Behav Genet (2014) 44:254–268
DOI 10.1007/s10519-013-9628-4

ORIGINAL RESEARCH

Three-and-a-Half-Factor Model? The Genetic and Environmental Structure of the CBCL/6–18 Internalizing Grouping

Sanja Franić · Conor V. Dolan · Denny Borsboom · Catherina E. M. van Beijsterveldt · Dorret I. Boomsma

Received: 19 June 2013 / Accepted: 24 October 2013 / Published online: 7 November 2013

Abstract In the present article, multivariate genetic item analyses were employed to address questions regarding the ontology and the genetic and environmental etiology of the Anxious/Depressed, Withdrawn, and Somatic Complaints syndrome dimensions of the Internalizing grouping of the Child Behavior Checklist/6–18 (CBCL/6–18). Using common and independent pathway genetic factor modeling, it was examined whether these syndrome dimensions can be ascribed a realist ontology. Subsequently, the structures of the genetic and environmental influences giving rise to the observed symptom covariation were examined. Maternal ratings of a population-based sample of 17,511 Dutch twins of mean age 7.4 ($SD = 0.4$) on the items of the Internalizing grouping of the Dutch CBCL/6–18 were analyzed. Applications of common and independent pathway modeling demonstrated that the Internalizing syndrome dimensions may be better understood as a composite of unconstrained genetic and environmental influences than as causally relevant entities generating the observed symptom covariation. Furthermore, the results indicate a common genetic basis for anxiety, depression, and withdrawn behavior, with the distinction between these syndromes being driven by the individual-specific environment. Implications for the substantive interpretation of these syndrome dimensions are discussed.

Keywords CBCL · Internalizing problems · Depression · Anxiety · Common pathway model · Independent pathway model · Genetic item analysis

Introduction

The development of taxonomy of psychiatric symptoms has traditionally been challenging. Difficulties in delineating between diagnostic categories, arising from issues such as overlapping features of multiple disorders, inconsistent empirical evidence regarding the factor structure of psychometric instruments, definitional issues arising from high comorbidity rates, debates regarding dimensional versus categorical conceptualization, and unknown degree of etiological overlap between symptoms or sets of symptoms, have notoriously hampered the attempts of arriving at a classification of psychopathology that would gain univocal support from the empirical researchers and the clinical practitioners alike. In children, these issues are further exacerbated by the developmental aspect of the disorders: for instance, the same disorder may manifest itself through different symptoms over time, while identical symptoms may reflect distinct, temporally changing underlying conditions.

Symptoms of anxiety and depression, for instance, famously illustrate the aforementioned issues (Clark and Watson 1991; Brown 1996; Brady and Kendall 1992; Rapee et al. 2009; Mineka et al. 1998). The definitional and etiological questions surrounding these disorders (and their extremely high comorbidity rates) are as old as the systematic study of the disorders itself. Are these two highly

Electronic supplementary material The online version of this article (doi:10.1007/s10519-013-9628-4) contains supplementary material, which is available to authorized users.

S. Franić (✉) · C. V. Dolan · C. E. M. van Beijsterveldt · D. I. Boomsma
Department of Biological Psychology, Faculty of Psychology and Education, VU University Amsterdam, Van der Boechorststraat 1, 1081 BT Amsterdam, The Netherlands
e-mail: s.franic@vu.nl

D. Borsboom
Department of Psychological Methods, University of Amsterdam, Amsterdam, The Netherlands

comorbid disorders manifestations of a single syndrome, or separate entities with overlapping features? To what extent are their etiologies shared? Is their symptom overlap a reflection of inadequacies of the current diagnostic systems, or an indication of a shared etiology? These and similar questions have stimulated ample and diverse theoretical development, and motivated a vast amount of research. The theories range from those postulating anxiety and depression as different points along a single continuum, to those conceptualizing them as conceptually and empirically distinct phenomena (Clark 1989).

This complexity, inherent to the study of psychiatric disorders, is further compounded by a lack of agreement in evaluating and understanding the structure of psychometric instruments used to assess psychopathology. The Child Behavior Checklist (CBCL, Achenbach and Rescorla 2001), for instance, is one of the most widely used instruments in assessing childhood psychopathology. It has been translated into over 85 languages, and more than 6,000 publications from over 65 countries report its applications in both the practical and the research context. However, when faced with critical empirical and psychometric evaluations, the syndrome dimensions postulated in the CBCL do not always stand up to scrutiny. In possibly the most comprehensive critical psychometric/empirical evaluation of the CBCL to date, Hartman et al. 1999 compellingly demonstrated that the 8-factor cross-informant model of the CBCL (Achenbach and Rescorla 2001, described below) fails to adequately describe the empirical data across multiple cultures under study, in both population-based and clinical samples. Furthermore, if violations of distributional assumptions, invariably present in the analysis of CBCL data, are taken into account when evaluating model fit, the conclusions of the studies indicating acceptable or nearly acceptable fit are often undermined (see Hartman et al. 1999). Upon close scrutiny, it therefore appears that the postulated 8-factor structure of the CBCL does not consistently survive critical confrontation with empirical data. This, naturally, raises questions about the instrument's validity: what do the CBCL syndrome dimensions measure, given the lack of unambiguous empirical support for the proposed 8-factor structure?

In the present paper, we propose that multivariate genetic item analyses (e.g., Heath et al. 1989; Kendler et al. 1987; van den Berg et al. 2007; Eaves 1983; Neale et al. 2005; Waller and Reise 1992; Franic et al. 2012b), as first applied to individual psychiatric symptoms by Kendler et al. (1987), can be used to illuminate some of the aforementioned issues. Specifically, genetic item analyses can be employed to examine how some of the difficulties in delineating the CBCL syndrome dimensions may arise as a function of the complexity of the latent genetic and environmental structures that underlie the observed symptom covariation. In addition, the applications of this type of analysis can contribute to the discussion on whether the current CBCL dimensions may be conceptualized as well-defined, coherent entities exerting causal influence on item covariation (i.e., whether they can be ascribed a realist ontology; Borsboom et al. 2003), or are better considered an unconstrained amalgamation of genetic and environmental influences. In the present article, we focus on the Internalizing grouping of the CBCL (items of the CBCL pertaining to intropunitive emotions and moods), with the aim of answering two principal questions: (1) Can one interpret the Internalizing syndrome dimensions of the CBCL substantively and causally? (2) What is the structure of the genetic and environmental influences giving rise to the observed (i.e., phenotypic) symptom covariation? We do not place primary emphasis on detailed phenotypic dimensionality assessment, and use it mainly insofar as it serves as a gateway into exploring the latent genetic and environmental dimensionality.

Methods

Data

The data were obtained from the Netherlands Twin Register at VU University Amsterdam (Bartels et al. 2007; van Beijsterveldt et al. 2012), and consist of maternal ratings of a population-based sample of 17,511 twins (including 3,023 MZ and 5,599 DZ complete twin pairs) of mean age 7.4 ($SD = 0.4$) on the Internalizing grouping of the Dutch version of the Child Behavior Checklist for ages 6–18 (CBCL/6–18; Achenbach and Rescorla 2001[1]). The CBCL/6–18 is a 140-item questionnaire used to assess problem behaviors and competencies in children, as reported by their parents. The cross-informant model of the CBCL (Achenbach and Rescorla 2001) was derived through the application of principal components analysis, and consists of eight correlated syndrome dimensions, broadly clustered into those pertaining to internalizing problems (the Internalizing grouping) and those pertaining to externalizing problems (the Externalizing grouping). The Internalizing grouping of the CBCL is a scale designed to measure disturbances in intropunitive emotions and moods in children, and consists of three subscales (i.e., syndrome dimensions): Anxious/Depressed (AD), Withdrawn (W), and Somatic Complaints (SC), containing 31 discrete items (listed in Fig. 1) in total. Responses are given on a three-point scale: "not true", "somewhat or sometimes true",

[1] The study had permission to permission to use, reproduce and reformat the CBCL.

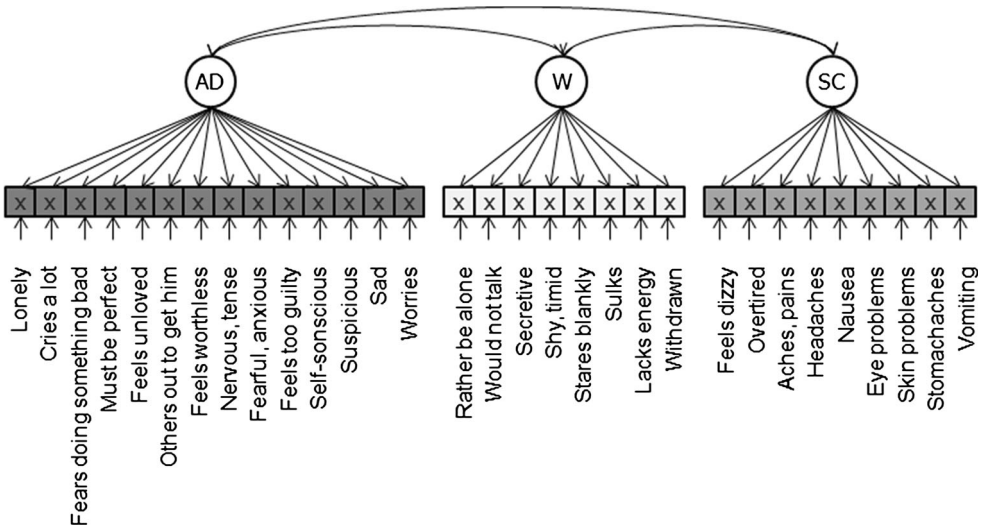

Fig. 1 The syndrome dimensions and item content of the CBCL/6-18 Internalizing grouping. *AD* Anxious/Depressed, *W* Withdrawn, *SC* Somatic Complaints

and "very true or often true". A path-diagrammatic representation of the three syndrome dimensions of the Internalizing grouping is given in Fig. 1.

