this bias results in a higher resemblance in MZ as well as in DZ twins [Bartels et al., 2007]. In studies that include reports of multiple raters, e.g., both parents for both twins, an unbiased estimate of the effect of shared environmental factors on childhood psychopathology can be obtained by estimating the effect of C on the part of the variance the parents agree upon. Previous multiple-rater studies in 7-year-old twins [Hudziak et al., 2003; van der Valk et al., 2003; Boomsma et al., 2005; Bartels et al., 2007; van Grootheest et al., 2007; Abdellaoui et al., 2008] showed that rater bias might account for around 10–30% of the phenotypic

variance. These studies focused on broad measures of internalizing and externalizing measures or on anxious depression, thought problems, and aggression. Such estimates are lacking for measures of childhood psychopathology that reflect symptoms associated with common psychiatric disorders as defined by the DSM-IV [American Psychiatric Association, 2000]. The effects of genetic and environmental factors on psychopathology have been shown to vary with age and so could the effects of rater bias [Bartels et al., 2007; Bergen et al., 2007; Burt, 2009; Polderman et al., 2015]. It is therefore important to analyze groups of children with a narrow age range, when considering the impact of rater bias.

Here, we analyze data on childhood psychopathology in a large sample of twins whose parents participate in the Netherlands Twin Register (NTR). Data were collected close to their seventh birthday in 12,310 twin pairs of which 8,480 twin pairs (67.4%) have ratings available from both parents. We analyzed the DSM-oriented affective, anxiety, somatic, attention-deficit/hyperactivity, oppositional-defiant, conduct and obsessive-compulsive problems scales of the Child Behavior Checklist (CBCL) [Achenbach and Rescorla, 2001]. The CBCL was originally developed to assess behavioral and emotional problems across a series of empirically defined scales based on exploratory (EFA) and confirmatory (CFA) factor analysis [Achenbach and Rescorla, 2001; Franic et al., 2014]. In contrast, items defining DSM-oriented scales were selected when 14 out of 22 experienced child psychiatrists and psychologist judged the item to be highly consistent with the relevant DSM-IV diagnostic category [Achenbach and Rescorla, 2001; Nelson et al., 2001]. A limited number of studies is available for the DSM-oriented scales. In one study in 398 Italian twin pairs [Spatola et al., 2010], the heritability for five DSM scales (i.e., affective, anxiety, attention, oppositional-defiant and conduct problems) varied between 34% and 74% in twins aged 8–11 and between 53% and 82% in twins aged 12–17. Significant effects of C were found for affective (39%) and anxiety (30%) problems in children aged 8–11. A recent Chinese twin study [Chen et al., 2015] in 658 twin pairs aged 6–18 reported modest genetic influences (19–37%) and substantial influences of C (54–67%) for the affective, anxiety, and somatic DSM-oriented CBCL scales. Chen et al. [2015] discussed that these discrepancies in genetic and environmental estimates across studies could be due to context differences, or to the broad age ranges and reports of one rater analyzed. For the obsessive-compulsive scale of the CBCL, analyzing maternal and paternal ratings yielded heritability estimates between 46% and 59% for boys and girls and an effect of C between 10% and 15% in 8,083 twin pairs at age 7 [van Grootheest et al., 2007].

We investigated whether parents differ in the assessments of their children, whether parental agreement depends on the child's gender and to what extent differences between children are explained by genetic, shared and non-shared environmental factors by simultaneously analyzing maternal and paternal ratings of mono- and dizygotic twins.

## METHODS

### Subjects

All participants were registered by their parents with the NTR shortly after birth [Boomsma et al., 2006; van Beijsterveldt et al.,

2013]. When the twins were 1, 2, 3, 5, 7, 10, and 12 years old, parents received a survey from the NTR. The first twins were registered in 1987 and recruitment and data collection is ongoing. For the current study, data of 7-year-old twins from birth cohorts 1986–2005 were analyzed. The surveys were mailed to the parents close to the twin's seventh birthday and reminders were sent after 2–4 months resulting in a response rate of 55% [van Beijsterveldt et al., 2013]. The final sample contained 2,079 monozygotic male (MZm), 2,324 monozygotic female (MZf), 2,086 dizygotic male (DZm), 1,924 dizygotic female (DZf), and 3,897 dizygotic opposite-sex twin pairs (DOS). More mothers ( $N = 12,085$ ) than fathers ( $N = 8,516$ ) completed the survey. For the same-sex twin pairs, zygosity was determined by DNA polymorphisms for 1,752 pairs, and otherwise by items in the survey about physical resemblance. Zygosity determination based on items about physical resemblance and DNA polymorphisms are in agreement in more than 93% of the twin pairs [Rietveld et al., 2000].

## Measures

Children's behavioral and emotional problems were assessed using the CBCL [Achenbach and Rescorla, 2001]. The CBCL is a rating scale for parents of children from 6 to 18 years old. It contains 118 specific items that are rated on a three-point scale (0–2; not true, somewhat true, very true). We analyzed variation among children's behavioral and emotional problems using the DSM-oriented CBCL scales that are consistent with diagnostic categories of the American Psychiatric Association's [1994] Diagnostic and Statistical Manual, 4th Edition (DSM-IV), namely the DSM-oriented affective, anxiety, somatic, attention-deficit/hyperactivity, oppositional-defiant, conduct and obsessive compulsive problems scales. Good validity for the DSM-oriented scales of the CBCL was reported in a US sample, with 80% of referred and non-referred children classified correctly and correlations with DSM-IV diagnostic categories ranging between 0.43 and 0.80 [Achenbach and Rescorla, 2001]. In a sample of Dutch children, the anxiety problems scale moderately predicted the presence or absence of a clinical DSM-IV anxiety disorder diagnosis, while the affective problem scale closely predicted DSM-IV major depression [Ferdinand, 2008]. In 2001, Nelson et al. created the obsessive-compulsive DSM-oriented scale after CFA and Andersen and Bilenberg [2012] confirmed high sensitivity and moderate specificity. Internal consistency in our sample, as reflected by Cronbach's  $\alpha$ , was on average 0.63, ranging from 0.49 to 0.76, which is comparable to findings in Spanish validation studies [Lacalle et al., 2012; Lacalle Sistere et al., 2014].

