Genetic and environmental influences upon the CBCL/6-18 DSM-oriented scales: similarities and differences across three different computational approaches and two age ranges

Chiara A. M. Spatola · Richard Rende · Marco Battaglia

Abstract Inasmuch as the newly established DSM-oriented CBCL/6-18 scales are to be increasingly employed to assess clinical/high-risk populations, it becomes important to explore their aetiology both within the normal- and the extreme range of variation in general population samples and to compare the results obtained in different age groups. We investigated by the Quantitative Maximum Likelihood, the De Fries-Fulker, and the Ordinal Maximum Likelihood methods the genetic and environmental influences upon the five DSM-oriented CBCL/6-18 scales in 796 twins aged 8–17 years belonging to the general population-based Italian Twin Registry. When children were analysed together regardless of age, most best-fitting solutions yielded genetic and non-shared environmental factors as the sole influences for DSM-oriented CBCL/6-18 behaviours, both for the normal and the extreme variations. When analyses were conducted separately for two age groups, shared environmental influences emerged consistently for Affective and Anxiety Problems in children aged 8–11. Oppositional-Defiant, Attention Deficit/Hyperactivity, and Conduct Problems appeared—with few exceptions—influenced only by genetic and non-shared environmental factors in both age groups, according to all three computational approaches. The De Fries-Fulker method appeared to be more sensitive in detecting shared environmental effects. Analysing the same set of data with different analytic approaches leads to better-balanced views on the aetiology of psychopathological behaviours in the developmental years.

Keywords DSM-oriented CBCL/6-18 scales · De Fries-Fulker Method · Shared environment · Twins · Age differences

C. A. M. Spatola (✉) · M. Battaglia
The Academic Centre for the Study of Behavioural Plasticity,
'Vita-Salute' San Raffaele University, 20 via Stamira d'Ancona,
20127 Milan, Italy
e-mail: chiara.spatola@hsr.it

R. Rende
Department of Psychiatry and Human Behavior,
Brown University, Providence, RI, USA

M. Battaglia
The Department of Clinical Neurosciences,
San Raffaele Institute, Milan, Italy

M. Battaglia
The Department of Child Psychiatry, Eugenio Medea Scientific
Institute, Bosisio Parini, Italy

Introduction

Parent-rated questionnaires that cover a wide array of psychopathological signs and symptoms continue to be relied upon to assess behavioural and emotional problems in children and adolescents. The Child Behaviour Checklist (CBCL) [2] is arguably the most widely used screening instrument in both the research and clinical settings, due in part to its good reliability and validity across different cultures [3, 7, 34, 59]. In addition to sampling an extensive array of problematic behaviours and to assessing their degree of severity, the CBCL scoring system also provides cut-off thresholds to discriminate between non-clinical, 'borderline', and clinically relevant behaviours [2].

The traditional CBCL subscales, generated on the basis of principal component analyses and variance-sharing of the 118 items on problem behaviours represent an empirically based assessment instrument that has been adopted worldwide. Cross-cultural comparisons [9] confirm the substantial

replicability of the psychometric structure of the CBCL, and the reliability, validity, and sensitivity of the scales, albeit with some differences regarding the mean and cut-off values across different cultures (see Achenbach and Rescorla [2] for further details). Some studies have employed CBCL items to build new scales which have proven effective in predicting juvenile bipolar disorder [17, 23].

Recently, a new scoring system based on consensus between clinicians [2] has been developed to allow better correspondence to the currently employed DSM-IV diagnostic criteria and syndromes and further foster the applicability of the CBCL as a screening instrument to clinical samples [1, 2, 30]. The new CBCL DSM-oriented scales (DOS) have thus been developed by including 55 items that had been rated as 'very consistent' with corresponding DSM-IV criteria by expert clinicians [1, 2] within six different syndrome scales, namely Affective, Anxiety, Somatic, Attention Deficit/Hyperactivity, Oppositional-Defiant, and Conduct Problems.

While the amount of empirical data presently available on the validity of the DOS is relatively small, it suggests satisfactory diagnostic efficiency and discrimination [28, 30]. For example, Ferdinand [18] recently showed that the CBCL DOS Anxiety- and Affective Problems scales predict the corresponding clinical DSM-IV diagnoses in a referred sample.

The traditional scales of the CBCL have been used in a number of behavioural genetic studies in order to provide insight on the magnitude of genetic and environmental influences on psychopathology in childhood [6, 25, 36, 42, 54]. Recently, we have examined the aetiological influences on the CBCL DOS as defined by Achenbach and Rescorla [2] in twins belonging to the Italian Registry by quantitative ML approach; our results show that the phenotypic variance of the new DOS scales is explained by only additive genetic and unique environmental agents. Moreover, heritability estimates were in general higher for the DOS than for the 'classical' CBCL scales [49].

Inasmuch as the CBCL 6/18 is likely to be increasingly employed for screening and assessing children who are potentially at risk for psychopathology, it becomes important to estimate whether the DOS' causal structure can be confirmed across normal and extreme variation. While the quantitative maximum-likelihood (ML) approach we employed to explore the causal structure of the DOS [49] is the most commonly employed to investigate the aetiology of individual differences within the normal range of variation, the particular mix of genetic and environmental influences on extreme scores may be substantially different [40, 42]. Thus, methods that specifically examine the aetiology of extremes are clinically important so that the most appropriate inferences may be made for levels of psychopathology that portend psychiatric disorder.

Two complimentary methods, the DeFries-Fulker (DF) analysis [11] and the ordinal ML method, are better suited to investigate the genetic and environmental causes of extreme scores. The DF analysis is a regression method designed to compute genetic and environmental effects on clinically significant levels of symptoms, considering also quantitative information for co-twins of probands, rather than simply classifying them as 'affected' or 'unaffected' [27, 40–42]. The ordinal ML approach investigates the genetic and environmental factors underlying extreme scores as represented in categorical terms, such as in the classical biomedical 'case' versus 'control' approach. Both DF analysis and the ordinal ML approach share the critical objective of examining the aetiology of extreme scores, and given that there is no a priori reason to favour one over the other, both are to be considered when conducting biometrical analyses of extreme scores.

An additional issue to consider is that a bulk of studies on problem behaviours show that the proportion of genetic and environmental contribution to individual differences vary depending on the age of children [5, 29, 44]: Silberg et al. [48] suggested that differences in the genetic or environmental contributions to problem behaviours in pre-pubertal versus post-pubertal youth can indicate the presence of different subtypes of the same disorder, or even different psychological syndromes.

While it is important to have estimates of genetic and environmental influences for problem behaviours in childhood and adolescence that are based on samples of normally developing subjects, clinicians, and epidemiologists are especially interested in knowing if, and to what extent, such estimates also apply to the more deviant scores within the same population analysed by different computational approaches, as this piece of information can influence policies in prevention and therapy.