Approach

Genetic covariance structure modeling (Martin and Eaves 1977) is the application of structural equation modeling (Bollen 1989; Kline 2005) to data collected in genetically informative samples, such as samples of twins (Neale and Cardon 1992; Franic et al. 2012b). In the classical twin design, the sample consists of monozygotic (MZ) and dizygotic (DZ) twin pairs. DZ twins share an average of 50 % of their segregating genes, while MZ twins entirely share their segregating DNA (Falconer and Mackay 1996; van Dongen et al. 2012). In the present analyses, the covariance structure of the phenotypes (i.e., observed traits, symptoms) is modeled as a function of latent factors representing three sources of individual differences: additive genetic (A), shared environmental (C) and individual-specific environmental (E) sources. Additive genetic influences are modeled by one or more A factors, which represent the total additive effects of genes relevant to the phenotypes. Based on quantitative genetic theory (Falconer and Mackay 1996; Mather and Jinks 1971), the A factors are known to correlate 1 across MZ twins and 0.5 across DZ twins. Environmental influences affecting the phenotype of both twins in an identical way, thereby increasing their similarity beyond what is expected based on genetic resemblance alone, are represented by one or more C factors. Therefore, by definition, the C factors correlate unity across twins (regardless of zygosity). All environmental influences causing the phenotype of two family members to differ are represented by one or more E factors. Thus, by definition, the E factors are correlated 0 across twins.[2] The expected covariance structure in a multivariate twin model is thus:

$$\begin{matrix}\Sigma_{11}\Sigma_{12} = & \Sigma_A + \Sigma_C + \Sigma_E & r_A\Sigma_A + \Sigma_C \\ \Sigma_{21}\Sigma_{22} & (r_A\Sigma_A + \Sigma_C)^t & \Sigma_A + \Sigma_C + \Sigma_E,\end{matrix} \quad (1)$$

where, given p phenotypes, $\boldsymbol{\Sigma}_{11}$ ($\boldsymbol{\Sigma}_{22}$) is the p × p covariance matrix of twin 1 (twin 2), $\boldsymbol{\Sigma}_{12}$ ($\boldsymbol{\Sigma}_{21}$) is the twin 1–twin 2 p × p covariance matrix, and $\boldsymbol{\Sigma}_A$, $\boldsymbol{\Sigma}_C$ and $\boldsymbol{\Sigma}_E$ are the additive genetic, shared environmental, and unique environmental p × p covariance matrices, respectively. The coefficient r_A is the additive genetic twin correlation (1 for MZ twins, 0.5 for DZ twins).

Figure 2 depicts two examples of the multivariate twin models used in the present study. The first model is a common pathway model (Kendler et al. 1987), also known as the psychometric factor model (McArdle and Goldsmith 1990). In a common pathway model, all of the A, C, and E influences on the item responses are mediated by a latent variable, henceforth referred to as the psychometric factor (factors P_1 and P_2 in Fig. 2). P_1 and P_2 may be viewed as latent phenotypic factors, e.g., 'anxiety' or 'depression'. In common pathway models, the psychometric factor acts as a mediator of the genetic and environmental effects, and the factor loadings represent common pathways from the A, C, and E factors to the observed item responses.

The second model is the independent pathway model (Kendler et al. 1987), also known as the biometric factor

[2] In addition, the phenotype may be influenced by non-additive genetic effects (D), which are the result of interactions of alleles within the same locus (genetic dominance) or across different loci (epistasis). These will not be modeled in the present paper, as the classical twin design does not allow for simultaneous estimation of A, D, C, and E effects. The choice between modeling C and D effects was informed by preliminary univariate item analyses, which showed most of the items to conform better to an ACE than to an ADE model. We note, however, that this does not exclude the presence of non-additive genetic influences (Keller and Coventry 2005).

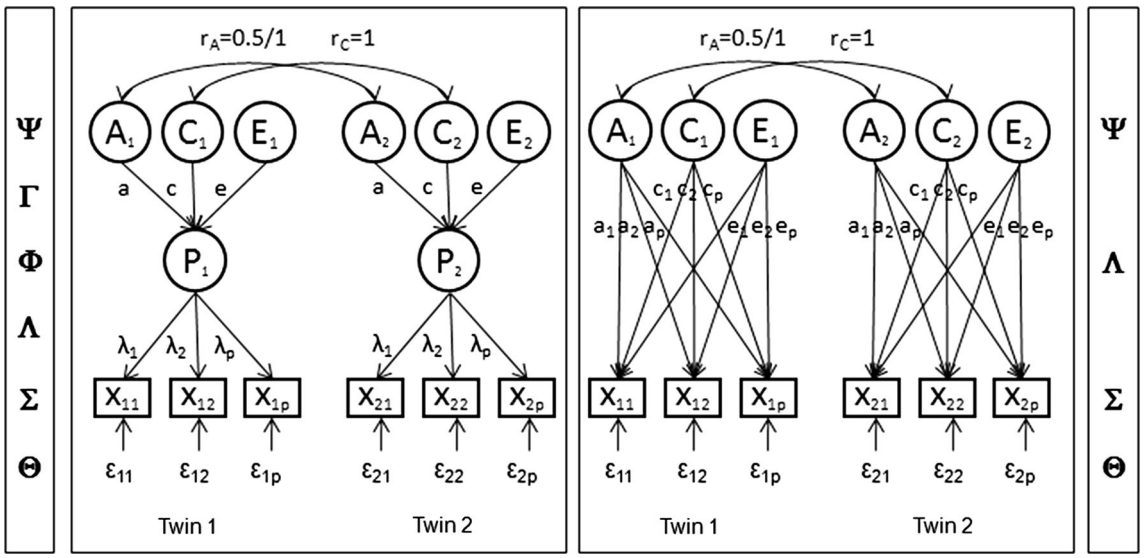

Fig. 2 A common pathway (*left*) and an independent pathway (*right*) genetic factor model. *Matrix names* on the sides correspond to the notation in the text

model (McArdle and Goldsmith 1990). This model is represented in the right panel of Fig. 2. In the independent pathway model, there is no phenotypic latent variable that mediates the genetic and environmental effects on the item responses. Rather, the A, C, and E factors influence item responses directly. In terms of the phenotypic (i.e., observed) covariance matrix of the item responses (i.e., $\mathbf{\Sigma}_{11} = \mathbf{\Sigma}_{22}$), we can convey the common and the independent pathway models, respectively, as follows:

$$\begin{aligned}\Sigma_{11} &= \Sigma_{22} = \Lambda\Phi\Lambda^{t} + \Theta_{cp} = \Lambda(\Phi_A + \Phi_C + \Phi_E)\Lambda^{t} + \Theta_{cp} \\ &= \Lambda\left(\Gamma_A\Gamma_A^{t} + \Gamma_C\Gamma_C^{t} + \Gamma_E\Gamma_E^{t}\right)\Lambda^{t} + \Theta_{cp} \\ \Sigma_{21} &= \Sigma_{12}^{t} = \Lambda(\Phi_A + \Phi_C)\Lambda^{t} + \Theta_{cp21} \\ &= \Lambda\left(r_A\Gamma_A\Gamma_A^{t} + \Gamma_C\Gamma_C^{t}\right)\Lambda^{t} + \Theta_{cp21}, \end{aligned} \quad (2)$$

and

$$\begin{aligned}\Sigma_{11} &= \Sigma_{22} = \Lambda_A\Psi_A\Lambda_A^{t} + \Lambda_C\Psi_C\Lambda_C^{t} + \Lambda_E\Psi_E\Lambda_E^{t} + \Theta_{ip} \\ &= \Lambda_A\Lambda_A^{t} + \Lambda_C\Lambda_C^{t} + \Lambda_E\Lambda_E^{t} + \Theta_{ip} \\ \Sigma_{21} &= \Sigma_{12}^{t} = r_A\Lambda_A\Lambda_A^{t} + \Lambda_C\Lambda_C^{t} + \Theta_{ip21}. \end{aligned} \quad (3)$$

Here $\mathbf{\Psi}_A$, $\mathbf{\Psi}_C$, and $\mathbf{\Psi}_E$ are the covariance matrices of the A, C, and E factors in the two models. In the common pathway model the covariance matrix of the psychometric factor, $\mathbf{\Phi}$, equals $\Phi_A + \Phi_C + \Phi_E$, i.e., $\Gamma_A\Gamma_A^{t} + \Gamma_C\Gamma_C^{t} + \Gamma_E\Gamma_E^{t}$, where $\mathbf{\Phi}_A$, $\mathbf{\Phi}_C$, and $\mathbf{\Phi}_E$ denote the A, C, and E variance components of $\mathbf{\Phi}$, and $\mathbf{\Gamma}_A$, $\mathbf{\Gamma}_C$, and $\mathbf{\Gamma}_E$ are the vectors of factor loadings $\mathbf{\Gamma}_A = [a]$, $\mathbf{\Gamma}_C = [c]$, $\mathbf{\Gamma}_E = [e]$. Note that in both models the diagonal matrices $\mathbf{\Theta}$ (denoted $\mathbf{\Theta}_{cp}$ and $\mathbf{\Theta}_{ip}$, as they may vary over the models) contain the residual variances of the items in the model. The residual covariance matrices may be subjected to their own decomposition, i.e., $\Theta = \Theta_A + \Theta_C + \Theta_E$ and $\Theta_{21} = r_A\Theta_A + \Theta_C$ (Neale and Cardon 1992).