## Genetic Epidemiological Analyses and Multiple Rater Models

Since MZ twins share (nearly) all their genetic material, while DZ twins share, on average, 50% of their segregating genes, a higher phenotypic MZ twin correlation indicates that genetic factors play a role. If the DZ twin correlation is higher than half of the MZ twin correlation, shared environmental effects also are of importance. If the MZ twin correlation is more than twice as high than the DZ twin correlation, it is inferred that non-additive genetic factors

contribute to the phenotypic variance, in addition to additive genetic factors. Finally, the remaining part of the variance is attributed to non-shared environment effects and measurement error. This information can be captured in structural equation modeling, where the variance is decomposed into additive (A) and non-additive (D) genetic, shared environmental (C) and non-shared environmental (E) components. Since C and D have an opposing effect on the DZ twin correlations, their effects cannot be estimated simultaneously in the classical twin design. Based on the observed correlation pattern, an ACE or ADE model is tested. When ratings from mothers and fathers are available, it is possible to decompose the variances of the two ratings into a part the parents agree upon, the common part, and into two uncorrelated parts reflecting the disagreement, the rater-specific parts using the psychometric model (Fig. 1) [Hewitt et al., 1992]. These common and rater-specific parts of the variance can be further decomposed in variance explained by A, C or D, and E. For the common part, information comes from the cross-twin-cross-rater-correlations, i.e., the correlations between the maternal ratings of twin 1 and the paternal ratings of twin 2. Rater bias is excluded from the common part, but is contained in the rater-specific shared environmental influence ( $C_m$  and  $C_f$ ). Therefore, C on the common part is an unbiased estimate of the effect of shared environmental factors. Furthermore, whether the rater-specific parts also reflect true behavior of the child is inferred from the significance of the genetic influences on these rater-specific parts ( $A_m$  and  $A_f$ ), as it is unlikely that measurement error leads to an estimation of genetic influences [Neale and Cardon, 1992].

The DSM-oriented scales were highly skewed, as is usual given the relatively large number of subjects with no or little psychiatric symptoms in population-based cohorts. Treating the scores as normally distributed variables would result in parameter bias and it is therefore recommended to categorize the data and fit a threshold model [Derks et al., 2004]. Dividing the data using the clinical cut-off scores consistent with the DSM-IV diagnostic categories would neglect the variation between individuals below the diagnostic threshold; previous work suggested no etiological demarcation between variation within the normal variation and at the extreme end [Markon et al., 2011]. In addition, dividing the data into three approximately equally sized groups instead of two groups yields more power [van der Sluis et al., 2013]. Therefore, the scores were divided into three (low, middle, and high) groups, and analyzed as categorical data with two thresholds [Derks et al., 2004]. In the threshold model, it is assumed that the categorical trait has an underlying continuous distribution of liability [Falconer and Mackay, 1996]. Data from boys and girls were divided into three more or less equally sized groups (percentages of children in the first two groups are shown in Table I). As can be seen in the table, for each scale, the groups included children with identical problems scores for paternal and maternal ratings (see Table I). Consequently, differences in thresholds for boys and girls and paternal and maternal ratings reflect differences in prevalence rates of 0, 1, or 2 scores.

The analyses were performed in OpenMx [Boker et al., 2011]. In the baseline model, we estimated the thresholds for maternal and paternal ratings in boys and girls and the polychoric correlations, reflecting the correlations on the liability distribution. In addition