Thus, the objectives of the present study are (1) to evaluate how the estimates of the genetic and environmental influences for CBCL/6-18 DOS vary, depending on whether psychopathology is approached by the quantitative, DF analysis, or the ordinal methods, and (2) to evaluate results obtained in two different age groups of twin children and adolescents.

Methods

Participants

This study is based on the responses provided to the CBCL 6/18 by parents of children belonging in the Italian Twin Registry (ITR), a database including all possible twins in the Italian population [50], from which subjects aged 8–17 years, living in the industrialized province of Milan

and in the suburban province of Lecco were sampled for psychometric studies of different nature and aims [35, 36, 49]. The recruitment methods and the socio-demographic characteristics are reported in detail elsewhere [36, 49]. Briefly, 398 twin pairs were recruited; the participation rate of 56% did not differ between families living in the industrialized Milan province compared with those living in the suburban Lecco province [36, 49]; the mean age of children and parents did not differ in the families who agreed, versus those who declined, participation into the study (respectively, 13.06 ± 2.60 vs. 13.11 ± 2.31, $P = 0.62$ for children; 46.06 ± 0.23 vs. 46.71 ± 0.35, $P = 0.10$ for mothers), and the maternal educational level and percentage of full-time employment were similar in participating and non-participating families (university degree: respectively, 17.7% vs. 16.3% of mothers, $P = 0.65$; mothers with full-time employment, respectively, 54 and 52%) [36, 49], with figures that closely reflect the average of Italian north-western population [26] and suggest the sample's representativeness of the general population of this part of the country. Moreover, all CBCL mean scores of twins in this sample were similar to those found in a national probability sample [19], with the exception of Rule-Breaking Behaviour, which was significantly higher (1.6 vs. 1.3, $P = 0.001$), and Aggressive Behaviour, which was significantly lower (4.6 vs. 6.3, $P = 0.001$), than in the national sample [49]. Comparisons based on t tests showed that the mean values of internalising problems DOS were significantly higher in girls (with the exception of Affective Problems, which was similar across genders), whereas the mean values of externalising problems scales were significantly higher in boys. Variances were significantly higher in boys for the externalising problems DOS, but did not differ between genders for the internalising problems DOS [49].

Measures

Twins' zygosity was assigned by the parent-rated Goldsmith questionnaire [21], which has an accuracy of determination of 99.4% [54]; according to its algorithm, there were 74 monozygotic (MZ) males, 70 MZ females, 134 same-sex dizygotic (DZS, 53 males, 81 females), and 120 different sex dizygotic (DZU) twin pairs. The zygosity of these 398 pairs was not different from that of the larger ITR population, and the MZ/DZS/DZU ratio was 1.1/1.0/0.9. The CBCL/6-18 was filled by one of the parents for both twins, with 77% percent of questionnaires completed by the mother, and the remaining 23% by the father, without significant differences in the mean values of the CBCL scales owing to parental sex [49]. In the present study the CBCL scores were obtained using the new scoring system based on DSM-oriented categories, which maintains the original

CBCL point scale. (range: 0–not true, 1–somewhat/sometimes true, 2–very true or often true). All the analyses were therefore fitted to each of the six CBCL DOS.

The procedures were accepted by the ethics committee of each participating institution, and for all participants parents signed a declaration of consent.

Model fitting analyses

Preliminary analyses showed that for all CBCL 6/18 DOS twin correlations did not differ significantly in DZU pairs and DZS pairs, suggesting that the same genes and shared environmental factors operate in both genders [6]. In consideration of this finding, and of the difficulty to adequately control for sex effects in a moderate-size sample like ours, we combined DZU and DZS twins together for the analyses, without assessing sex effects. The variances of all CBCL/6-18 DOS did not differ between MZ and DZ (DZU + DZS) twins, according to the Levene test for equality of variances (range: $P = 0.09$ for "Conduct Problems" subscale; $P = 0.87$ for "Affective Problems" subscale).

In order to explore the causal factors that underlie between-individuals differences in problem behaviours by taking into account both continuous variation and extreme scores, and to evaluate whether results differ quantitatively and qualitatively across different computational approaches, we analysed the individual scores obtained in the five CBCL/6-18 DOS by Structural Equation Modelling through three alternative model fitting methods, namely

1. Quantitative ML
2. DF analysis
3. Ordinal ML

All three computational approaches have the fundamental aim of quantifying the relative role of (1) genetic effects, which can be additive (A, resulting from the sum of all individual effects of alleles at different contributing loci) or non-additive (D, resulting from the effect of dominance/epistasis at one or more contributing loci); (2) shared environmental (C) effects, attributable to those environmental factors—such as family-wide experiences—that increase similarities within a twin pair; and (3) unique (i.e. individual-specific) environmental (E) influences that increase differences within a twin pair, also including measurement error. However, in a study conducted on twins reared together, dominant genetic and shared environmental components of variance are confounded and therefore cannot be estimated simultaneously [16].

The model assumes a correlation between twins' additive genetic influences of 1.0 for MZ pairs and of 0.5 for DZ pairs and a correlation between twins' shared environmental influences of 1.0 for both MZ and DZ twin pairs (i.e. shared environmental influences are assumed of equal

magnitude for MZ and DZ twins) [32]. While all three computational approaches share these basic elements, they also have unique features and specific trade offs.

The quantitative ML approach is the most widely employed method to explore the aetiology of individual
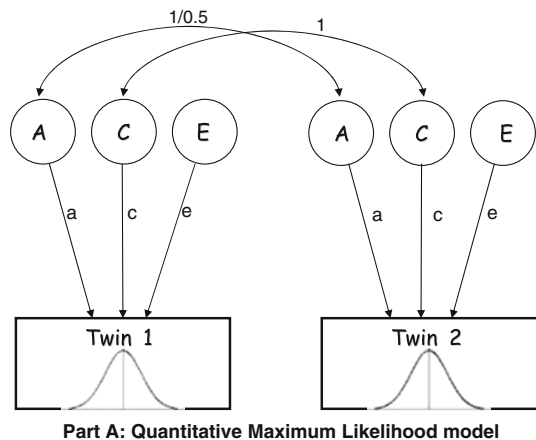
Part A: Quantitative Maximum Likelihood model

Part B: DeFries-Fulker multiple regression model

Part C: Ordinal Maximum Likelihood model

Fig. 1 Synoptic representation of the Quantitative Maximum Likelihood Model, De Fries-Fulker Regression Model, and of the Ordinal Maximum Likelihood Model. Observed variables are depicted in squares and latent variables in circles. Abbreviations: *A* additive genetic influences; *C* shared environmental influences; *E* non-shared environmental influences