In the present paper, we distinguish between genetic factor models (introduced above), and phenotypic factor models. By 'phenotypic factor model' we refer to the factor model as usually formulated and applied in psychological research. The term 'phenotypic' is used because the model is applied only to the observed (i.e., phenotypic) covariation; no genetic information is used.[3] The 8-factor cross-informant model of the CBCL and the 5-factor model of personality (McCrae and Costa 1999; McCrae and John 1992) are examples of a phenotypic factor model.

The common pathway model bears a number of similarities to the phenotypic factor model. Notably, both the phenotypic factor model and the common pathway model are based on the premise that all covariation in item responses is attributable to one or more latent variables. In phenotypic factor modeling, this can be formulated in terms of measurement invariance: influences of all external variables affecting covariation in item responses run only via the latent variable (Mellenbergh 1989; Meredith 1993). Likewise, in common pathway modeling one assumes that all of the A, C, and E influences on item covariation run only via the psychometric factor. That is, there are no direct effects of A, C, and E on the items. The assumption

[3] This is the standard application of the factor model to data collected in unrelated subjects, or when no information is available on genetic relatedness. We note, however, that if genome-wide DNA marker data are available in unrelated subjects, these could be used in a GTCA-like approach (Yang et al. 2011) to explore genetic covariance structures.

of full mediation of external influences by a latent variable has strong implications. For instance, different external variables affecting a set of item responses via the same latent variable exert the same magnitude of influence relative to each other on all the items that depend on that latent variable. For instance, if an A and a C variable affect a set of items via the same psychometric factor, then the magnitude of influence exerted by the variable A on any individual item will be a scalar multiple of the magnitude of influence exerted by the variable C on that same item, and this scalar multiple (k) will be a constant across all the items depending on this psychometric factor. This means that one can derive a common pathway model from an independent pathway model by imposing proportionality constraints on the factor loadings, such that $a_1/a_2 = c_1/c_2 = e_1/e_2 = k$ (following the notation in the right panel of Fig. 2).

Thus, the common pathway model makes explicit an assumption of the phenotypic latent variable model concerning the sources of item covariation—all influences on item covariation run via the phenotypic latent variable. This means, barring cases of model equivalence, that a latent variable model cannot hold unless the corresponding common pathway model holds (Franic et al. 2012a). Because any given latent variable hypothesis implies a corresponding common pathway model, a refutation of that common pathway model constitutes evidence against the latent variable hypothesis.

For this reason, one may test the latent variable hypothesis by comparing the fit of a common pathway model to that of a corresponding independent pathway model. Specifically, if a model in which all of the A, C, and E factors exert direct influence on the phenotype fits the data statistically better than a model in which these influences are mediated by a phenotypic latent variable, this would provide evidence against the hypothesis that the effects on the observed item covariation are completely mediated by the phenotypic latent variable. In that case the latent factors employed in the phenotypic factor model are no more than an amalgamation of the direct influences of the A, C, and E factors on the observed item responses. If, on the other hand, an independent pathway model does not fit the data better than the corresponding common pathway model, this would provide support for the structure employed in the common pathway model, and substantiation for the corresponding phenotypic latent variable model. Comparison of an independent pathway model and a common pathway model may be conducted using a likelihood ratio test, because, as shown above, a common pathway model can be derived from an independent pathway model by imposing appropriate proportionality constraints on the factor loadings (i.e., the models are nested).

Analyses

In the present analyses, the outlined methodology was used to examine the substantive interpretability of the Internalizing syndrome dimensions of the CBCL (Anxious/Depressed, Withdrawn, Somatic Complaints). The phenotypic dimensionality of the 31 items was assessed using exploratory (EFA) and confirmatory (CFA) factor analysis. In this part of the analyses, the data were treated as if the sample consisted of genetically unrelated individuals. As treating observations from the same family as independent may result in biased test statistics, we performed a correction for clustering available in MPlus, which has been shown to work well in this context (Rebollo et al. 2006). The EFA was performed using the oblique geomin rotation. Split-half validation was used, i.e., EFA was performed on one randomly selected half of the sample ($N = 8756$), and CFA on the other ($N = 8755$).

Based on the results of the phenotypic dimensionality assessment, a common pathway model was formulated: in this model, the phenotypic factors obtained in the EFA and the CFA were retained, and their variation decomposed into A, C, and E components, as illustrated in the top panel of Fig. 4. Subsequently, an independent pathway model was specified. This model is equal to the common pathway model in the number of the latent A, C, and E factors (i.e., the dimensions of the $\mathbf{\Psi}_A$, $\mathbf{\Psi}_C$, and $\mathbf{\Psi}_E$ matrices are equal across the two models), but it disposes of the phenotypic factors, i.e., it allows for the items to load directly on the A, C, and E factors. By comparing the fit of the common and the independent pathway model, we address the first focal question of whether one can interpret the syndrome dimensions of the CBCL Internalizing grouping substantively and causally.

To address the second research question, namely one concerning the dimensionality and the factor structure of the genetic and environmental effects that underlie the observed symptom covariation, independent pathway modeling was employed in an exploratory manner. First, the covariance matrix of the 31 symptoms was decomposed into A, C, and E components, i.e., the unconstrained 31×31 $\mathbf{\Sigma}_A$, $\mathbf{\Sigma}_C$, and $\mathbf{\Sigma}_E$ matrices (Eq. 1) were estimated. Subsequently, EFA was applied to each of these matrices to obtain an indication of their dimensionality. Examining the dimensionality and the factor structure of the genetic and environmental effects which jointly act to produce the observed symptom structure provides insight into the observed symptom covariation, as the structure emerging in the phenotypic analyses (EFA, CFA) depends directly on the structure and relative magnitude of the underlying genetic and environmental components; for instance, a strongly prevailing unidimensional C component will make the phenotypic structure appear unidimensional.

Table 1 Standardized factor loadings obtained in the phenotypic EFA and CFA

Item	EFA ($N = 8756$)							CFA ($N = 8755$)						
	3-Factor solution			4-Factor solution				3-Factor model			4-Factor model			
	AD	W	SC	D	A	W	SC	AD	W	SC	D	A	W	SC
Lonely	0.734	−0.068	0.001	0.670	−0.054	0.090	0.016	0.680			0.706			
Cries a lot	0.438	0.039	0.103	0.390	0.008	0.124	0.111	0.555			0.578			
Fears doing bad	0.393	0.165	0.049	0.309	0.471	−0.050	0.005	0.580				0.619		
Must be perfect	0.314	0.174	0.062	0.245	0.453	−0.042	0.019	0.499				0.530		
Feels unloved	0.988	−0.268	−0.055	0.885	0.008	−0.103	−0.061	0.708			0.736			
Others out to get him	0.852	−0.141	0.001	0.767	−0.006	0.013	0.004	0.695			0.725			
Feels worthless	0.819	0.012	−0.028	0.720	0.249	0.006	−0.052	0.791			0.821			
Nervous, tense	0.277	0.303	0.142	0.215	0.363	0.148	0.122	0.634				0.678		
Fearful, anxious	0.197	0.426	0.121	0.139	0.388	0.250	0.105	0.650				0.695		
Feels too guilty	0.499	0.183	0.046	0.411	0.470	−0.019	0.004	0.710				0.759		
Self-conscious	0.079	0.622	0.032	−0.002	0.569	0.347	−0.002	0.620				0.664		
Suspicious	0.628	0.032	0.028	0.556	0.030	0.139	0.037	0.640			0.664			
Sad	0.729	0.027	0.010	0.653	0.056	0.126	0.017	0.768			0.796			
Worries	0.484	0.191	0.100	0.406	0.366	0.061	0.072	0.740			0.767			
Rather be alone	0.141	0.408	−0.039	0.139	−0.048	0.474	0.006		0.542				0.541	
Would not talk	−0.001	0.675	−0.031	0.004	0.026	0.678	0.027		0.661				0.662	
Secretive	−0.019	0.812	−0.083	−0.033	0.090	0.806	−0.029		0.743				0.743	
Shy, timid	−0.191	0.795	−0.017	−0.228	0.424	0.562	−0.019		0.569				0.572	
Stares blankly	0.174	0.457	0.050	0.180	−0.096	0.546	0.112		0.668				0.667	
Sulks	0.403	0.065	0.106	0.356	0.000	0.154	0.117		0.547				0.546	
Lacks energy	0.134	0.339	0.050	0.150	−0.134	0.441	0.099		0.567				0.566	
Withdrawn	0.141	0.718	−0.145	0.138	0.026	0.740	−0.082		0.778				0.778	
Feels dizzy	0.205	−0.011	0.478	0.193	−0.001	0.031	0.476			0.620				0.620
Overtired	0.216	0.164	0.318	0.197	0.035	0.192	0.329			0.752				0.753
Aches, pains	0.026	0.018	0.726	0.023	0.038	0.015	0.720			0.778				0.778
Headaches	−0.050	0.024	0.709	−0.042	0.017	0.022	0.703			0.639				0.638
Nausea	0.022	−0.089	0.883	0.032	−0.046	−0.043	0.876			0.741				0.741
Eye problems	−0.025	0.077	0.334	−0.015	0.001	0.078	0.337			0.380				0.380
Skin problems	−0.017	0.060	0.280	−0.023	0.072	0.023	0.274			0.302				0.302
Stomachaches	0.004	0.020	0.732	−0.007	0.120	−0.036	0.719			0.722				0.722
Vomiting	−0.049	−0.017	0.724	−0.040	−0.009	−0.004	0.718			0.609				0.609
Factor determinacies	0.963	0.963	0.963	0.953	0.873	0.935	0.946							

The analyses were performed using Mplus (Muthén and Muthén 2007), Mx (Neale 2000), and R (R Core Team 2013). Given the discrete nature of the items, we fitted discrete factor models (i.e., we assumed the discrete indicator variables to be a realization of a continuous normal[4] latent process, and modeled polychoric correlations; Flora and Curran 2004; Wirth and Edwards 2007) using the robust weighted least squares estimator (WLSMV; Muthén and Muthén 1998–2007). The polychoric correlations between the 31 items and between the 62 (31 per twin) items served as input in the phenotypic and the genetic factor analyses, respectively. In evaluating model fit, the comparative fit index (CFI), the Tucker Lewis index (TLI), and the root mean square error of approximation (RMSEA) were used. As both the sample size and the models employed were large, the Chi square statistic was of limited use as an overall fit measure (Jöreskog 1993), and was used only to test local hypotheses concerning comparisons of nested models, as these comparisons are associated with a smaller approximation error.