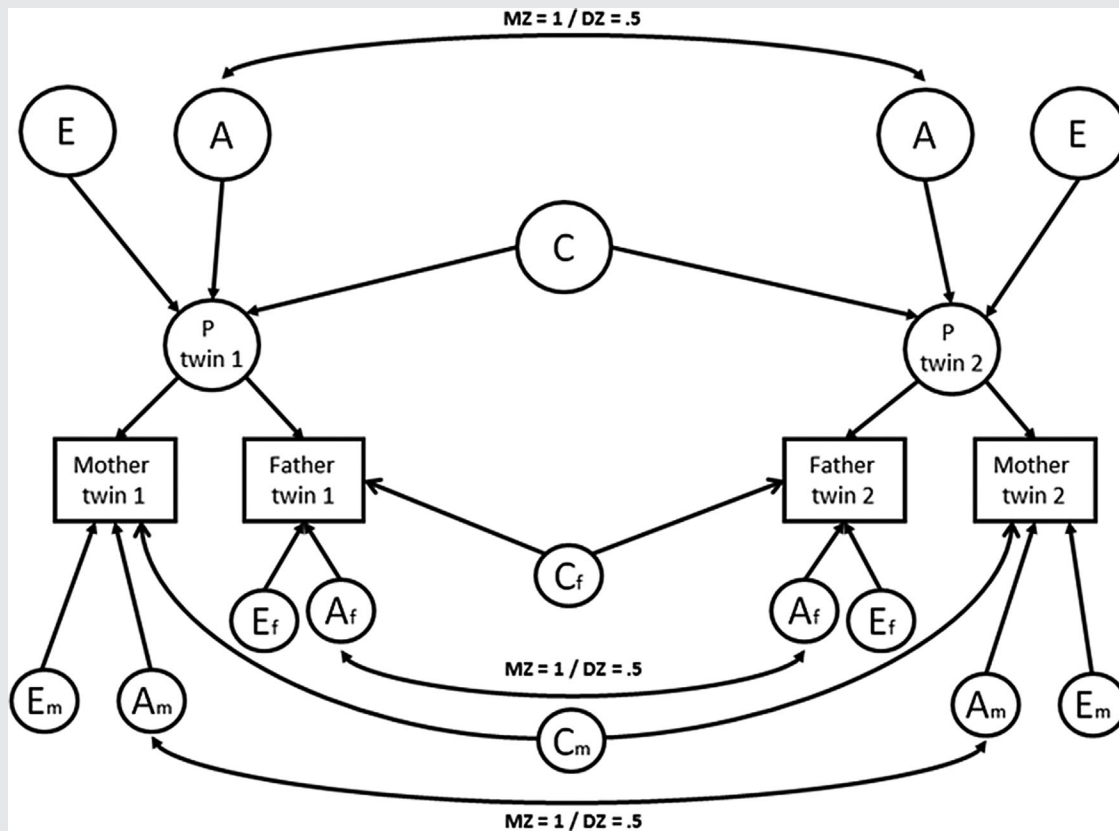


FIG. 1. The psychometric model. Maternal and paternal ratings are linear functions of the latent phenotypes of the twins (P twin 1 and P twin 2), and rater specific variance ( $A_m$ ,  $C_m$ ,  $E_m$ ,  $A_f$ ,  $C_f$ , and  $E_f$ ). When constraining  $A_m$  and  $A_f$  (underlined) to zero, the model represents a restricted rater bias model with  $C_m$  and  $C_f$  representing mother's and father's bias and  $E_m$  and  $E_f$  representing residual error.

to the correlations between paternal and maternal ratings, twin correlations for the maternal and the paternal ratings and the cross-twin-cross rater correlations were estimated for MZm, DZm, MZf, DZf, and DOS twin pairs. Sex differences in prevalence were analyzed by testing whether the thresholds could be constrained to be equal for mother and father ratings and for boys and for girls per rater. Next, we tested whether the parental agreement depended on zygosity or sex of the offspring, i.e., is parental agreement similar in MZ and DZ twins, and in boys and girls. Lastly, sex differences in the correlations were analyzed by constraining correlations between same-sex male and female twin pairs to be equal. This provides a test of quantitative sex differences in genetic architecture. To investigate qualitative sex effects, i.e., whether different genes operate in boys and girls, correlations between DZ same-sex and DOS twins were constrained to be equal [Vink et al., 2012]. The fit of these models was compared to the more general model.

Significance testing was based on the likelihood ratio test, where the negative log-likelihood ( $-2LL$ ) of the constrained model is subtracted from the  $-2LL$  of the more general saturated model. The difference between the  $-2LL$  of the two models follows a  $\chi^2$  distribution with degrees of freedom ( $df$ ) equal to the amount of constraints. If the difference in fit is significant, the more saturated model should be retained. If the difference in fit is not statistically

significant, the constrained model should be retained to achieve the best fitting and most parsimonious model.

Based on these outcomes, the psychometric model as explained above and depicted in Figure 1 was applied to the data to estimate the influences of A, C (or D), and E on the common and rater specific parts. Ninety-five percent confidence intervals were calculated to evaluate whether common and rater-specific A, C, or D were significant and whether estimates were similar in mothers and fathers.

## RESULTS

### Descriptives

The thresholds for boys and girls of maternal and paternal ratings in the 7-year-old Dutch twins are presented in Table I. The mean problem scores of the untransformed data and their standard deviations are given in Supplementary Table SI. In the model as estimated in OpenMx, including thresholds and correlations, mothers scored higher than fathers as reflected by the lower thresholds for the former ( $P < 0.001$  for all scales) (Supplementary Table SII). Furthermore, significant differences between boys and girls in the thresholds were observed ( $P < 0.001$  for all scales,  $P = 0.03$  for anxiety problems). Overall, girls scored higher on

TABLE I. The Thresholds for the Liability Distributions and the Percentages of Children in the Three Groups (Low, Middle, and High) for the Maternal (M) and Paternal (P) Ratings of the Different DSM-Oriented CBCL Scales in Boys and Girls