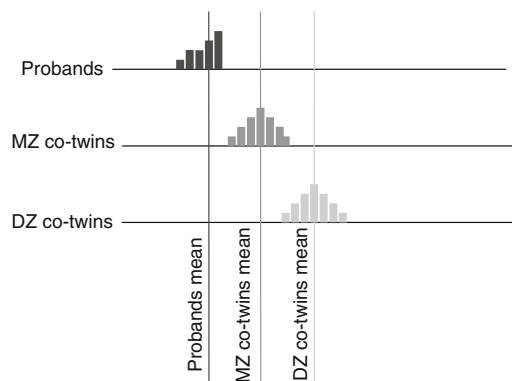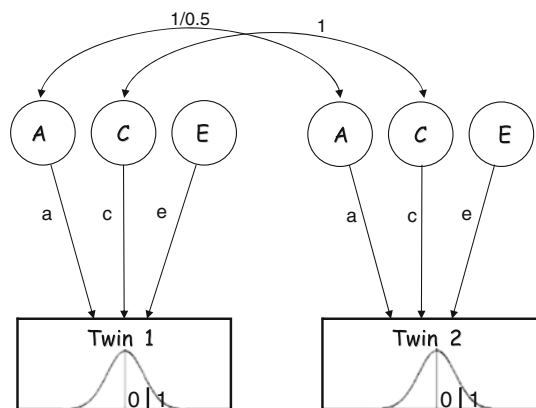differences in the normal range variation and capitalises upon decomposition of the MZ-DZ within-pairs covariance (see Fig. 1, part A). The proportion of the variance accounted for by genetic and environmental influences is calculated by squaring the parameters 'a', 'c', and 'e' and dividing them by the total variance ($a^2 + c^2 + e^2$). The model derives inferences about the relative role of genetic and environmental influences on continuous and normally distributed phenotypic scores and has less applicability when these distributional features are violated [12]. Since the large majority of CBCL/6-18 DOS in our sample had non-normally distributed scores (skewness range: 0.90–2.73; kurtosis range: 0.54–9.49), they were transformed by natural logarithm, to approximate normal distributions. After transformation, skewness and kurtosis were, respectively, between −0.25 and 1.22 and between −1.03 and 0.99 [49].

The DF analysis is a method based on standard multiple regression that allows to estimate heritability of extreme scores on a dimension [11]. Even if this model requires an arbitrary cut-off to be set upon the dimension of interest in order to identify the "cases" (probands), it capitalises on the quantitative scores of co-twins of probands, rather than classifying them as "cases" or "controls".

The model provides a simultaneous estimate of the genetic and shared environmental influences from the difference between the mean scores of MZ and DZ probands, their cotwins and the general population mean on the phenotype by examining the regression of the cosibling's mean as a function of genetic relatedness to the proband [11]. Common notation for the model is as follows:

$$C_i = b_3 P_i + b_4 R_i + b_5 P_i R_i + A$$

where the score for co-twin i (C_i) is predicted from the phenotypic score for twin i (P_i), a coefficient that denotes the genetic relationship between twins (R_i), each weighted by the estimated regression coefficient, and the intercept constant (A).

Differently from the quantitative ML methods that reflect an *individual differences* approach, the DF analysis yields *group differences* estimates, denoting to what extent the mean differences between an extreme group (i.e. probands) and the population on a dimensional measure, can be ascribed to genetic and environmental factors [37].

The general strategy of this method is quite straightforward: if an extreme value is affected by genes, the scores of DZ co-twins of probands should regress further to the population mean than scores of MZ co-twins; if environmental influences common to both twins play a role in influencing the extreme values, the scores of co-twins can be predicted from scores of probands without differences across zygosity groups; finally, if neither genes nor shared environment play an influence on the phenotype, co-twin

scores regress right back to the population mean (see Fig. 1, part B). We applied the Mx model-fitting implementation of the DF-analysis which refines the regression approach by basing the analysis on pairs rather than individuals [38]. Before model fitting analyses, individual scores were transformed according to the DF model so that the proband's mean is 1 and the general population mean is 0.

The third approach is the the Ordinal ML; this is the closest to the clinical 'case' and 'control' approach. By adopting an ordinal model-fitting approach we specified a threshold to separate 'normality' from 'pathology' upon the traits under study. This approach assumes phenotypic discreteness (or semi-continuity) of traits imposed upon of the same normally distributed underlying multi-factorial continuum of risk [39]. As such, it is the closest approximation to the classical medical distinction between affected and unaffected subjects while assuming a multi-factorial underlying liability, whereby several environmental and genetic determinants add to each other to influence the manifestation of an illness once a critical threshold is reached (see Fig. 1, part C).

While the cut-off values to separate 'normal' from 'clinical' and 'borderline' scores in the empirically defined CBCL scales are well-defined across different cultures [9, 19], no such distinction is yet available for the DOS in an Italian population. For both the DF analyses and the Ordinal ML approach we chose to set a relatively lenient threshold at the 85th percentile to, respectively, define 'probands' and 'cases', as suitable to increase sensitivity in an exploratory study of a general population sample like ours.

Since the age span of the children in this study is wide (8–17), and the contribution of genetic and environmental influences upon behaviour can vary in development, analyses were carried out both for subjects of all ages together, and split into two groups: 8–11 (prepubertal) and 12–17 (adolescent) years.

All the analyses were conducted from raw data using the structural-equation modelling program Mx [31]. The goodness-of fit index of each model is provided by the χ^2 goodness of fit and twice the log likelihood (−2LL), which then converted to the Akaike Information Criterion (AIC = −2LL−2df), that considers the rule of parsimony. Thus, the determination of the best-fitting model among nested models was based on the lowest AIC value, reflecting the optimal balance between model fitting and parsimony [22].

Results

Table 1 shows the twin correlations for each of the five CBCL/6-18 DOS. For all scales the MZ correlations were higher than the DZ correlations, indicating genetic effects on phenotypic variance. For the Attention Deficit/Hyperactivity Problems and the Conduct Problems scales, the MZ correlations were considerably greater than twice the DZ correlations, suggesting the possible role of non-additive genetic factors.

Table 2 shows the mode l fitting results obtained by applying the three different computational approaches of (a) Quantitative ML, (b) DF analysis, and (c) Ordinal ML in the whole sample.

We found considerable consistency of results across the three different approaches for the Anxiety, Oppositional Defiant, and Conduct Problems scales, whereby only genetic factors account for familial similarities, regardless of the computational approach.

For the Affective Problems scale, the quantitative ML approach and the ordinal ML method yielded similar results, with both approaches suggesting genetic factors as the only causes for familial similarities; when we applied the DF analysis regression model, however, we found that both an AE and a CE model could fit almost equally well to the data, with the latter being slightly preferable by virtue of the AIC.