[4] Tests of departures from underlying bivariate normality indicated that the normality assumption was tenable for all items.

Results

Phenotypic analyses (EFA and CFA)

The results of phenotypic dimensionality assessment are presented in Fig. 3 and Table 1. EFA produced two well-fitting solutions: a 3- and a 4-factor solution (Fig. 3). Interestingly, in both solutions, the items of the Anxious/Depressed scale appear to cluster into those pertaining to anxiety ('Fears doing something bad', 'Must be perfect', 'Nervous, tense', 'Fearful, anxious', 'Feels too guilty', 'Self-conscious') and those pertaining to depression ('Lonely', 'Cries a lot', 'Feels unloved', 'Others out to get him', 'Feels worthless', 'Suspicious', 'Sad', 'Worries'). In contrast to the Anxious/Depressed scale, the Somatic Complaints scale displayed a clearly unidimensional structure. The same is true of the Withdrawn scale, with the exception of the item 'Sulks', which consistently clustered with the items pertaining to depression, and the item 'Shy, timid', which in the 4-factor solution cross-loaded highly on the 'Anxious' factor.

The 4-factor solution, in which anxiety and depression form separate clusters, and the standard CBCL cross-informant model containing the Anxious/Depressed, Withdrawn, and Somatic Complaints scales, were subsequently tested in CFA. The models and the fit measures are shown in Fig. 3. As can be seen from the Figure, the two models differed only minimally in terms of model fit: CFI = 0.877 versus 0.891, TLI = 0.944 versus 0.950, RMSEA = 0.037 versus 0.035 for the 3- versus the 4-factor models, respectively. In the light of the well-established difficulty in distinguishing phenotypically the dimensions of anxiety and depression, this finding is perhaps not entirely unexpected.

Genetic covariance structure modeling

Based on the results of the phenotypic dimensionality assessment, a 3- and a 4-factor common pathway model were formulated. These are depicted in the top panel of Fig. 4. In both models, the common factors obtained in the phenotypic analyses (Anxious/Depressed, Withdrawn, and Somatic Complaints for the 3-factor model, and Anxious, Depressed, Withdrawn, and Somatic Complaints for the 4-factor model) were retained, and the contributions of the A, C, and E factors to their variation were assessed. As can be seen in Fig. 4, the fit of the two common pathway models was virtually indistinguishable: CFI = 0.947 versus 0.952, TLI = 0.962 versus 0.966, RMSEA = 0.028 versus 0.026 for the 3- versus the 4-factor model, respectively.

Subsequently, based on the two common pathway models, the two independent pathway models depicted in the lower panel of Fig. 4 were formulated. In these models, the A, C, and E factors employed in the common pathway analyses were retained, but the psychometric factors were disposed of, i.e., the items were allowed to load directly on the A, C, and E factors. Again, the fit of the two independent pathway models was virtually indistinguishable: CFI = 0.977 versus 0.976, TLI = 0.982 versus 0.982, RMSEA = 0.019 versus 0.010 for the 3- versus the 4-factor-based model, respectively.

Addressing the first focal question of whether an independent pathway model fits the data appreciably better than a common pathway model, we compared the general fit of the models, and carried out likelihood ratio tests of the proportionality constraints mentioned above.[5] These tests revealed both the 3- and the 4-factor-based independent pathway models to fit the data better than their common pathway versions ($\chi^2 = 1554$, df = 24, $p < 0.0001$ for the 3-factor-based models, $\chi^2 = 1084$, df = 21, $p < 0.0001$ for the 4-factor-based models). This implies that the common pathway models, in which phenotypic latent variables mediate all of the A, C and E influences, fail to convey entirely accurately the genetic and environmental effects on the items.

In the second set of analyses, EFA was employed to evaluate separately the dimensionality and the factor structure of the genetic and environmental influences that underlie the observed symptom covariation. Specifically, we evaluated the dimensionalities of the $\mathbf{\Sigma}_A$, $\mathbf{\Sigma}_C$ and $\mathbf{\Sigma}_E$ covariance matrices given in Eq. 1. The results are shown in Fig. 5. An inspection of the scree plots in the Figure clearly indicates a 1-dimensional C structure. The structures of A and E matrices remain, however, somewhat less clear. To explore the A and E structures further, the present EFA results were used as a basis for specifying a number of competing independent pathway models with varying dimensionalities of the $\mathbf{\Sigma}_A$, $\mathbf{\Sigma}_C$ and $\mathbf{\Sigma}_E$ covariance matrices. An overview of these models, including the fit measures and inter-factor correlations, is given in Supplementary Table 1. Overall, a comparison of the models suggested a model with 2A, 1C, and 4E factors as the best-fitting model with acceptable inter-factor correlations (CFI = 0.978, TLI = 0.983, RMSEA = 0.018). This model is depicted in Fig. 6, and parameter estimates are given in Table 2. It should, however, be noted that the models did not differ considerably in terms of model fit; therefore the structure in Fig. 6 need not necessarily be conclusive. What the present results do strongly suggest, however, is a unidimensional C

[5] For WLSMV estimators the standard approach of taking the difference between Chi square values and the corresponding degrees of freedom is not appropriate because the chi-square difference is not chi-square distributed (Muthén and Muthén 1998–2007). We therefore performed chi-square difference testing using scaling correction factors (Satorra and Bentler 2001).

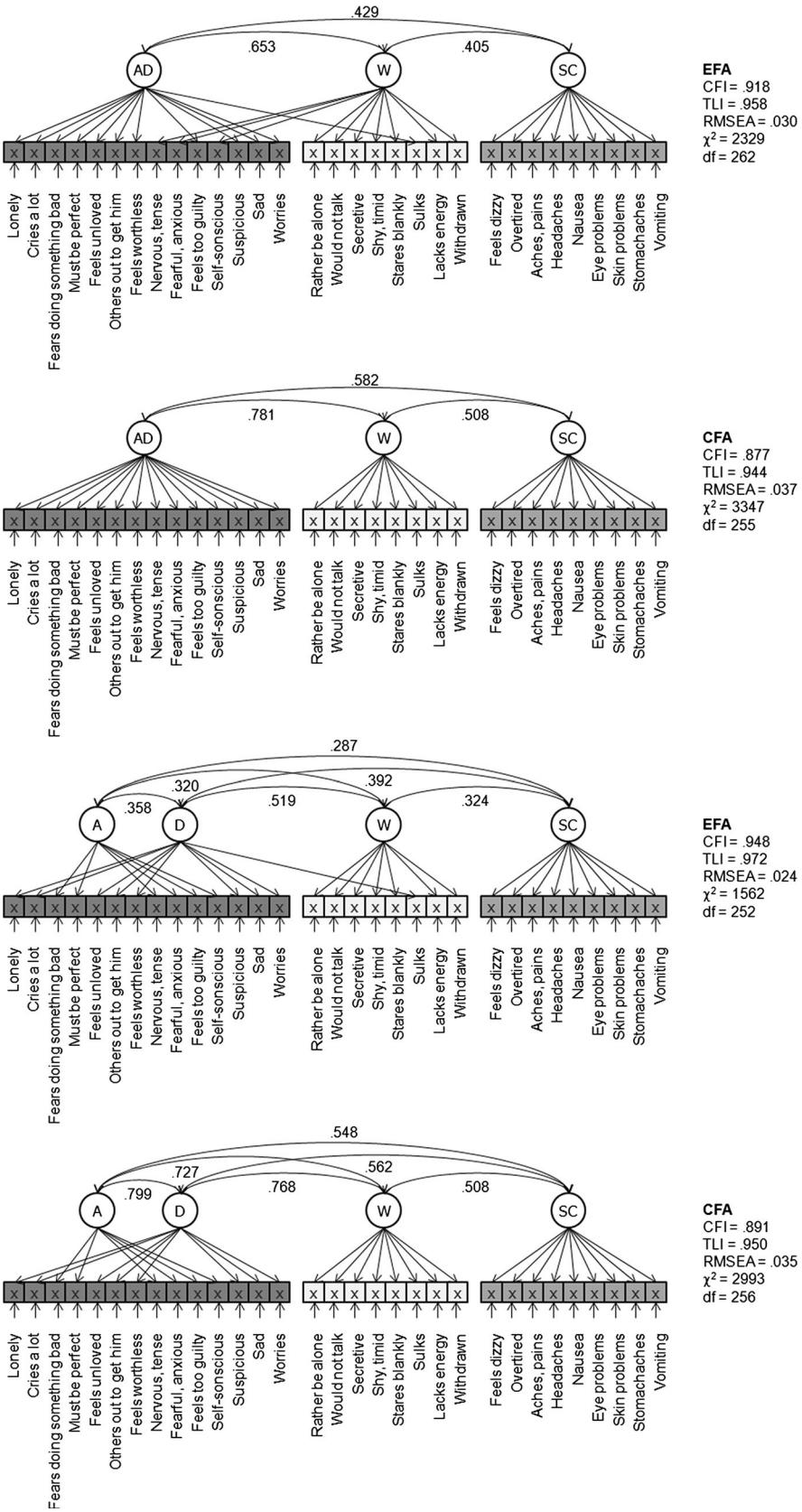

Fig. 3 Results of the phenotypic EFA and CFA. In EFA solutions only the highest factor loading for each item is depicted (the omitted factor loadings equal 0.057 on average; the depicted factor loadings equal 0.57 on average)