	Affective		Anxiety		Somatic		ADHD		Oppositional defiant		Conduct		Obsessive-compulsive	
	M	P	M	P	M	P	M	P	M	P	M	P	M	P
Boys														
Thresh1	-0.09	0.08	-0.16	-0.04	0.32	0.51	-0.20	-0.11	-0.35	-0.20	-0.23	-0.14	0.07	0.23
Thresh2	0.53	0.76	0.50	0.61	0.92	1.14	0.56	0.47	0.50	0.64	0.28	0.38	0.72	0.91
Low	46.4%	53.2%	43.6%	48.4%	62.6%	69.5%	42.1%	45.6%	36.3%	42.1%	40.9%	44.4%	52.8%	59.1%
Middle	23.8%	24.4%	25.6%	24.5%	19.5%	17.7%	29.1%	22.5%	32.9%	31.8%	20.1%	20.4%	23.6%	22.8%
High	29.8%	22.4%	30.8%	27.1%	17.9%	12.8%	28.8%	31.9%	30.9%	26.1%	38.7%	35.2%	23.6%	18.1%
Girls														
Thresh1	-0.14	-0.01	-0.21	-0.07	0.20	0.43	0.15	0.22	-0.13	0.02	0.15	0.20	0.02	0.21
Thresh2	0.51	0.71	0.46	0.60	0.79	1.05	0.89	0.78	0.79	0.91	0.72	0.76	0.73	0.92
Low	44.4%	49.6%	41.7%	47.2%	57.9%	66.6%	56%	58.7%	44.8%	50.8%	56%	57.9%	50.8%	58.3%
Middle	25.1%	26.6%	26%	25.4%	20.6%	18.7%	25.3%	19.5%	33.7%	31.1%	20.4%	19.7%	25.9%	23.8%
High	30.5%	23.8%	32.3%	27.4%	21.5%	14.7%	18.7%	21.8%	21.5%	18.1%	23.6%	22.4%	23.3%	17.9%
Low	0	0	0	0	0	0	0-1	0-1	0-1	0-1	0	0	0	0
Middle	1	1	1	1	1	1	2-3	2-3	2-3	2-3	1	1	1	1
High	2-15	2-17	2-12	2-11	2-14	2-14	4-10	4-10	4-10	4-10	2-23	2-20	2-15	2-13

The last three rows show the problem scores that were included into the three groups for each scale, with threshold 1 separating the low and middle groups and threshold 2 separating the middle and high groups. ADHD: attention deficit/hyperactivity disorder.

the affective, anxiety, somatic, and OCD scales and boys on the ADHD, ODD, and CD scales.

### Correlations Between Parents and Between Twins

The cross-rater correlations varied between 0.60 and 0.75 for both boys and girls within MZ and DZ twins (Table II). The correlations were fairly similar for all scales. The parental agreement did neither depend on the zygosity of the twins ( $P > 0.05$  for all scales), nor on sex except for CD where agreement was higher in boys ( $P < 0.001$  for CD,  $P > 0.01$  for the other scales) (Supplementary Table SII). Overall, parental agreement is similar in boys and girls and in MZ and DZ twins.

Furthermore, Table II shows the polychoric twin correlations for the maternal and paternal ratings. The MZ correlations were always higher than the DZ correlations, indicating that additive genetic factors play a role. With the exception of ADHD, the DZ twin correlations were higher than half of the MZ twin correlations which suggests influences from C. The MZ correlations for ADHD were more than twice as large as the DZ twin correlations, pointing to a role for D, besides A. The cross-twin-cross-rater correlations were higher for MZ than for DZ twins, implying that the common part of the variance is also influenced by genetic factors. There were no significant differences ( $P > 0.01$  for all scales) between the correlations of the same-sex male and female twin pairs (quantitative sex differences), or between the DZ same-sex and DOS twins for any of the scales (qualitative sex differences) (Supplementary Table SII).

### Genetic and Environmental Influences on Psychopathology

Rater-specific genetic influences were significant for all DSM-oriented scales ( $P < .001$ ) (Supplementary Table SII). This indicates that in addition to a common phenotype assessed by mother and father, the unique part of each parent's ratings reflects true behavior of the child.

Table III shows the overall heritability estimate and the overall effects of shared and nonshared environment (A, C/D, and E contributions to the common plus maternal or paternal rater-specific part) on the left. The individual parameter estimates and their confidence intervals of the A, C/D, and E contributions to the common and rater-specific variance are displayed in the middle. The two columns on the right give the total percentage of variance explained by the common part and the standardized parameter estimates of A, C or D, and E solely on the common part of the DSM-oriented CBCL-scales. This C estimate is unbiased. As an example, for affective problems, combining genetic influences on the commonly assessed part with the genetic influence on the maternal rater-specific part results in a heritability estimate of 62% using maternal ratings (43 [Acommon] + 19 [Am] in Table III). The commonly assessed part of affective problems explains 65% (43 [A] + 9 [C] + 13 [E]) of the total variance and the contribution of C on the commonly assessed part is 14% (9 [Ccommon]/65 [A + C + E common]) (shown under estimates on the % common in Table III).

TABLE II. Polychoric Cross-Rater Correlations for Boys and Girls Within MZ and DZ Twins, Polychoric Twin Correlations of the Maternal (M) and Paternal (P) Ratings, and Cross-Twin-Cross-Rater Correlations for the DSM-Oriented CBCL Scales

Cross-rater	Affective			Anxiety			Somatic			ADHD			Oppositional defiant			Conduct			Obsessive-compulsive			
	Boys	Girls	P	Boys	Girls	P	Boys	Girls	P	Boys	Girls	P	Boys	Girls	P	Boys	Girls	P	Boys	Girls	P	
	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	M	
<b>Twin correlations</b>																						
MZm	0.80	0.82		0.74	0.77		0.71	0.73		0.82	0.79		0.84	0.88		0.88	0.91		0.68	0.74		
DZm	0.47	0.49		0.39	0.42		0.44	0.38		0.26	0.26		0.55	0.55		0.67	0.66		0.40	0.37		
MZf	0.80	0.80		0.75	0.80		0.74	0.74		0.80	0.81		0.82	0.84		0.87	0.88		0.67	0.70		
DZf	0.48	0.49		0.40	0.48		0.48	0.40		0.23	0.25		0.54	0.57		0.68	0.70		0.64	0.62		
DOS	0.50	0.50		0.46	0.49		0.45	0.44		0.27	0.34		0.55	0.54		0.63	0.66		0.32	0.33		
<b>Cross-twin cross-rater</b>																						
MZm	0.54			0.49			0.47			0.62			0.65			0.60			0.44			
DZm	0.30			0.23			0.22			0.14			0.35			0.43			0.19			
MZf	0.53			0.49			0.46			0.59			0.58			0.54			0.40			
DZf	0.31			0.25			0.21			0.08			0.33			0.40			0.12			
DOS	0.31			0.27			0.25			0.17			0.33			0.41			0.20			

ADHD: attention deficit/hyperactivity disorder.