Table 3 presents the proportions of variance accounted for by genetic, shared environmental, and non-shared environmental factors with their confidence intervals, as

Table 1 Twin correlations (95% confidence intervals) for the five DOS

	Entire sample		Age 8–11		Age 12–17	
	MZ	DZ	MZ	DZ	MZ	DZ
Affective Problems	0.63 (0.57;0.69)	0.35 (0.27;0.43)	0.42 (0.34;0.50)	0.39 (0.31;0.47)	0.69 (0.64;0.75)	0.33 (0.24;0.42)
Anxiety Problems	0.50 (0.43;0.58)	0.32 (0.23;0.41)	0.33 (0.24;0.42)	0.30 (0.21;0.39)	0.58 (0.51;0.65)	0.33 (0.24;0.42)
Attention Deficit/Hyperactivity Problems	0.60 (0.53;0.67)	0.19 (0.09;0.29)	0.67 (0.62;0.73)	0.15 (0.05;0.25)	0.57 (0.50;0.64)	0.21 (0.11;0.31)
Oppositional Defiant Problems	0.58 (0.51;0.65)	0.36 (0.28;0.45)	0.49 (0.42;0.56)	0.30 (0.21;0.39)	0.65 (0.60;0.71)	0.39 (0.31;0.47)
Conduct Problems	0.71 (0.67;0.77)	0.29 (0.20;0.38)	0.64 (0.58;0.70)	0.14 (0.04;0.24)	0.75 (0.71;0.79)	0.34 (0.26;0.43)

DOS DSM-oriented scales; *MZ* monozygotic; *DZ* dizygotic

Table 2 Model fitting and model comparisons for the quantitative ML, the DF-analysis, and the ordinal ML approach

CBCL/6–18 DSM-oriented scales	Model	Quantitative ML					DF-analysis (threshold 85%)					Ordinal ML (threshold 85%)				
		−2LL	DF	χ^2_{df}	P	AIC	−2LL	DF	χ^2_{df}	P	AIC	−2LL	DF	χ^2_{df}	P	AIC
Affective Problems	ACE	971.594	790	–	–	−608.406	1286.206	792	–	–	−297.794	714.434	793	–	–	−871.566
	AE	**971.810**	**791**	**0.216₁**	**0.6421**	**−610.190**	1287.815	793	1.609₁	0.2046	−298.185	**714.956**	**794**	**0.522₁**	**0.4700**	**−873.044**
	CE	986.899	791	15.305₁	0.0001	−595.101	**1287.039**	**793**	**0.833₁**	**0.3614**	**−298.961**	716.220	794	1.786₁	0.1814	−871.780
Anxiety Problems	ACE	884.348	790	–	–	−695.652	1512.646	792	–	–	−71.354	783.882	793	–	–	−802.118
	AE	**884.914**	**791**	**0.566₁**	**0.4519**	**−697.086**	**1512.742**	**793**	**0.096₁**	**0.7567**	**−73.258**	**783.882**	**794**	**0.000₁**	**1.0000**	**−804.118**
	CE	891.105	791	6.757₁	0.0093	−690.895	1515.960	793	3.314₁	0.0687	−70.040	787.227	794	3.345₁	0.0674	−800.773
Attention Deficit/Hyperactivity Problems	ACE	1196.937	790	–	–	−383.063	**1326.614**	**792**	–	–	**−257.386**	709.138	793	–	–	−876.862
	AE	**1196.937**	**791**	**0.000₁**	**1.0000**	**−385.063**	1336.265	793	9.651₁	0.0019	−249.735	**709.138**	**794**	**0.000₁**	**1.0000**	**−878.862**
	CE	1219.957	791	23.02₁	0.0001	−362.043	1345.887	793	19.273₁	0.0001	−240.113	716.330	794	7.192₁	0.0073	−871.670
Oppositional Defiant Problems	ACE	821.134	790	–	–	−758.866	1553.288	792	–	–	−30.712	803.052	793	–	–	−782.948
	AE	**821.540**	**791**	**0.406₁**	**0.5240**	**−760.460**	**1554.119**	**793**	**0.831₁**	**0.3620**	**−31.881**	**803.260**	**794**	**0.208₁**	**0.6483**	**−784.740**
	CE	833.278	791	12.144₁	0.0005	−748.722	1556.925	793	3.637₁	0.0565	−29.075	809.892	794	6.840₁	0.0089	−778.108
Conduct Problems	ACE	888.802	790	–	–	−691.198	1147.979	792	–	–	−436.021	674.003	793	–	–	−911.997
	AE	**888.802**	**791**	**0.000₁**	**1.0000**	**−693.198**	**1148.029**	**793**	**0.050₁**	**0.8231**	**−437.971**	**674.003**	**794**	**0.000₁**	**1.0000**	**−913.997**
	CE	929.966	791	41.164₁	0.0001	−652.034	1152.364	793	4.385₁	0.0363	−433.636	680.504	794	6.501₁	0.0108	−907.496

Best-fitting models are shown in bold

−2LL minus twice the log-likelihood; *DF* degrees of freedom; χ^2_{df} Chi-square goodness of fit with "df" degrees of freedom; *AIC* −2LL−2(DF)

Table 3 Proportions of variance accounted for by genetic, shared environmental and non-shared environmental factors as estimated from the full and the best-fitting models

CBCL/6-18 DSM-oriented scales	Model	Quantitative ML			Model	DF-analysis (threshold 85%)			Model	Ordinal ML (threshold 85%)		
		A	C	E		A	C	E		A	C	E
Affective Problems	ACE	0.57 (0.30;0.70)	0.05 (0.00;0.27)	0.38 (0.30;0.48)	ACE	0.25 (−0.29;0.78)	0.25 (−0.13;0.64)	0.50 (0.29;0.72)	ACE	0.44 (0.00;0.81)	0.19 (0.00;0.81)	0.37 (0.19;0.63)
	AE	**0.63 (0.54;0.71)**	–	**0.37 (0.29;0.46)**	CE	–	**0.42 (0.29;0.55)**	**0.58 (0.45;0.71)**	AE	**0.66 (0.45;0.82)**	–	**0.34 (0.18;0.55)**
Anxiety Problems	ACE	0.43 (0.11;0.38)	0.09 (0.00;0.32)	0.48 (0.38;0.61)	ACE	0.39 (−0.03;0.81)	0.05 (−0.27;0.37)	0.56 (0.40;0.72)	ACE	0.57 (0.00;0.74)	0.00 (0.00;0.18)	0.43 (0.26;0.67)
	AE	**0.54 (0.43;0.63)**	–	**0.46 (0.37;0.57)**	AE	**0.45 (0.32;0.59)**	–	**0.55 (0.41;0.68)**	AE	**0.57 (0.35;0.74)**	–	**0.43 (0.26;0.65)**
Attention Deficit/ Hyperactivity Problems	ACE	0.56 (0.41;0.66)	0.00 (0.00;0.09)	0.44 (0.34;0.55)	ACE	1.08 (0.61;1.56)	**−0.50 (−0.82;−0.19)**	0.42 (0.22;0.62)	ACE	0.55 (0.21;0.79)	0.00 (0.00;0.47)	0.45 (0.21;0.75)
	AE	**0.56 (0.45;0.66)**	–	**0.44 (0.34;0.55)**	AE	**0.37 (0.21;0.53)**	–	**0.63 (0.47;0.79)**	AE	**0.55 (0.25;0.79)**	–	**0.45 (0.21;0.75)**
Oppositional Defiant Problems	ACE	0.53 (0.24;0.68)	0.07 (0.00;0.29)	0.40 (0.32;0.51)	ACE	0.43 (−0.01;0.87)	0.15 (−0.17;0.47)	0.42 (0.25;0.60)	ACE	0.69 (0.18;0.90)	0.10 (0.00;0.49)	0.21 (0.10;0.40)
	AE	**0.61 (0.51;0.69)**	–	**0.39 (0.31;0.49)**	AE	**0.62 (0.47;0.77)**	–	**0.38 (0.23;0.53)**	AE	**0.80 (0.64;0.90)**	–	**0.20 (0.10;0.36)**
Conduct Problems	ACE	0.71 (0.60;0.78)	0.00 (0.00;0.09)	0.29 (0.22;0.37)	ACE	0.61 (0.04;1.18)	−0.05 (−0.46;0.37)	0.44 (0.21;0.66)	ACE	0.70 (0.21;0.86)	0.00 (0.00;0.36)	0.30 (0.14;0.53)
	AE	**0.71 (0.63;0.78)**	–	**0.29 (0.22;0.37)**	AE	**0.55 (0.36;0.74)**	–	**0.45 (0.26;0.64)**	AE	**0.70 (0.47;0.86)**	–	**0.30 (0.14;0.53)**