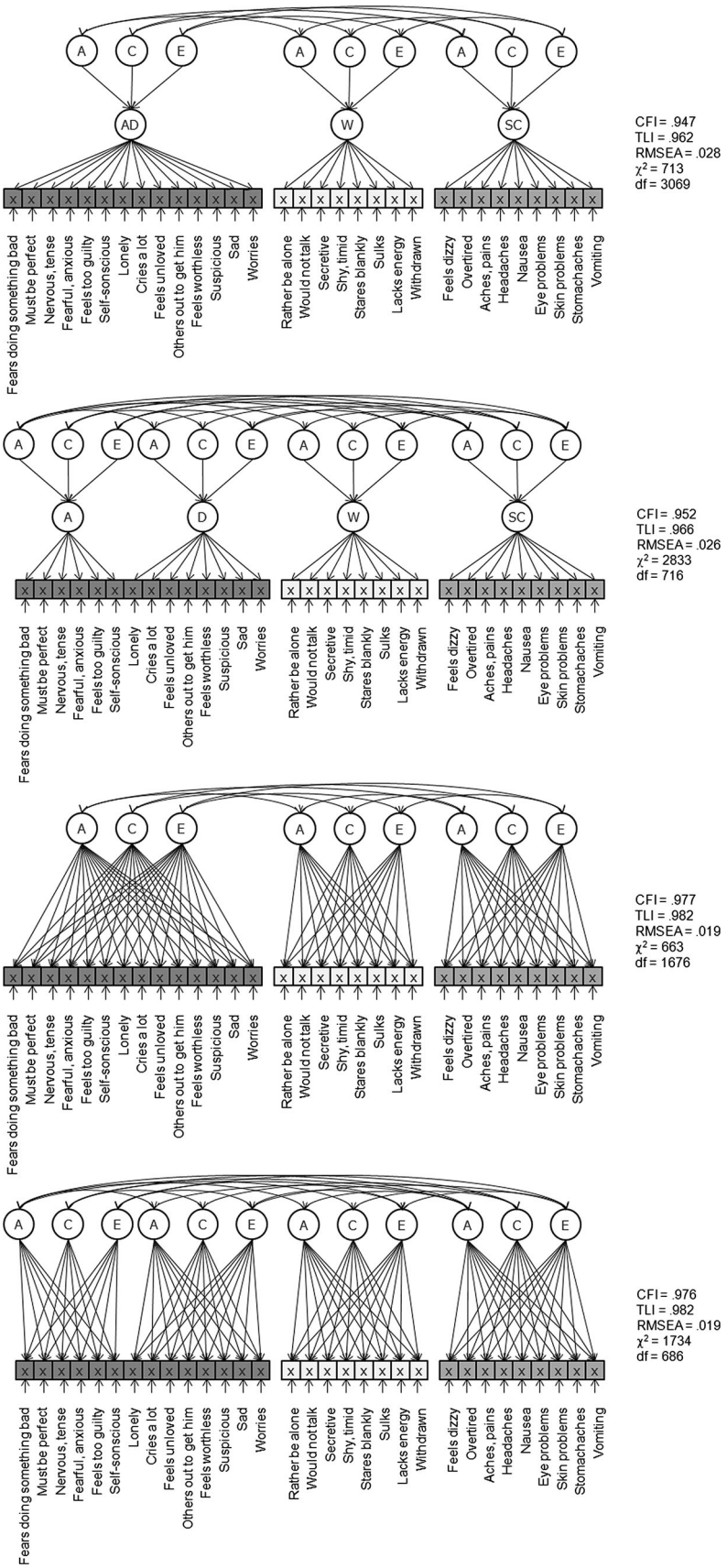

Fig. 4 The common (*upper panel*) and independent (*lower panel*) pathway models fitted to the data

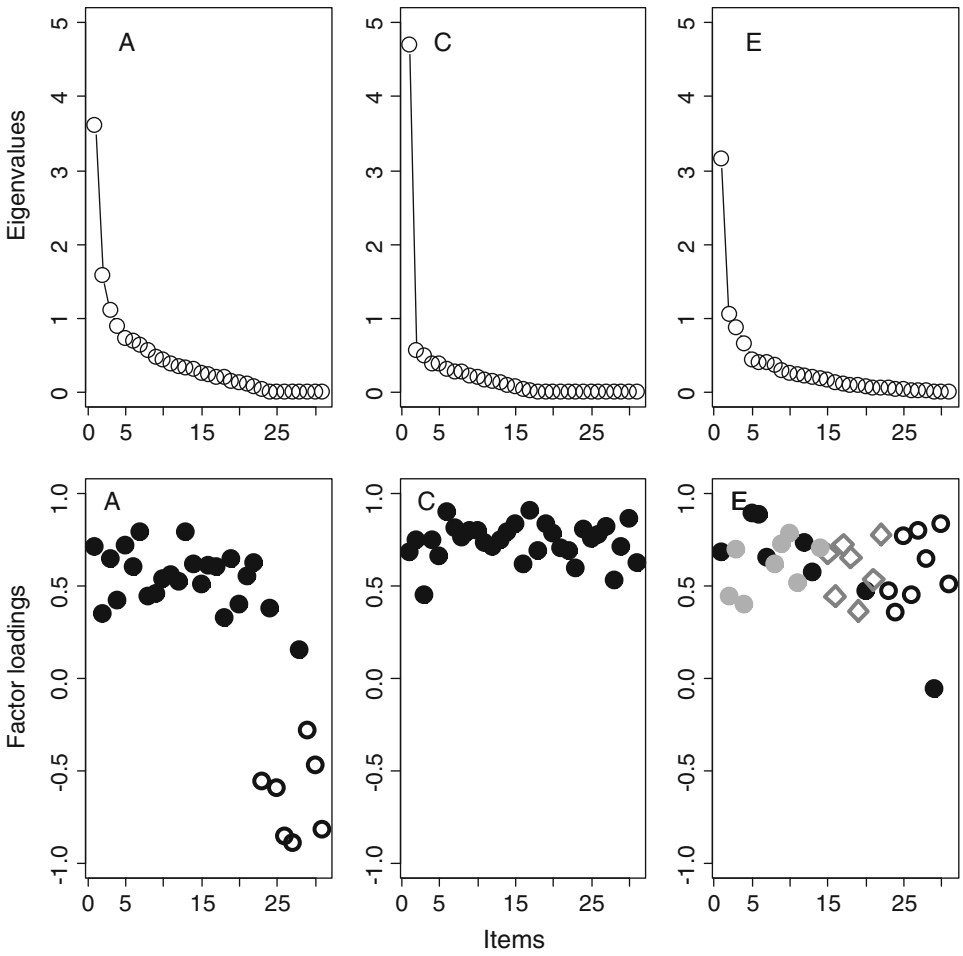

Fig. 5 Eigenvalues of the $\mathbf{\Sigma}_A$, $\mathbf{\Sigma}_C$, and $\mathbf{\Sigma}_E$ matrices (*upper panel*) and factor loadings obtained in EFA solutions with 2 A, 1 C, and 4 E factors (*lower panel*). *Colors/shapes* code for different latent factors. Only the highest factor loading for each item is shown

structure, and multidimensional (but mutually differing) A and E structures. These structures may also be discerned in Fig. 7, which gives a graphical representation of the $\mathbf{\Sigma_A}$, $\mathbf{\Sigma_C}$, and $\mathbf{\Sigma_E}$ covariance matrices (Epskamp et al. 2012).

Finally, the results of variance component estimation are given in Table 2. Overall, around 50 % of the variance in the CBCL Internalizing symptoms is explained by the common A, C, and E factors, the remaining half being due to residual (symptom-specific) factors. The overall symptom heritability (defined as the heritability due to both the common and the symptom-specific factors) is 50 % on average. The mean proportions of the phenotypic variance explained by the C and E factors are 20 and 30 %, respectively (last three columns Table 2). These proportions are relatively stable across all symptom clusters, with symptoms of depression being somewhat less heritable than the others (41 vs. 51–65 % on average). Interestingly, the high item heritability is predominantly due to the item-specific, rather than the common A factors, while the C component is primarily due to the common C factor, with the item-specific factors accounting for a negligible portion of the variance.

Discussion

The present article aimed at answering two principal questions: one pertaining to the ontological nature of the syndrome dimensions postulated in the CBCL cross-informant model, and the other pertaining to the factor structure of the genetic and environmental influences that underlie the observed symptom covariation.