ADHD yielded the highest heritability with estimates of 81% and 80% for maternal and paternal ratings. For all other scales, except CD, adding the common and rater-specific parts resulted in heritability estimates between 54% and 64% (bold in Table III). For CD, lower heritability estimates were reported, namely 44% and 46%. For all scales, higher genetic influences were found on the commonly assessed variance (ranging between 52% and 82%) than on either the maternal or paternal ratings (ranging between 44% and 80%). The larger contribution of genetic effects to the common part of the ratings is due to lower contributions of C. Based on the 95% confidence intervals, differences in rater-specific genetic influences between parents were observed for somatic problems, ODD and OCD, with significantly higher estimates for father ratings than for mother ratings. For the remaining scales, the rater-specific genetic influences were equal between fathers and mothers, resulting in equal total heritability estimates for the parents. Furthermore, significant rater-specific C, which includes rater bias, was found for all scales (ranging between 7% and 20%), except for ADHD for which the effect of D was estimated. Rater-specific C was smaller for the affective, anxiety, and OCD problems scale (8–13%) than for the somatic, ODD, and CD problems scale (10–20%). Maternal-specific C was significantly higher for somatic, ODD, and OCD problems than paternal-specific C. The estimates of the proportion of the common variance explained by C was significant for affective (14%), ODD (13%), and CD problems (37%) (Table III). Recall that this estimation of C is free of rater bias.

## DISCUSSION

We provided an overview of maternal and paternal symptom scores and the genetic, shared, and unshared environmental contributions to individual variation in DSM-oriented problem scales measured at age 7: affective, anxiety, somatic, attention-deficit/hyperactivity (ADHD), oppositional-defiant (ODD), conduct (CD), and obsessive-compulsive (OCD) problems. Maternal ratings exceeded paternal ratings of child psychopathology for all scales, regardless of the child's sex. Furthermore, parents agreed to a large extent (correlations between 0.60 and 0.75) on the ranking of the problems in their children. Differences in the unique rater parts were explained by significant genetic influences as well as by rater bias. Hence, parents assess unique aspects of their child's behavior, but are also somewhat biased in their assessment. Regarding the contributions of genetic and environmental effects, the following points are noteworthy: the estimates of A, C or D, and E were comparable in boys and girls and largely comparable across raters, heritability estimates were generally around 60%. Estimates were comparable over scales, except for ADHD which was also influenced by D, and for CD which yielded lower heritability estimates and higher estimates for the influence of C. On the common part, unbiased shared environmental influences were significant for affective problems (13%), ODD (13%), and CD (37%). This signifies that for the remaining scales, it is possible that the common environmental effects as found in our and previous studies on the rater specific parts of the parental ratings or on reports of one rater are due to rater bias.

In this study, the correlations between parental measures are quite large, this indicates that parents generally agree on the

TABLE III. The Estimates of the Additive Genetic (A), Shared (C) or Non-additive Genetic (D) and Non-shared Environmental (E) Contributions for the Maternal and Paternal Ratings, for the Common and Rater-Specific Parts and for the Common as the Total Part

	Maternal ratings	Paternal ratings	Common part	Rater-specific mother's part	Rater-specific father's part	Total % common	Estimates on the % common
Affective							
A	<b>62</b>	<b>62</b>	43 [40-49]	19 [14-24]	19 [14-23]	65	66
C	17	18	9 [4-11]	8 [4-13]	9 [7-14]		<u>14</u>
E	20	19	13 [11-13]	7 [5-9]	6 [4-8]		20
Anxiety							
A	<b>63</b>	<b>64</b>	46 [41-51]	17 [11-22]	18 [12-23]	64	72
C	11	14	2 [0-7]	9 [4-12]	12 [7-17]		<u>3</u>
E	26	22	16 [15-17]	10 [8-11]	6 [4-8]		25
Somatic							
A	<b>55</b>	<b>64</b>	47 [45-49]	8 [2-11]	17 [15-20]	68	69
C	18	10	0 [0-0]	18 [18-20]	10 [7-10]		<u>0</u>
E	28	27	21 [19-21]	7 [4-7]	6 [5-8]		31
ADHD							
A	23	22	2 [0-12]	21 [19-23]	20 [19-22]	73	3
D	58	58	58 [48-60]	0 [0-0]	0 [0-0]		79
E	19	19	13 [13-15]	6 [5-7]	6 [4-8]		18
Oppositional defiant							
A	<b>56</b>	<b>61</b>	50 [49-55]	6 [4-6]	11 [10-14]	69	72
C	27	24	9 [6-11]	18 [17-21]	15 [13-17]		<u>13</u>
E	17	14	10 [9-11]	7 [6-8]	4 [3-5]		15
Conduct							
A	<b>44</b>	<b>46</b>	34 [29-39]	10 [6-14]	12 [7-17]	65	52
C	44	43	24 [21-28]	20 [15-24]	19 [15-23]		<u>32</u>
E	12	10	7 [6-9]	5 [4-7]	3 [2-5]		11
Obsessive-compulsive							
A	<b>54</b>	<b>63</b>	41 [39-44]	13 [6-18]	22 [15-29]	62	66
C	13	7	0 [0-0]	13 [10-17]	7 [2-13]		<u>0</u>
E	33	29	21 [19-23]	12 [10-15]	8 [5-11]		34