Best-fitting models are shown in bold

A additive genetic influences; C shared environmental influences; E non-shared environmental influences

determined by applying the three different methods in the whole sample.

The Attention Deficit/Hyperactivity Problems scale appeared to be strongly influenced by genetic factors: both quantitative and ordinal model fitting analyses yielded a best-fitting model encompassing additive genetic and non-shared environmental causal factors only. The DF regression model, which, differently from ML approaches, does not impose the constraint of positive values for parameter estimates, showed a negative estimate for the shared environmental component and an heritability estimate greater than one for this subscale, as a consequence of the more than twice difference between MZ and DZ phenotypic correlations [8].

Table 4 shows the best-fitting models and parameters' estimates (with their confidence intervals) when the three different computational models are applied to subjects aged 8–11 years and subjects aged 12–17 years separately. The best-fitting solutions for the internalising categories of DOS Anxiety Problems and DOS Affective Problems recognise environment as the sole source of variation for the 8–11 years age span, and additive genetic and environmental unique contributions for the 12–17 years age span across all three computational approaches. For the DOS Attention Deficit/Hyperactivity, Oppositional Defiant, and Conduct Problems, results in the two age ranges and across all three computational approaches closely resembled those obtained in the whole sample, in that additive genetic and environmental unique contributions appeared to be the only relevant sources of individual variation, albeit with three notable exceptions, all of which were observed under the DF approach: common environmental influences showed negative estimates for Attention Deficit/Hyperactivity and Conduct Problems, while Oppositional Defiant Problems in the 12–17 age range were found entirely environmental in nature.

Discussion

One of the main issues for developmental psychopathology concerns the aetiological connection between normal and abnormal behaviour [42]. In order to clarify the extent to which clinically relevant problems are aetiologically different from variation within the normal range, we explored the genetic and environmental factors influencing the five CBCL/6-18 DOS by implementing three different methods. By, respectively, focusing upon the normal and the abnormal domains, the quantitative and the ordinal ML methods reflect two conceptual and methodological approaches to aetiological research that are to some extent at the opposite. On the other hand, the DF analysis provides a framework that can help integrate these two approaches

Table 4 Proportions of variance accounted for by genetic, shared environmental and non-shared environmental factors in two separated age groups (8–11 and 12–17), as estimated from the best-fitting models

CBCL/6-18 DSM-oriented scales	Quantitative ML						DF analysis (threshold 85%)						Ordinal ML (threshold 85%)					
	Age 8–11			Age 12–17			Age 8–11			Age 12–17			Age 8–11			Age 12–17		
	A	C	E	A	C	E	A	C	E	A	C	E	A	C	E	A	C	E
Affective Problems	–	0.39 (0.24;0.53)	0.61 (0.47;0.76)	0.69 (0.58;0.76)	–	0.31 (0.24;0.42)	–	0.28 (0.06;0.51)	0.72 (0.49;0.94)	0.64 (0.43;0.85)	–	0.36 (0.15;0.57)	–	0.51 (0.17;0.76)	0.49 (0.24;0.83)	0.66 (0.40;0.84)	–	0.34 (0.16;0.60)
Anxiety Problems	–	0.30 (0.13;0.44)	0.70 (0.56;0.87)	0.61 (0.49;0.71)	–	0.39 (0.29;0.51)	–	0.22 (0.04;0.40)	0.78 (0.60;0.96)	0.54 (0.38;0.70)	–	0.46 (0.30;0.62)	–	0.16 (0.00;0.47)	0.84 (0.53;1.00)	0.71 (0.48;0.87)	–	0.29 (0.13;0.52)
Attention Deficit/ Hyperactivity Problems	0.62 (0.42;0.76)	–	0.38 (0.24;0.58)	0.53 (0.38;0.64)	–	0.47 (0.36;0.62)	1.12 (0.47;1.76)	–0.58 (–1.01;–0.14)	0.46 (0.19;0.74)	1.07 (0.44;1.70)	–0.47 (–0.90;–0.05)	0.40 (0.13;0.67)	0.34 (0.00;0.81)	–	0.66 (0.18;1.00)	0.63 (0.28;0.87)	–	0.37 (0.13;0.72)
Oppositional Defiant Problems	0.49 (0.28;0.64)	–	0.51 (0.36;0.72)	0.67 (0.57;0.76)	–	0.33 (0.24;0.43)	0.51 (0.25;0.77)	–	0.49 (0.23;0.75)	–	0.50 (0.37;0.63)	0.50 (0.37;0.63)	0.74 (0.39;0.93)	–	0.26 (0.07;0.61)	0.82 (0.63;0.93)	–	0.18 (0.07;0.37)
Conduct Problems	0.57 (0.35;0.72)	–	0.43 (0.28;0.65)	0.76 (0.67;0.83)	–	0.24 (0.17;0.33)	0.94 (0.19;1.69)	–0.49 (–1.02;0.05)	0.55 (0.24;0.85)	0.64 (0.40;0.88)	–	0.36 (0.12;0.60)	0.54 (0.00;0.90)	–	0.46 (0.10;0.99)	0.73 (0.47;0.89)	–	0.27 (0.11;0.53)

A additive genetic influences; *C* shared environmental influences; *E* non-shared environmental influences

[10]. A critical comparison of the results can also help highlighting the respective methodological strengths and weakness proper of each approach.