The first question relates to a longstanding discussion in philosophy of science. The latent variable model, arguably the predominant measurement model in psychology, invariably invokes a latent variable which is hypothesized to underlie a set of observed variables (i.e., item responses, symptoms). The ontological nature of such latent variables

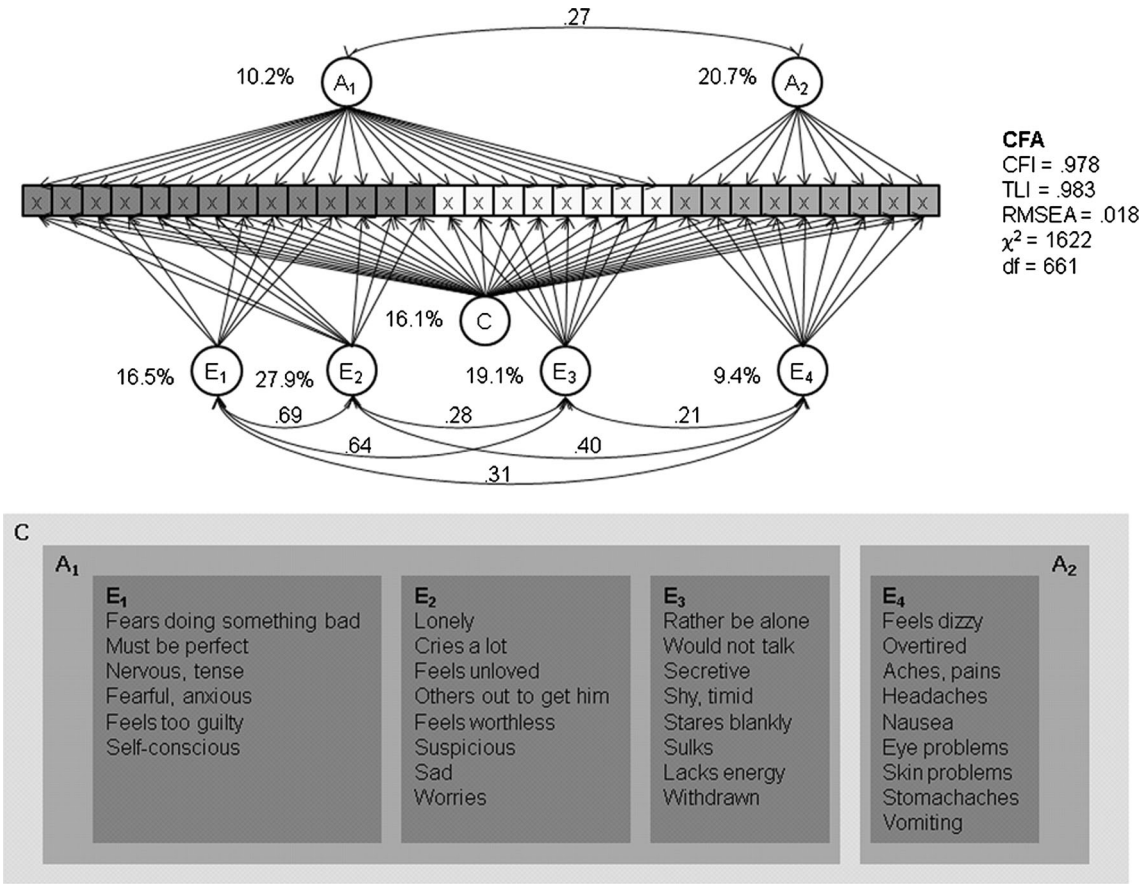

Fig. 6 The 2A 1C 4E independent pathway model. Item residuals are not depicted but are estimated in the model. The mean percentages of item variance explained by each factor are given. Items are listed below, and their allocation to factors is indicated by the *color of the panels*

has long been a subject of debate. On the theoretical side of the debate, broadly speaking, two principal (and mutually opposing) accounts of the latent variable are commonly invoked. In the realist view, the latent variable signifies a real entity which is assumed to exist independently of measurement, and is characterized by a causal relationship with its indicators: for instance, because a child is depressed, they exhibit symptoms such as excessive crying and feelings of sadness and worthlessness. The opposing, constructivist account, posits the latent variable as nothing more than a statistical construct used to simplify observations; in this view, this construct need not exist independently of measurement (Borsboom et al. 2003).

Empirical contributions to this debate have, to our knowledge, been scarce, although the existence and causal relevance of specific latent constructs such as depression and general intelligence have long been a source of controversy. Genetic factor modeling, as applied in the present article, may inform the discussion from an empirical perspective: by comparing the fit of a common pathway model, in which the latent phenotypic variables mediate all genetic and environmental effects on item covariation (the model therefore incorporating a realist hypothesis concerning the nature of those variables, or at least bring consistent with a realist perspective), to the fit of an independent pathway model (which bears no realist commitment regarding the phenotypic variable), one may test the latent variable hypothesis.

In the present case, neither the common pathway model featuring the three Internalizing syndrome dimensions of the CBCL (Anxious/Depressed, Withdrawn, and Somatic Complaints), nor the common pathway model postulating anxiety and depression as separate entities, survived confrontation with the independent pathway models. This invites reconsideration of the substantive interpretation of the dimensions in question, as it follows that these dimensions are better understood as a composite of unconstrained genetic and environmental influences than as well-defined entities that plausibly exist independently of measurement and statistical procedures (e.g., as natural kinds, Kendler et al. 2011).

This does not necessarily undermine the practical utility of the CBCL; we do not doubt its usefulness for diagnostic purposes, especially given that the broad structure found in

Table 2 The 2A 1C 4E model: Proportions of item variance explained by the common A, C, and E factors (first three columns), by the item-specific A, C, and E factors (next three columns), by all common factors relevant to the item $\left(\lambda^2_{\text{common}} = \lambda^2_{\text{A}} + \lambda^2_{\text{C}} + \lambda^2_{\text{E}}\right)$, by all residual factors relevant to the item $\left(\lambda^2_{\text{residual}} = \lambda^2_{\text{resA}} + \lambda^2_{\text{resC}} + \lambda^2_{\text{resE}}\right)$, and by all the A, C, and E factors relevant to the item, respectively $\left(\text{Total } \lambda^2_{\text{A}} = \lambda^2_{\text{A}} + \lambda^2_{\text{resA}}, \text{ etc.}\right)$

Item	λ^2_{A1}	λ^2_{C}	λ^2_{E2}	λ^2_{resA}	λ^2_{resC}	λ^2_{resE}	λ^2_{common}	$\lambda^2_{residual}$	Total λ^2_{A}	Total λ^2_{C}	Total λ^2_{E}
Lonely	0.01	0.17	0.36	0.37	0.02	0.08	0.54	0.46	0.37	0.18	0.44
Cries a lot	0.02	0.23	0.06	0.56	0.00	0.13	0.30	0.70	0.58	0.23	0.19
Feels unloved	0.00	0.16	0.45	0.39	0.00	0.00	0.61	0.39	0.39	0.16	0.45
Others out to get him	0.01	0.24	0.34	0.38	0.00	0.04	0.58	0.42	0.39	0.24	0.38
Feels worthless	0.05	0.21	0.44	0.29	0.00	0.01	0.70	0.30	0.34	0.21	0.45
Suspicious	0.04	0.30	0.10	0.37	0.08	0.11	0.44	0.56	0.41	0.38	0.21
Sad	0.03	0.27	0.30	0.36	0.01	0.03	0.60	0.40	0.39	0.28	0.33
Worries	0.11	0.21	0.19	0.31	0.10	0.07	0.51	0.49	0.42	0.32	0.27
$\overline{\lambda^2}$	0.03	0.22	0.28	0.38	0.03	0.06	0.54	0.46	0.41	0.25	0.34
			λ^2_{E1}								
Fears doing something bad	0.08	0.10	0.21	0.36	0.26	0.00	0.39	0.62	0.44	0.36	0.21
Must be perfect	0.06	0.08	0.18	0.42	0.00	0.26	0.32	0.68	0.48	0.08	0.44
Nervous, tense	0.12	0.19	0.11	0.40	0.00	0.18	0.42	0.58	0.52	0.19	0.29
Fearful, anxious	0.18	0.16	0.12	0.36	0.00	0.18	0.46	0.54	0.53	0.16	0.30
Feels too guilty	0.08	0.22	0.30	0.35	0.06	0.00	0.59	0.41	0.43	0.28	0.29
Self-conscious	0.42	0.08	0.07	0.25	0.00	0.17	0.58	0.42	0.67	0.09	0.24
$\overline{\lambda^2}$	0.16	0.14	0.17	0.36	0.05	0.13	0.46	0.54	0.51	0.19	0.30
			λ^2_{E3}								
Rather be alone	0.04	0.09	0.22	0.37	0.00	0.28	0.35	0.65	0.41	0.09	0.50
Would not talk	0.14	0.15	0.15	0.40	0.16	0.00	0.44	0.56	0.54	0.31	0.15
Secretive	0.21	0.11	0.32	0.31	0.00	0.05	0.64	0.36	0.52	0.11	0.37
Shy, timid	0.43	0.03	0.04	0.33	0.00	0.17	0.50	0.50	0.77	0.03	0.20
Stares blankly	0.04	0.21	0.23	0.48	0.00	0.05	0.47	0.53	0.51	0.21	0.28
Sulks	0.03	0.25	0.02	0.49	0.00	0.21	0.30	0.70	0.52	0.25	0.23
Lacks energy	0.02	0.17	0.12	0.66	0.00	0.03	0.31	0.69	0.68	0.17	0.15
Withdrawn	0.14	0.09	0.43	0.37	0.00	0.00	0.65	0.37	0.51	0.09	0.43
$\overline{\lambda^2}$	0.13	0.14	0.19	0.43	0.02	0.10	0.46	0.54	0.56	0.16	0.29
	λ^2_{A2}		λ^2_{E4}								
Feels dizzy	0.07	0.22	0.08	0.25	0.04	0.34	0.37	0.63	0.32	0.26	0.43
Overtired	0.03	0.37	0.08	0.41	0.00	0.11	0.47	0.53	0.44	0.37	0.19
Aches, pains	0.22	0.15	0.24	0.30	0.09	0.00	0.61	0.39	0.52	0.23	0.24
Headaches	0.26	0.09	0.11	0.23	0.00	0.30	0.47	0.53	0.49	0.09	0.42
Nausea	0.46	0.11	0.10	0.09	0.13	0.10	0.68	0.32	0.55	0.25	0.20
Eye problems	0.03	0.04	0.10	0.70	0.00	0.14	0.17	0.83	0.73	0.04	0.23
Skin problems	0.06	0.05	0.00	0.64	0.00	0.24	0.11	0.89	0.70	0.05	0.24
Stomachaches	0.30	0.14	0.13	0.25	0.00	0.19	0.56	0.44	0.55	0.14	0.31
Vomiting	0.44	0.10	0.00	0.00	0.33	0.12	0.54	0.46	0.44	0.44	0.12
$\overline{\lambda^2}$	0.21	0.14	0.09	0.32	0.07	0.17	0.44	0.56	0.53	0.21	0.27
Overall $\overline{\lambda^2}$	0.13	0.16	0.18	0.37	0.04	0.12	0.47	0.53	0.50	0.20	0.30