The estimates of the total additive genetic (A) [depicted in bold], shared (C) or non-additive genetic (D) and non-shared environmental (E) contributions to the common + rater-specific part for the maternal and paternal ratings are on the left side of the table. The parameter estimates [confidence intervals] of the A, C/D and E contributions to the common and rater-specific parts are in the middle. The last two columns give the total percentage of variance explained by the common part and the standardized parameter estimates of A, C or D and E on solely the common part of the DSM-oriented CBCL-scales. The reliable shared environmental effect is underlined.

ranking of their child. However, scores on maternal ratings were, on average, higher than scores on paternal ratings for all problem scales, i.e., mothers report on average more behavioral and emotional problems of their children than fathers did. This is in agreement with prior studies in 7-year-old children using different scales of the CBCL [van der Valk et al., 2003; Boomsma et al., 2005; Abdellaoui et al., 2008; Langberg et al., 2010; Mascendaro et al., 2012]. Since the cutoffs for subclinical or clinical scores as defined in the manual are the same for maternal and paternal ratings, children rated by their mother will more often pass the threshold than children rated by their father due to this difference between raters. This is especially important in situations where these scores are used to screen children for psychiatric symptoms, for example, to decide which children may benefit from an intervention, either prevention or treatment or to decide which children are eligible for inclusion in research studies. Consequently, it could be considered to apply different cutoffs for (sub)clinical scores for reports from mothers and fathers. Furthermore, our findings confirm that relying on a single parent for rating probably results in biased estimates of the influences of C [Hudziak et al., 2003; van der Valk et al., 2003; Boomsma et al., 2005; Bartels et al., 2007; van Grootheest et al., 2007; Abdellaoui et al., 2008]. We report reliable influences of C, free of rater bias, for affective problems, ODD, and CD in 7-year-olds only. Our C estimates were smaller than reported by Spatola et al. [2010] in Italian twins aged 8–11 and Chen et al. [2015] in Chinese twins aged 6–18, which can be explained by the fact that we used a smaller age range and multiple informants. Neither did we find a significant effect of C for OCD on the common part like van Grootheest et al. [2007], who also used the common perception shared by both parents, not confounded by rater bias, to estimate a unbiased C of 10% on OCD in 7-year-old twins. However, our study used a larger sample size and categorized the data. Derks et al. [2004] showed that the use of categorized data leads to unbiased estimates of genetic and environmental effects in L-shaped distributed data.

Furthermore, we found a high contribution of C (37%) on CD, which is discrepant with the two earlier studies that estimated C on the DSM-oriented CD scale [Spatola et al., 2010; Bertolotti et al., 2014]. Besides rater bias, there are two other biases that can confound C; namely imitation among twins and assortative mating. When twins imitate one another more than other siblings do, this can result in higher estimates of C, as this behavior increases the correlations of both monozygotic and dizygotic twins, but affects the dizygotic variances to a greater extent [Carey, 1986, 1992]. This can be detected by analyzing whether there are differences in the MZ and DZ thresholds [Carey, 1992]. This was not the case in our sample for CD ( $P=0.78$ ). Assortative mating, the tendency for people to mate with those who are more similar to themselves, can lead to an increase in genetic similarity in dizygotic twins, but not in monozygotic twins (as they already share 100% of their genes), which can also confound C. Spousal resemblance has been reported for scores on the DSM-oriented scales in a population-based sample, but not to a greater extent for antisocial problems than for other psychopathologies [Wesseldijk et al., 2016] and therefore cannot explain the high contribution of C solely on CD. Future research on the role of C on CD in young

children is therefore recommended. Genetically informative designs, such as children-of-twins or adoption studies are most suitable since these designs offer possibilities to account for genotype-environment correlations [O'Connor et al., 1998; McAdams et al., 2014].

This study has several strengths and weaknesses. The distribution of the different CBCL DSM-oriented scales were highly skewed, and we therefore analyzed the data using a threshold model, resulting in lower statistical power compared to an analysis of continuous data [Derks et al., 2004]. However, the parameter estimates in a threshold model are more accurate than in an analysis of continuous data [Posthuma and Boomsma, 2000]. While scores on the CBCL DSM-oriented scales are associated with the presence or absence of DSM diagnoses [Achenbach and Rescorla, 2001], they are not the same. Therefore, the heritability estimates apply only to the questionnaire scales. The major strength of this study is that we have fully explored informant effects and the influence of genetic and environmental factors on the rarely studied DSM-oriented scales of the CBCL in a large 7-year-old twin sample.

To conclude, this study shows that, besides the substantial genetic influence on the common and rater-specific parts for all scales, there appears to be a reliable effect of C only for affective, ODD and CD problems and not for anxiety, somatic and OCD problems. Additionally, fathers and mothers assess their child's psychopathology differently and this should be evaluated when using parental reports.

## ACKNOWLEDGMENTS

We are grateful to all twin families participating in the Netherlands Twin Register.