Genetic and environmental influences on CBCL/6-18 DOS in the sample as a whole

Our main finding is an overall agreement across the three different approaches in showing that most emotional and behavioural problems measured by the DOS are influenced mainly by genetic and non-shared environmental factors, both in the normal and in the extreme variations when all children aged 8–17 are pooled together. This confirms our previous findings with the quantitative ML approach about the aetiology of the CBCL DOS [49]. In general, this finding implies that there is a continuum of genetic risk that contributes to individual differences in emotional and behavioural problems as well as to the expression of clinically-relevant levels of symptomatology.

In this sense, there are no unique indicators of genetic or environmental risk on extreme scores of the CBCL DOS. That said, as these quantitative methods only determine the magnitude of the overall genetic effect and carry no implications as to specific regions of the genome, it is not to be assumed that there cannot be sets of genes unique to the extremes.

Despite this overall convergence across approaches, some notable exceptions were obtained by the application of the DF approach. First, when we applied the DF model to the Affective Problems subscale, the estimates from the full ACE model indicated a role of similar size for both additive genetic and shared environmental factors, while the best-fitting solution indicated only environmental factors as responsible for extreme group membership. A greater influence of shared environmental factors on extreme scores of depressive symptoms compared with normal variation has been previously reported [10, 15, 42, 43] (but see also Gjone et al. [20] for opposite findings, and the review by Rice et al. [44]). Our results with the DF analysis applied to Affective Problems DOS expands this work because of the clinical relevance of the CBCL DOS and reinforces the conclusion that shared environmental influences should be considered in aetiological models of depressive symptoms in childhood. In particular, the inference is that there may be environmental influences that promote concordance for siblings in terms of extreme affective symptoms. Although our study did not include putative environmental measures, future work could push the field forward by examining candidate social contextual factors (e.g. SES, parenting skills, and stress) that may be especially detectable at the extremes of the affective symptom distribution.

By applying the DF methodology to Attention Deficit/ Hyperactivity Problems subscale, we obtained a saturated

model that apparently yielded the best compromise between fit and parsimony. However, the estimate of the shared environmental component was lower than zero, with a heritability estimate greater than one, which has previously been reported by Waller [58] as indicative of non-additive genetic effects on the phenotype under the DF approach. Thus, rather than supporting a 'true' role for shared environmental influences, these figures may suggest a role for non-additive genetic factors. Accordingly, when the shared environmental estimate was constrained to be zero we obtained, as expected, estimates similar to those provided by the more conventional quantitative and ordinal model-fitting procedures. Both additive and non-additive genetic components' effects on Attention Deficit/Hyperactivity Problems have been reported by previous investigations [13, 24, 51].

Genetic and environmental influences on CBCL/6-18 DOS in pre-pubertal versus adolescent twins

Separate analyses in the two distinct age groups helped to further characterize our findings, by revealing that all three computational approaches converge in finding shared environmental influences as the sole sources for familial similarity in younger, pre-pubertal children, whereas additive genetic influences appear to be the sole sources for familial similarity in adolescents, for anxiety and depressive symptoms. These results suggest a change in the aetiology of internalizing symptoms, with genetic factors gaining importance in the transition from childhood to adolescence. It has been suggested that such increase in heritability during adolescence may partially reflect active gene-environment correlation [44], a phenomenon which we cannot investigate via the classical twin design, however. Our results with the Affective Problems scale in the two age groups agree with those of Boomsma et al. [5], which found similar trends for the CBCL juvenile bipolar scale. Consistently, Thapar and McGuffin [52], found that shared environmental influences on the one side, and genetic factors on the other, were substantial in explaining variance of children's (8–11 years) and adolescents' (12–16 years) depressive symptoms, respectively.

Other studies, however, found opposite trends: Gjone et al. [20], reported a reduction of heritability with age on the CBCL internalizing scale in twins aged 5–15, and Legrand et al. [29] found a moderate genetic contribution to self-reported anxiety measured in 10–12 years old, but not in 16–18 years old, female twins.

Our result of a substantial influence of shared environmental factors upon internalising problems supports investigations that take into account specific and focused analyses of the familial environment. In this regard, a meta-analysis by van der Bruggen et al. [56] found a substantial

association between child anxiety and parental control, with a stronger effect in school-aged children.

When we analysed Attention Deficit/Hyperactivity Problems, Oppositional Defiant and Conduct Problems in the two age groups, results generally replicated the findings obtained with the whole sample, with genetic and non-shared environmental factors as the sole sources of variation, albeit with the notable exception of shared environmental factors as the only source of familial similarities for Oppositional Defiant Problems in adolescents, found with the DF approach.

The same reasoning outlined above in discussing the negative estimates of shared environment in DF models applies also to the results in the two age groups. The relevance of genetic influences across age groups was previously reported for attention problems and hyperactivity [45] and for aggressive behaviours [25, 53].

Checks and balances of three computational approaches

It is often the case that when psychopathological behaviours are measured in general population subjects, the observed distributions depart significantly from normality. In fact, with behavioural phenotypes it is often found that the greater amount of subjects in the sample scores near to zero, and the 'affected individuals' fall within a relatively wide range of scores: this often yields L-shaped distributions [55]. When non-normally distributed traits are analysed by the quantitative ML approach, there is risk for specific bias of parameter estimations, even after reduction of skewness and kurtosis by data transformation [12]. When dealing with these non-normally distributed phenotypes, the ordinal ML approach is at lower risk for biased parameter estimations [55]. However, this same approach adopts one or more empirical thresholds that are often arbitrarily superimposed upon a sample distribution. Consequently, the information yielded by subjects who bear a score lower than the threshold for the 'affection state'- and are thereby defined 'non affected'-, is equated to null, with the consequence of loosing statistical power [12]. On the contrary, the DF analysis utilises the quantitative scores of co-twins of probands, instead of simply classifying them as "cases" or "controls" and is suitable also with non-normally distributed traits. By this method not only aetiological information about extremes can be acquired, but also quantitative variation in severity of impairment can be taken into account. This is not to say that DF analysis is without drawbacks: in the presence of non-additivity the risk for overestimation of additive genetic influences at the expenses of shared environmental factors is present, as seen also in our results with the Attention Deficit/Hyperactivity Problems scale.

At least four main potential limitations should be considered. The first lies in the relatively small size of our

sample, yielding limited power and preventing us from addressing other relevant issues, such as sex differences: although most DZU phenotypic correlations in our study do not differ significantly from DZS correlations, the inclusion of all DZ pairs in an unique category may have influenced the parameters' estimates. Since several studies found evidence of sex differences in the genetic architecture of externalizing- but not of internalizing, problems [24, 48, 57], our findings are perhaps more reliable for the internalized- than for the externalized traits.