The $\overline{\lambda^2}$ rows give the mean proportion of item variance explained per item cluster (Depressed, Anxious, Withdrawn, and Somatic Complaints, respectively), and the Overall $\overline{\lambda^2}$ row gives the mean proportion of item variance explained across all items

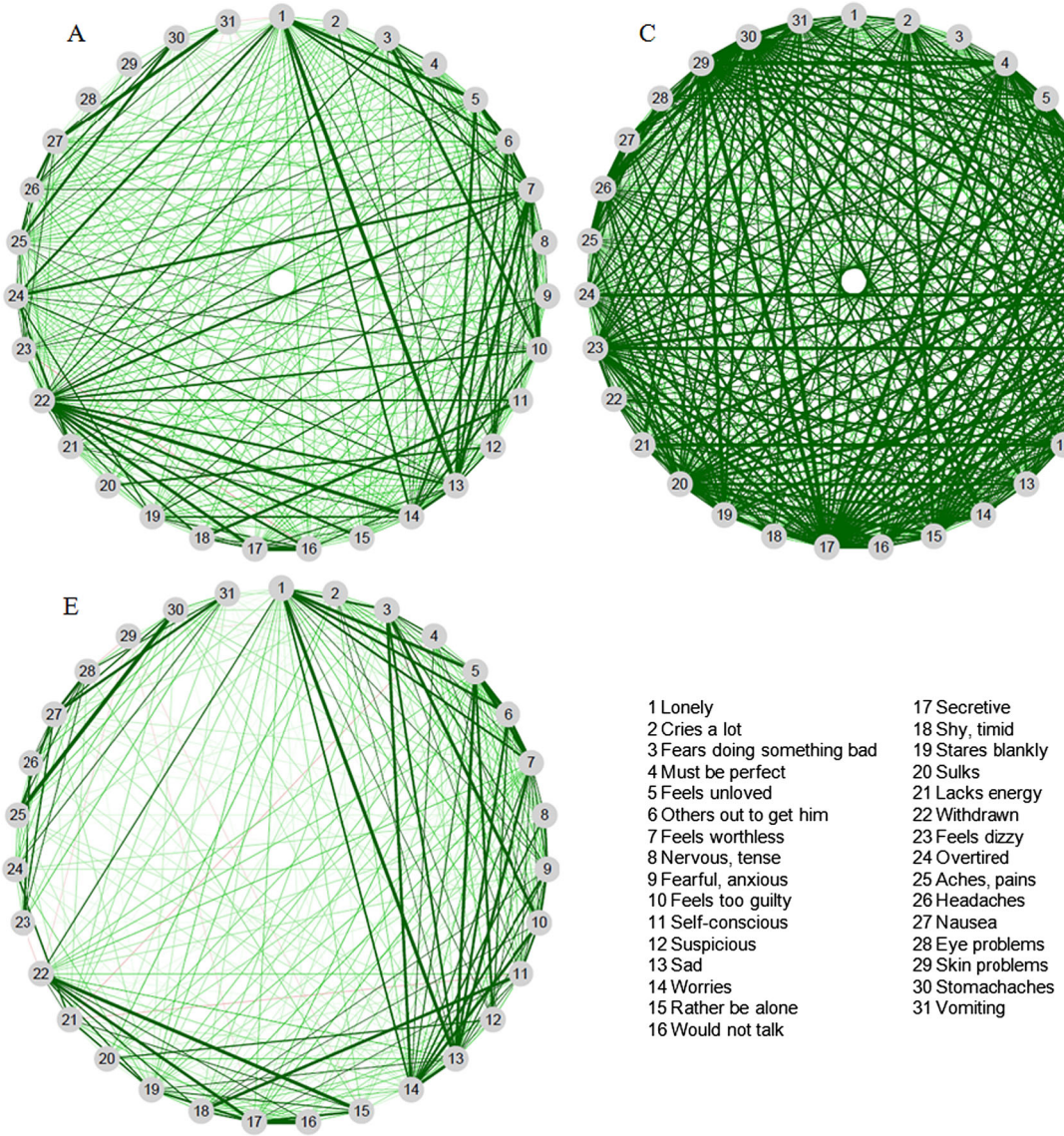

Fig. 7 Graphical representation of the correlation structures of the $\mathbf{\Sigma_A}$, $\mathbf{\Sigma_C}$, and $\mathbf{\Sigma_E}$ matrices. Nodes (i.e., *circles*) represent symptoms. The thickness of the edges (i.e., of the *lines connecting the nodes*) represents the strength of correlations between the symptoms. For instance, the thickness of the line connecting item 1 ("Lonely") to item 13 ("Sad") in the "A" *graph* represents the magnitude of the additive genetic correlation between these two symptoms

our analyses is in line with the current item allocation of the CBCL. Furthermore, the reasons for rejecting the common pathway structure may be local (due only to a subset of observed variables) and therefore the violation may be accommodated by addition of parameters or by removal of offending variables. What the present results do suggest, however, is that the three syndrome dimensions, as currently defined, do not appear to represent homogeneous entities in the Borsboom et al. (2003) sense, but are rather an amalgam of several different genetic and environmental structures. Clearly, the ascription of causal forces to such amalgams is problematic.

The second research question pertains to the structure of the genetic and environmental influences that give rise to the observed symptom covariation. Interestingly, the results suggest mutually differing additive genetic, common environmental and unique environmental structures. The 2-dimensional additive genetic structure distinctly affects symptoms of anxiety, depression, and withdrawal, on the one hand, and somatic complaints, on the other. The 4-dimensional unique environmental structure affects each of these symptom clusters distinctly, while the common environment acts uniformly across the entire range of internalizing symptoms. This partly replicates the findings of previous

multivariate investigations into the genetic and environmental sources of symptom covariation, which demonstrate a common genetic diathesis for anxiety and depression, with the distinction between these disorders being driven by the individual-specific environment (e.g., Kendler et al. 1987; Middeldorp et al. 2005; Kendler et al. 1992).

The present results put the aforementioned difficulties in delineating between the diagnostic categories of anxiety and depression into a clearer perspective. Anxiety and depression appear to share a common genetic basis: a single set of genes affects the individual differences in predisposition to developing general anxiety-, depression- and withdrawal-related symptomatology. Previous research and theoretical work have amply demonstrated a possibility of a general factor accounting for shared symptoms of anxiety, depression, and possibly more broad neurotic symptomatology (with more specific factors accounting for the specific subtypes of symptoms) (e.g., Clark and Watson 1991). This general factor can conceivably be identified with the shared genetic predisposition found in the present analyses. While this shared predisposition constitutes a broad genetic vulnerability which may predispose children to developing general internalizing symptomatology, the specific form of symptomatology (anxiety, depression or withdrawal) will depend on the children's unique environmental influences. The common family environment,[6] interestingly, exerts an overall protective or predisposing effect on the entire set of internalizing symptoms, either increasing or lowering the chance of developing internalizing psychopathology across the board.

If one takes into account not only the structure, but also the relative magnitude of the A, C, and E influences found in the present analyses, an illuminating picture emerges. Consider the set of items pertaining to anxious, depressed, and withdrawn behaviors. Under the model depicted in Fig. 6, this item set is influenced by a unidimensional A and a unidimensional C structure. These unidimensional latent structures, which act to make the symptoms act alike (i.e., covary), collectively explain around a quarter of their total phenotypic variance (10.2 and 17 % of the relevant item variance is explained by the A_1 and by the C factor, respectively). The factors which facilitate the clustering of these symptoms into three separate groups (in particular, the E_1, E_2, and E_3 factors) explain around 22 % of their phenotypic variance. The remainder ($\sim$50 %) of the phenotypic variance is explained by item-specific factors. Given the balance in the magnitude of influence that these mutually differing structures exert on the item set, the inability of the phenotypic modeling to distinguish between several different models is not surprising. In fact, one could wonder how the phenotypic analyses could converge on a single model, if several different models, each equally relevant to the phenotypic structure, are correct.