## REFERENCES

- Abdellaoui A, Bartels M, Hudziak JJ, Rizzu P, van Beijsterveldt TC, Boomsma DI. 2008. Genetic influences on thought problems in 7-year-olds: A twin-study of genetic, environmental and rater effects. *Twin Res Hum Genet* 11(6):571–578.
- Achenbach TM, McConaughy SH, Howell CT. 1987. Child adolescent behavioral and emotional-problems—implications of cross-informant correlations for situational specificity. *Psychol Bull* 101(2):213–232.
- Achenbach TM, Rescorla L. 2001. Manual for the ASEBA school-age forms & profiles: An integrated system of multi-informant assessment. Burlington, VT: ASEBA. xii, p 238.
- American Psychiatric Association. 2000. Diagnostic and statistical manual of mental disorders, 4th ed., text rev ed. Washington, DC: American Psychiatric Association.
- Andersen PAS, Bilenberg N. 2012. Comparison of child behavior checklist subscales in screening for obsessive-compulsive disorder. *Dan Med J* 59(11):1–6.
- Bartels M, Boomsma DI, Hudziak JJ, van Beijsterveldt TC, van den Oord EJ. 2007. Twins and the study of rater (dis)agreement. *Psychol Methods* 12(4):451–466.
- Bergen SE, Gardner CO, Kendler KS. 2007. Age-related changes in heritability of behavioral phenotypes over adolescence and young adulthood: A meta-analysis. *Twin Res Hum Genet* 10(3):423–433.



- Bertoletti E, Michelini G, Moruzzi S, Ferrer G, Ferini-Strambi L, Stazi MA, Ogliari A, Battaglia M. 2014. A general population twin study of conduct problems and the auditory P300 waveform. *J Abnorm Child Psychol* 42(5):861–869.
- Boker S, Neale M, Maes H, Wilde M, Spiegel M, Brick T, Spies J, Estabrook R, Kenny S, Bates T, Mehta P, Fox J. 2011. OpenMx: An open source extended structural equation modeling framework. *Psychometrika* 76(2):306–317.
- Boomsma DI, de Geus EJ, Vink JM, Stubbe JH, Distel MA, Hottenga JJ, Posthuma D, van Beijsterveldt TC, Hudziak JJ, Bartels M, Willemsen G. 2006. Netherlands Twin Register: From twins to twin families. *Twin Res Hum Genet* 9(6):849–857.
- Boomsma DI, van Beijsterveldt CE, Hudziak JJ. 2005. Genetic and environmental influences on anxious/depression during childhood: A study from the Netherlands Twin Register. *Genes Brain Behav* 4(8):466–481.
- Burt SA. 2009. Rethinking environmental contributions to child and adolescent psychopathology: A meta-analysis of shared environmental influences. *Psychol Bull* 135(4):608–637.
- Carey G. 1986. Sibling imitation and contrast effects. *Behav Genet* 16(3):319–341.
- Carey G. 1992. Twin imitation for antisocial behavior: Implications for genetic and family environment research. *J Abnorm Psychol* 101(1):18–25.
- Chen TJ, Ji CY, Wang SS, Lichtenstein P, Larsson H, Chang Z. 2015. Genetic and environmental influences on the relationship between ADHD symptoms and internalizing problems: A Chinese twin study. *Am J Med Genet Part B* 9999:1–7.
- Derks EM, Dolan CV, Boomsma DI. 2004. Effects of censoring on parameter estimates and power in genetic modeling. *Twin Res Hum Genet* 7(6):659–669.
- Duhig AM, Renk K, Epstein MK, Phares V. 2000. Interparental agreement on internalizing, externalizing, and total behavior problems: A meta-analysis. *Clin Psychol-Sci Pr* 7(4):435–453.
- Falconer DS, Mackay TFC. 1996. Introduction to quantitative genetics. Harlow, Essex, UK: Longmans Green.
- Ferdinand RF. 2008. Validity of the CBCL/YSR DSM-IV scales anxiety problems and affective problems. *J Anxiety Disord* 22(1):126–134.
- Franic S, Dolan CV, Borsboom D, van Beijsterveldt CE, Boomsma DI. 2014. Three-and-a-half-factor model? The genetic and environmental structure of the CBCL/6-18 internalizing grouping. *Behav Genet* 44(3):254–268.
- Hewitt JK, Silberg JL, Neale MC, Eaves LJ, Erickson M. 1992. The analysis of parental ratings of children's behavior using LISREL. *Behav Genet* 22(3):293–317.
- Hudziak JJ, Althoff RR, Stanger C, van Beijsterveldt CEM, Nelson EC, Hanna GL, Boomsma DI, Todd RD. 2006. The obsessive compulsive scale of the child behavior checklist predicts obsessive-compulsive disorder: A receiver operating characteristic curve analysis. *J Child Psychol Psc* 47(2):160–166.
- Hudziak JJ, van Beijsterveldt CE, Bartels M, Rietveld MJ, Rettew DC, Derks EM, Boomsma DI. 2003. Individual differences in aggression: Genetic analyses by age, gender, and informant in 3-, 7-, and 10-year-old Dutch twins. *Behav Genet* 33(5):575–589.
- Lacalle M, Ezpeleta L, Domenech JM. 2012. DSM-oriented scales of the child behavior checklist and youth self-report in clinically referred Spanish children. *Span J Psychol* 15(1):377–387.
- Lacalle Sistere M, Domenech Massons JM, Granero Perez R, Ezpeleta Ascaso L. 2014. Validity of the DSM-oriented scales of the child behavior checklist and youth self-report. *Psicothema* 26(3):364–371.
- Langberg JM, Epstein JN, Simon JO, Loren REA, Arnold LE, Hechtman L, Hinshaw SP, Hoza B, Jensen PS, Pelham WE, Swanson JM, Wigal T. 2010. Parent agreement on ratings of children's attention deficit/hyperactivity disorder and broadband externalizing behaviors. *J Emot Behav Disord* 18(1):41–50.
- Luoma I, Koivisto AM, Tamminen T. 2004. Fathers' and mothers' perceptions of their child and maternal depressive symptoms. *Nord J Psychiat* 58(3):205–211.
- Markon KE, Chmielewski M, Miller CJ. 2011. The reliability and validity of discrete and continuous measures of psychopathology: A quantitative review. *Psychol Bull* 137(5):856–879.
- Mascendaro PM, Herman KC, Webster-Stratton C. 2012. Parent discrepancies in ratings of young children's co-occurring internalizing symptoms. *Sch Psychol Q* 27(3):134–143.
- McAdams TA, Neiderhiser JM, Rijdsdijk FV, Narusyte J, Lichtenstein P, Eley TC. 2014. Accounting for genetic and environmental confounds in associations between parent and child characteristics: A systematic review of children-of-twins studies. *Psychol Bull* 140(4):1138–1173.
- Neale MC, Cardon LR, North Atlantic Treaty Organization. 1992. Methodology for genetic studies of twins and families. Dordrecht: Kluwer Academic Publishers. xxv, p 496.
- Nelson EC, Hanna GL, Hudziak JJ, Botteron KN, Heath AC, Todd RD. 2001. Obsessive-compulsive scale of the child behavior checklist: Specificity, sensitivity, and predictive power. *Pediatrics* 108(1):E14.
- O'Connor TG, Deater-Deckard K, Fulker D, Rutter M, Plomin R. 1998. Genotype-environment correlations in late childhood and early adolescence: Antisocial behavioral problems and coercive parenting. *Dev Psychol* 34(5):970–981.
- Polanczyk GV, Salum GA, Sugaya LS, Caye A, Rohde LA. 2015. Annual research review: A meta-analysis of the worldwide prevalence of mental disorders in children and adolescents. *J Child Psychol Psychiatry* 56(3):345–365.
- Polderman TJ, Benyamin B, de Leeuw CA, Sullivan PF, van Bochoven A, Visscher PM, Posthuma D. 2015. Meta-analysis of the heritability of human traits based on fifty years of twin studies. *Nat Genet* 47(7):702–709.
- Posthuma D, Boomsma DI. 2000. A note on the statistical power in extended twin designs. *Behav Genet* 30(2):147–158.
- Rietveld MJ, van Der Valk JC, Bongers IL, Stroet TM, Slagboom PE, Boomsma DI. 2000. Zygosity diagnosis in young twins by parental report. *Twin Res Hum Genet* 3(3):134–141.
- Sollie H, Larsson B, Morch WT. 2013. Comparison of mother, father, and teacher reports of ADHD core symptoms in a sample of child psychiatric outpatients. *J Atten Disord* 17(8):699–710.
- Spatola CA, Rende R, Battaglia M. 2010. Genetic and environmental influences upon the CBCL/6-18 DSM-oriented scales: Similarities and differences across three different computational approaches and two age ranges. *Eur Child Adolesc Psychiatry* 19(8):647–658.
- Thomsen PH. 2000. Obsessions: The impact and treatment of obsessive-compulsive disorder in children and adolescents. *J Psychopharmacol* 14(2):31–37.
- van Beijsterveldt CE, Groen-Blokhuis M, Hottenga JJ, Franic S, Hudziak JJ, Lamb D, Huppertz C, de Zeeuw E, Nivard M, Schutte N, Swagerman S, Glasner T, van Fulpen M, Brouwer C, Stroet T, Nowotny D, Ehli EA, Davies GE, Scheet P, Orlebeke JF, Kan KJ, Smit D, Dolan CV, Middeldorp CM, de Geus EJ, Bartels M, Boomsma DI. 2013. The Young Netherlands Twin Register (YNTR): Longitudinal twin and family studies in over 70,000 children. *Twin Res Hum Genet* 16(1):252–267.

