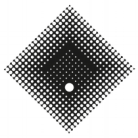
VICTORIA UNIVERSITY
MELBOURNE AUSTRALIA

Re-evaluation of the Latent Structure of Common Childhood Disorders: Is There a General Psychopathology Factor (P-Factor)?

This is the Published version of the following publication

Gomez, Rapson, Stavropoulos, Vasileios, Vance, Alasdair and Griffiths, Mark (2019) Re-evaluation of the Latent Structure of Common Childhood Disorders: Is There a General Psychopathology Factor (P-Factor)? International Journal of Mental Health and Addiction, 17 (2). pp. 258-278. ISSN 1557-1874

The publisher's official version can be found at
https://link.springer.com/article/10.1007%2Fs11469-018-0017-3
Note that access to this version may require subscription.

Downloaded from VU Research Repository https://vuir.vu.edu.au/39641/

International Journal of Mental Health and Addiction (2019) 17:258–278
https://doi.org/10.1007/s11469-018-0017-3

ORIGINAL ARTICLE

Re-evaluation of the Latent Structure of Common Childhood Disorders: Is There a General Psychopathology Factor (P-Factor)?

Rapson Gomez[1] · Vasileios Stavropoulos[1] · Alasdair Vance[2] · Mark D. Griffiths[3]

Published online: 8 November 2018
© The Author(s) 2018

Abstract

In the field of psychopathology, there is high comorbidity between different disorders. Traditionally, support for two broad correlated dimensions of internalizing and externalizing symptoms has consistently emerged for children and adolescents. To date, oblique 2 and 3 first-order factor models (factors for externalizing and internalizing, and fear, distress, and externalizing) and bi-factor models with the corresponding two and three group factors have been suggested for common internalizing and eternalizing child and adolescent disorders. The present study used confirmatory factor analyses to examine the relative support for these models in adolescents (≥ 12 to 18 years; $N = 866$) and children (6 to < 12 years; $N = 1233$) and the reliability and convergent and divergent validities of the psychopathology factor (P-factor) and group factors in the optimum bi-factor model. All participants were from a clinic and underwent Diagnostic and Statistical Manual of Mental Disorders, 4th Edition clinical diagnosis. The findings showed that the bi-factor model with two group factors (internalizing and externalizing) was the optimum model for both children and adolescents. For both groups, findings showed relatively higher reliability for the P-factor than the group factors, although the externalizing group factor showed substantial reliability in adolescents, and both the externalizing and internalizing group factors also showed substantial reliability in children. The factors of the optimum bi-factor model also showed good convergent and discriminant validities. The implications for theory and clinical and research practice related to psychopathology are discussed.

Keywords Psychopathology · Confirmatory factor analysis · P-factor · Internalizing and externalizing disorders · Reliability · Validity · Transdiagnostic approach

✉ Mark D. Griffiths
mark.griffiths@ntu.ac.uk

[1] Cairnmillar Institute, Melbourne, Australia

[2] The University of Melbourne, Melbourne, Australia

[3] Nottingham Trent University, Nottingham, UK

A robust finding for psychopathology (viewed either categorically or dimensionally) is the high comorbidity/co-occurrence across the different disorders and syndromes (Angold et al. 1999; Angold and Costello 2009; Krueger 1999; Krueger and Markon 2006). Traditionally, and based on factor analysis studies, support for two broad correlated dimensions of internalizing and externalizing symptoms has emerged consistently for children and adolescents (Achenbach 1966; Achenbach and Edelbrock 1978; Lahey et al. 2015; Patalay et al. 2015; Tackett et al. 2013). The internalizing dimension includes behaviors that have the propensity to express distress inwards, such as mood and anxiety disorders. In children and adolescents, the externalizing dimension includes behaviors that have the propensity to express distress outwards (Krueger 1999; Krueger and Finger 2001), such as the symptoms in attention deficit hyperactivity disorder (ADHD), oppositional defiant disorder (ODD), and conduct disorder (CD). Other researchers have proposed that the broad dimension of internalizing is better considered to comprise separate, but correlated, dimensions for fear-related behaviors/disorders or syndromes and distress-related behaviors/disorders or syndromes (Slade and Watson 2006; Vollebergh et al. 2001). Thus, psychopathology is viewed in this model in terms of three correlated broad dimensions: fear, distress, and externalizing (Martel et al. 2017; Lahey et al. 2012). More recently, other broad dimensions have been suggested, such as psychotic symptoms (Stochl et al. 2015), thought problems (Carragher et al. 2016; Caspi et al. 2014; Laceulle et al. 2015), and autism spectrum-related problems (Noordhof et al. 2015). Therefore, overall, at least two to four correlated broad dimensions have been proposed for psychopathology.