Second, while the present study contributes to a growing data base that has examined the aetiology of extreme levels of emotional and behavioural symptoms in youth, it is based on a general population sample, implying that the subjects who score high on different levels of symptomatology are relatively few. To address this issue, and enhance the power of analysis, a relatively low threshold (85th percentile, compared with the 95th/98th percentile employed by Achenbach) was adopted for the DF analyses. While this choice remains arbitrary, it should be remembered that our aim was that of measuring the effects of genetic and environmental factors upon extreme variation, not necessarily upon a strictly-defined 'clinical' phenotype.

A third limitation is the assessment at one time point only, which is a potential confounder for the estimation of the effects of idiosyncratic environmental factors, and for the distinction of these effects from measurement error. Finally, we based the study on parental reports only, which may limit the sensitivity of the assessment, especially for internalising symptoms [46], even though some data [47] report better agreement for internalising than for externalising conditions between children and parents.

A more general cautionary note encompasses the use of screening measures like the CBCL for detecting psychopathology and drawing information on the genetic and environmental influences on mental health and disorders. While direct interviews in general population subjects elicit more information of possible clinical relevance, these studies are by definition based on children who are mostly healthy, and relying on paper-and-pencil instruments is still viable in face of realistic cost-effectiveness considerations.

Conclusion

A better understanding of the relationships that link normal and atypical development through multilevel approaches is a recognised, major goal of contemporary developmental psychopathology: here we sought to explore the causal sources of variation (both within the normal range and at the extremes) of five CBCL/6-18 DSM-oriented scales. The three computational methods we employed—the quantitative ML, the DF analysis, and the ordinal ML approach—have unique

advantages and drawbacks, and yielded broadly similar results; they also showed differences, such as distinguishable sensitivity in detecting a role for family-wide (shared) environmental determinants for depressive problems in children. Thus, while the quantitative ML is typically adopted for epidemiological samples, the DF-analysis for selected samples (such as sib pairs ascertained via proband sib), and the ordinal ML for clinical samples, by comparing the results obtained through all three approaches one can obtain a better balanced view of the aetiology of psychopathological behaviours in the developmental years.

Moreover, by showing the general tendency to confirm the same amount of genetic and extra-genetic contributions across different scales and computational approaches, the present study confirms that the variance in the newly established DSM-oriented CBCL/6-18 scales is moderately to substantially accounted for by genetic factors, with little, or no, contribution from shared environmental factors when all children aged 8–17 are pooled together. Best-fit solutions from all three computational approaches, however, correspond in suggesting a role for shared environmental factors in affecting the risk for Anxiety and Affective Problems in children aged 8–11 years, consistently with recent epidemiological-, quantitative genetic-, and molecular genetic studies [4, 14, 33]. Clinicians may therefore pay special attention in assessing the presence of environmental risk/precipitating factors in children who are referred for these disorders.

Since all solutions show an important role for non-shared environmental factors in influencing the DOS scores, future studies may focus on how specified environmental moderators—such as measures of life adversities, which were not available for this sample—influence variation in these scales via univariate and multivariate approaches.

References

1. Achenbach TM, Dumenci L, Rescorla LA (2003) DSM-oriented and empirically based approaches to constructing scales from the same item pools. J Clin Child Adolesc Psychol 32:328–340
2. Achenbach TM, Rescorla LA (2001) Manual for the ASEBA school-age forms & profiles. ASEBA, Burlington, VT: University of Vermont, Research

3. Achenbach TM, Verhulst FC, Baron GD, Althaus M (1987) A comparison of syndromes derived from the Child Behavior Checklist for American and Dutch boys aged 6–11 and 12–16. J Child Psychol Psychiatry 28:437–453

4. Battaglia M, Pesenti-Gritti P, Medland SE, Ogliari A, Tambs K, Spatola CA (2009) A genetically informed study of the association between childhood separation anxiety, sensitivity to CO_2, panic disorder, and the effect of childhood parental loss. Arch Gen Psychiatry 66:64–71

5. Boomsma DI, Rebollo I, Derks EM, van Beijsterveldt TC, Althoff RR, Rettew DC, Hudziak JJ (2006) Longitudinal stability of the CBCL-juvenile bipolar disorder phenotype: a study in Dutch twins. Biol Psychiatry 60:912–920

6. 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 4:466–481

7. Carter AS, Grigorenko EL, Pauls DL (1995) A Russian adaptation of the Child Behavior Checklist: psychometric properties and associations with child and maternal affective symptomatology and family function. J Abnorm Child Psychol 23:661–684

8. Cherny SS, DeFries JC, Fulker DW (1992) Multiple regression analysis of twin data: a model-fitting approach. Behav Genet 22:489–497

9. Crijnen AA, Achenbach TM, Verhulst FC (1999) Problems reported by parents of children in multiple cultures: the Child Behavior Checklist syndrome constructs. Am J Psychiatry 156:569–574

10. Deater-Deckard K, Reiss D, Hetherington EM, Plomin R (1997) Dimensions and disorders of adolescent adjustment: a quantitative genetic analysis of unselected samples and selected extremes. J Child Psychol Psychiatry 38:515–525

11. DeFries JC, Fulker DW (1985) Multiple regression analysis of twin data. Behav Genet 15:467–473

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

13. Derks EM, Hudziak JJ, Dolan CV, van Beijsterveldt TC, Verhulst FC, Boomsma DI (2008) Genetic and environmental influences on the relation between attention problems and attention deficit hyperactivity disorder. Behav Genet 38:11–23

14. Dunn J, O'Connor TG, Levy I (2002) Out of the picture: a study of family drawings by children from step-, single-parent, and non-step families. J Clin Child Adolesc Psychol 31:505–512

15. Eley TC (1997) Depressive symptoms in children and adolescents: etiological links between normality and abnormality: a research note. J Child Psychol Psychiatry 38:861–865

16. Evans DM, Gillespie NA, Martin NG (2002) Biometrical genetics. Biol Psychol 61:33–51

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

18. Ferdinand RF (2008) Validity of the CBCL/YSR DSM-IV scales Anxiety Problems and Affective Problems. J Anxiety Disord 22:126–134

19. Frigerio A, Cattaneo C, Cataldo MG, Schiatti A, Molteni M, Battaglia M (2004) Behavioral and emotional problems among Italian children and adolescents aged 4 to 18 years as reported by parents and teachers. Eur J Psychol Assess 20:124–133

20. Gjone H, Stevenson J, Sundet JM, Eilertsen DE (1996) Changes in heritability across increasing levels of behavior problems in young twins. Behav Genet 26:419–426

21. Goldsmith HH (1991) A zygosity questionnaire for young twins: a research note. Behav Genet 21:257–269

22. Heath AC, Neale MC, Hewitt JK, Eaves LJ, Fulker DW (1989) Testing structural equation models for twin data using LISREL. Behav Genet 19:9–35