Finally, it should be mentioned that problems regarding the validity and reliability of children's self-reports and the consequent use of raters (parents, teachers) in the assessment of children's behavior may complicate assessment and subsequent interpretation. Rater bias (i.e., systematic effects on ratings originating from rater characteristics) is a widely recognized problem in research involving informants. In the context of twin and family studies, unmodeled rater bias is known to result in an overestimation of the shared environmental variance (Neale and Cardon 1992). Previous studies on internalizing symptoms have demonstrated a modest to nonexistent role of shared environment in the development of anxiety (Rapee et al. 2009; Gregory and Eley 2007; Hettema et al. 2001; Legrand et al. 1999), and a modest to moderate role of shared environment in the development of depression (Rice et al. 2002; Boomsma et al. 2005). Although this is consistent with the present findings, the extent to which our estimate of the shared environmental component is confounded by rater bias remains to be examined.

In summary, the present article utilized genetic item analyses to examine the ontology and the genetic and environmental etiology of the latent constructs 'Anxious/Depressed', 'Withdrawn', and 'Somatic Complaints', as defined in the CBCL/6–18 cross-informant model. The results (1) invite reconsideration of the substantive interpretation of these latent constructs, and (2) consistently with results of previous studies, demonstrate that additive genetics, common environment, and individual-unique environment each exert a distinct and mutually differing pattern of influence on internalizing symptoms. These results provide an informative context to the discussion on the phenotypic delineation between different syndromes or disorders, and contribute to our understanding of both the nature of the Internalizing syndrome dimensions and the etiology of internalizing behavior.

[6] See Carey (2009) for an alternative interpretation of C.

References

Achenbach TM, Rescorla LA (2001) ASEBA school age forms and profiles. ASEBA, Burlington

Bartels M, van Beijsterveldtl CEM, Derks EM, Stroet TM, Polderman TJC, Hudziak JJ, Boomsma DI (2007) Young Netherlands Twin Register (Y-NTR): a longitudinal multiple informant study of problem behavior. Twin Res Hum Genet 10(1):3–11

Bollen KA (1989) Structural equations with latent variables. Wiley, New York

Boomsma D, Van Beijsterveldt C, Hudziak J (2005) Genetic and environmental influences on Anxious/Depression during childhood: a study from the Netherlands Twin Register. Genes Brain Behav 4(8):466–481

Borsboom D, Mellenbergh GJ, van Heerden J (2003) The theoretical status of latent variables. Psychol Rev 110(2):203–219

Brady EU, Kendall PC (1992) Comorbidity of anxiety and depression in children and adolescents. Psychol Bull 111(2):244

Brown TA (1996) Validity of the DSM-III-R and DSM-IV classification systems for anxiety disorders. In: Rapee RM (ed) Current controversies in the anxiety disorders. Guilford Press, New York, pp 21–45

Carey G (2009) The problem with c2. In: Annual meeting of the Behavior Genetics Association, Minneapolis, MN, USA

Clark LA (1989) The anxiety and depressive disorders: descriptive psychopathology and differential diagnosis. In: Kendall PC, Watson D (eds) Anxiety and depression: distinctive and overlapping features. Academic Press, New York, pp 83–129

Clark LA, Watson D (1991) Tripartite model of anxiety and depression—psychometric evidence and taxonomic implications. J Abnorm Psychol 100(3):316–336

Eaves LJ (1983) Errors of inference in the detection of major gene effects on psychological test-scores. Am J Hum Genet 35(6):1179–1189

Epskamp S, Cramer AOJ, Waldorp LJ, Schmittmann VD, Borsboom D (2012) qgraph: network visualizations of relationships in psychometric data. J Stat Softw 48(4):1–18

Falconer DS, Mackay TFC (1996) Introduction to quantitative genetics. Longmans Green, Harlow

Flora DB, Curran PJ (2004) An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. Psychol Methods 9(4):466–491

Franić S, Dolan CV, Borsboom D, Hudziak JJ, Van Beijsterveldt CEM, Boomsma DI (2012a) Can genetics help psychometrics? Improving dimensionality assessment through genetic factor modeling. Psychol Methods 18(3):406–433

Franić S, Dolan CV, Borsboom D, Boomsma DI (2012b) Structural equation modeling in genetics. In: Hoyle RH (ed) Handbook of structural equation modeling. Guilford Press, New York, pp 617–635

Gregory AM, Eley TC (2007) Genetic influences on anxiety in children: what we've learned and where we're heading. Clin Child Fam Psychol Rev 10(3):199–212

Hartman CA, Hox J, Auerbach J, Erol N, Fonseca AC, Mellenbergh GJ, Novik TS, Oosterlaan J, Roussos AC, Shalev RS, Zilber N, Sergeant JA (1999) Syndrome dimensions of the Child Behavior Checklist and the Teacher Report Form: a critical empirical evaluation. J Child Psychol Psychiatry 40(7):1095–1116

Heath AC, Jardine R, Eaves LJ, Martin NG (1989) The genetic-structure of personality. II. Genetic item analysis of the EPQ. Pers Individ Differ 10(6):615–624

Hettema JM, Neale MC, Kendler KS (2001) A review and meta-analysis of the genetic epidemiology of anxiety disorders. Am J Psychiatry 158(10):1568–1578

Jöreskog KG (1993) Testing structural equation models. In: Bollen KA, Long SJ (eds) Testing structural equation models. SAGE, Newbury Park, pp 294–316

Keller MC, Coventry WL (2005) Quantifying and addressing parameter indeterminacy in the classical twin design. Twin Res Hum Genet 8(3):201–213

Kendler KS, Heath AC, Martin NG, Eaves LJ (1987) Symptoms of anxiety and symptoms of depression—same genes, different environments. Arch Gen Psychiatry 44(5):451–457

Kendler KS, Neale MC, Kessler RC, Heath AC, Eaves LJ (1992) Major depression and generalized anxiety disorder—same genes, (partly) different environments. Arch Gen Psychiatry 49(9):716–722

Kendler K, Zachar P, Craver C (2011) What kinds of things are psychiatric disorders? Psychol Med 41(06):1143–1150

Kline RB (2005) Principles and practice of structural equation modeling. Guilford Press, New York

Legrand LN, McGue M, Iacono WG (1999) A twin study of state and trait anxiety in childhood and adolescence. J Child Psychol Psychiatry 40(6):953–958

Martin NG, Eaves LJ (1977) Genetic-analysis of covariance structure. Heredity 38:79–95

Mather K, Jinks JL (1971) Biometrical genetics. Chapman and Hall, London

McArdle JJ, Goldsmith HH (1990) Alternative common factor models for multivariate biometric analyses. Behav Genet 20(5):569–608

McCrae RR, Costa PT (1999) A five-factor theory of personality. Guilford Press, New York

McCrae RR, John OP (1992) An introduction to the 5-factor model and its applications. J Pers 60(2):175–215

Mellenbergh GJ (1989) Item bias and item response theory. Int J Educ Res 13(2):127–143

Meredith W (1993) Measurement invariance, factor-analysis and factorial invariance. Psychometrika 58(4):525–543

Middeldorp CM, Cath DC, Van Dyck R, Boomsma DI (2005) The co-morbidity of anxiety and depression in the perspective of genetic epidemiology. A review of twin and family studies. Psychol Med 35(5):611–624

Mineka S, Watson D, Clark LA (1998) Comorbidity of anxiety and unipolar mood disorders. Annu Rev Psychol 49(1):377–412

Muthén LK, Muthén BO (1998–2007) Mplus user's guide. Muthén & Muthén, Los Angeles

Muthén LK, Muthén BO (2007) MPlus (5.1) [Computer software]. Muthén & Muthén, Los Angeles

Neale MC (2000) MxGui (1.7.03) [Computer software]. Virginia Commonwealth University, Richmond, VA

Neale MC, Cardon L (1992) Methodology for genetic studies of twins and families. Kluwer Academic Publishers B.V., Dordrecht

Neale MC, Lubke G, Aggen SH, Dolan CV (2005) Problems with using sum scores for estimating variance components: contamination and measurement noninvariance. Twin Res Hum Genet 8(6):553–568

R Core Team (2013) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. http://www.R-project.org/

Rapee RM, Schniering CA, Hudson JL (2009) Anxiety disorders during childhood and adolescence: origins and treatment. Annu Rev Clin Psychol 5:311–341

Rebollo I, de Moor MHM, Dolan CV, Boomsma DI (2006) Phenotypic factor analysis of family data: correction of the bias due to dependency. Twin Res Hum Genet 9(3):367–376

Rice F, Harold G, Thapar A (2002) The genetic aetiology of childhood depression: a review. J Child Psychol Psychiatry 43(1):65–79

Satorra A, Bentler PM (2001) A scaled difference Chi square test statistic for moment structure analysis. Psychometrika 66(4):507–514

van Beijsterveldt CEM, Groen-Blokhuis M, Hottenga JJ, Franić 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 K-J, Smit D, Dolan CV, Middeldorp CM, de Geus EJC, Bartels M, Boomsma DI (2012) 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 den Berg SM, Glas CAW, Boomsma DI (2007) Variance decomposition using an IRT measurement model. Behav Genet 37(4):604–616

van Dongen JP, Draisma HHM, Martin NG, Boomsma DI (2012) The continuing value of twin studies in the omics era. Nat Rev Genet 13(9):640–653

Waller NG, Reise SP (1992) Genetic and environmental-influences on item response pattern scalability. Behav Genet 22(2):135–152

Wirth RJ, Edwards MC (2007) Item factor analysis: current approaches and future directions. Psychol Methods 12(1):58–79

Yang J, Lee SH, Goddard ME, Visscher PM (2011) GCTA: a tool for genome-wide complex trait analysis. Am J Hum Genet 88(1):76