- van der Sluis S, Posthuma D, Nivard MG, Verhage M, Dolan CV. 2013. Power in GWAS: Lifting the curse of the clinical cut-off. *Mol Psychiatry* 18(1):2–3.
- van der Valk JC, van den Oord EJ, Verhulst FC, Boomsma DI. 2003. Using shared and unique parental views to study the etiology of 7-year-old twins' internalizing and externalizing problems. *Behav Genet* 33(4): 409–420.
- van Grootheest DS, Bartels M, Cath DC, Beekman AT, Hudziak JJ, Boomsma DI. 2007. Genetic and environmental contributions underlying stability in childhood obsessive-compulsive behavior. *Biol Psychiatry* 61(3):308–315.
- Vink JM, Bartels M, van Beijsterveldt TC, van Dongen J, van Beek JH, Distel MA, de Moor MH, Smit DJ, Minica CC, Ligthart L, Geels LM, Abdellaoui A, Middeldorp CM, Hottenga JJ, Willemsen G, de Geus EJ, Boomsma DI. 2012. Sex differences in genetic architecture of complex phenotypes? *PLoS ONE* 7(12):47371.
- Wesseldijk LW, Dieleman GC, Lindauer RJ, Bartels M, Willemsen G, Hudziak JJ, Boomsma DI, Middeldorp CM. 2016. Spousal resemblance in psychopathology: A comparison of parents of children with and without psychopathology. *Eur Psychiatry* 34:49–55.

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's web-site.