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

24. Hudziak JJ, Derks EM, Althoff RR, Rettew DC, Boomsma DI (2005) The genetic and environmental contributions to attention deficit hyperactivity disorder as measured by the Conners' Rating Scales–Revised. Am J Psychiatry 162:1614–1620

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

26. ISTAT (2003) Rapporto Annuale 2002 Roma: Istituto Nazionale di statistica

27. Kohler HP, Rodgers JL (2001) DF-analyses of heritability with double-entry twin data: asymptotic standard errors and efficient estimation. Behav Genet 31:179–191

28. Krol NP, De Bruyn EE, Coolen JC, van Aarle EJ (2006) From CBCL to DSM: a comparison of two methods to screen for DSM-IV diagnoses using CBCL data. J Clin Child Adolesc Psychol 35:127–135

29. Legrand LN, McGue M, Iacono WG (1999) A twin study of state and trait anxiety in childhood and adolescence. J Child Psychol Psychiatry 40:953–958

30. Lengua LJ, Sadowski CA (2001) Assessing the child psychopathology beast: a reply to Achenbach and Dumenci's (2001) commentary. J Consult Clin Psychol 69:703–705

31. Neale MC, Boker SM, Xie G, Maes H (2003) Mx: Statistical Modeling. Department of Psychiatry, Virginia Commonwealth, Richmond

32. Neale MC, Cardon LR (1992) Methodology for genetic studies of twins and families. Kluwer, Dordrecht, The Netherlands

33. Nobile M, Rusconi M, Bellina M, Marino C, Giorda R, Carlet O, Vanzin L, Molteni M, Battaglia M (2009) The influence of family structure, the TPH2 G-703T and the 5-HTTLPR serotonergic genes upon affective problems in children aged 10–14 years. J Child Psychol Psychiatry 50:317–325

34. Novik TS (1999) Validity of the child behaviour checklist in a Norwegian sample. Eur Child Adolesc Psychiatry 8:247–254

35. Ogliari A, Citterio A, Zanoni A, Fagnani C, Patriarca V, Cirrincione R, Stazi MA, Battaglia M (2006) Genetic and environmental influences on anxiety dimensions in Italian twins evaluated with the SCARED questionnaire. J Anxiety Disord 20:760–777

36. Pesenti-Gritti P, Spatola CA, Fagnani C, Ogliari A, Patriarca V, Stazi MA, Battaglia M (2008) The co-occurrence between internalizing and externalizing behaviors. A general population twin study. Eur Child Adolesc Psychiatry 17:82–92

37. Purcell S, Eley TC, Dale PS, Oliver B, Petrill SA, Price TS, Saudino KJ, Simonoff E, Stevenson J, Taylor E, Plomin R (2001) Comorbidity between verbal and non-verbal cognitive delays in 2-year-old: a bivariate twin analysis. 4:195-196-208

38. Purcell S, Sham PC (2003) A model-fitting implementation of the DeFries-Fulker model for selected twin data. Behav Genet 33:271–278

39. Reich T, James JW, Morris CA (1972) The use of multiple thresholds in determining the mode of transmission of semi-continuous traits. Ann Hum Genet 36:163–184

40. Rende R (1999) Adaptive and maladaptive pathways in development: a quantitative genetic perspective. In: La Buda M, Grigorenko EL, Ravich-Serbo I, Scarr S (eds) On the way to individuality: current methodological issues in behavioral genetics. Nova Science, New York

41. Rende R, Slomkowski C, Lloyd-Richardson E, Niaura R (2005) Sibling effects on substance use in adolescence: social contagion and genetic relatedness. J Fam Psychol 19:611–618

42. Rende R, Slomkowski C, Lloyd-Richardson E, Stroud L, Niaura R (2006) Estimating genetic and environmental influences on depressive symptoms in adolescence: differing effects on higher and lower levels of symptoms. J Clin Child Adolesc Psychol 35:237–243

43. Rende RD, Plomin R, Reiss D, Hetherington EM (1993) Genetic and environmental influences on depressive symptomatology in adolescence: individual differences and extreme scores. J Child Psychol Psychiatry 34:1387–1398

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

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

46. Salbach-Andrae H, Klinkowski N, Lenz K, Lehmkuhl U (2009) Agreement between youth-reported and parent-reported psychopathology in a referred sample. Eur Child Adolesc Psychiatry 18:136–143

47. Seiffge-Krenke I, Kollmar F (1998) Discrepancies between mothers' and fathers' perceptions of sons' and daughters' problem behaviour: a longitudinal analysis of parent-adolescent agreement on internalising and externalising problem behaviour. J Child Psychol Psychiatry 39:687–697

48. Silberg JL, Erickson MT, Meyer JM, Eaves LJ, Rutter ML, Hewitt JK (1994) The application of structural equation modeling to maternal ratings of twins' behavioral and emotional problems. J Consult Clin Psychol 62:510–521

49. Spatola CA, Fagnani C, Pesenti-Gritti P, Ogliari A, Stazi MA, Battaglia M (2007) A general population twin study of the CBCL/6–18 DSM-oriented scales. J Am Acad Child Adolesc Psychiatry 46:619–627

50. Stazi MA, Cotichini R, Patriarca V, Brescianini S, Fagnani C, D'Ippolito C, Cannoni S, Ristori G, Salvetti M (2002) The Italian Twin Project: from the personal identification number to a national twin registry. Twin Res 5:382–386

51. Taylor E (2009) Developing ADHD. J Child Psychol Psychiatry 50:126–132

52. Thapar A, McGuffin P (1994) A twin study of depressive symptoms in childhood. Br J Psychiatry 165:259–265

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

54. van Beijsterveldt CE, Verhulst FC, Molenaar PC, Boomsma DI (2004) The genetic basis of problem behavior in 5-year-old dutch twin pairs. Behav Genet 34:229–242

55. van den Oord EJ, Pickles A, Waldman ID (2003) Normal variation and abnormality: an empirical study of the liability distributions underlying depression and delinquency. J Child Psychol Psychiatry 44:180–192

56. van der Bruggen CO, Stams GJ, Bogels SM (2008) Research review: the relation between child and parent anxiety and parental control: a meta-analytic review. J Child Psychol Psychiatry 49:1257–1269

57. van der Valk JC, Verhulst FC, Stroet TM, Boomsma DI (1998) Quantitative genetic analysis of internalising and externalising problems in a large sample of 3-year-old twins. Twin Res 1:25–33

58. Waller NG (1994) A DeFries and Fulker regression model for genetic nonadditivity. Behav Genet 24:149–153

59. Yang HJ, Soong WT, Chiang CN, Chen WJ (2000) Competence and behavioral/emotional problems among Taiwanese adolescents as reported by parents and teachers. J Am Acad Child Adolesc Psychiatry 39:232–239