Independent of the different numbers of broad dimensions involved, a robust finding in these studies is that the proposed dimensions are consistently moderately to highly correlated with each other (Achenbach and Rescorla 2001; Angold and Costello 2009; Angold et al. 1999; Krueger and Markon 2006). For example, among children and adolescents, the correlations between internalizing and externalizing factors appear to range between 0.40 and 0.60 (Achenbach and Rescorla 2001). Similar correlations have been found for fear, distress, and externalizing factors (Angold and Costello 2009; Angold et al. 1999) and for the inter-correlation between internalizing, externalizing, and thought disorder problems (Lahey et al. 2004; Wright et al. 2013).

The high inter-correlations between the various broad psychopathology dimensions raise the possibility that an even broader overall psychopathology dimension could exist that potentially explains the co-variances found between them. Since 2012, a growing number of studies have examined this possibility using a type of confirmatory factor analysis (CFA) model called the bi-factor model (Reise 2012). A conventional bi-factor model is an orthogonal first-order factor model with a general factor in which all items (usually) load along with separate group factors for the different dimensions, after removing the variances accounted by the general factor. In such a model, the general factor captures the covariance across all the items, and the group factors capture the unique covariance of the items within the relevant dimensions, after accounting for their variance due to the general factor (Reise 2012). Thus, when applied to set of psychopathology constructs, the general factor or more specifically, the general psychopathology factor (or P-factor; Caspi et al. 2014) reflects the covariance across all the items (be they classified as problem behaviors, symptoms, disorders and/or syndromes) forming the broad dimensions (such as internalizing symptoms/disorders), and the group factors reflect the unique covariance of the broad dimensions, after accounting for the variance allocated to the general P-factor.

For a bi-factor model or a higher order factor model, the appropriate internal consistency reliability indices for the general factor and the group factors are omega hierarchical (ω_h) and

omega-subscale respectively (Brunner et al. 2012; Zinbarg et al. 2005). The values for ω_h and ω_t range from 0 to 1, with 0 indicating no reliability and 1 reflecting perfect reliability. According to Reise et al. (2013a, b), ω_h and ω_t values of at least 0.75 are preferred for meaningful interpretation of a scale. A number of other fit indices have also been proposed that could enable a more sophisticated and accurate interpretation of the dimensionality of the general factor (or P-factor) in the bi-factor model. These include explained common variance (ECV; Reise et al. 2013a, b), percentage of uncontaminated correlations (PUC; Bonifay et al. 2015), and the index of construct reliability (*H*; Hancock and Mueller 2001).

To date, bi-factor models of psychopathology have been examined for adults (Caspi et al. 2014; Lahey et al. 2012; Stochl et al. 2015), adolescents (Carragher et al. 2016; Castellanos-Ryan et al. 2016; Laceulle et al. 2015; Noordhof et al. 2015; Patalay et al. 2015), children (Lahey et al. 2015; Martel et al. 2017), and adolescents and children together (Tackett et al. 2013). Given the focus of the present study, past studies involving children and adolescents (Carragher et al. 2015; Castellanos-Ryan et al. 2016; Laceulle et al. 2015; Lahey et al. 2015; Martel et al. 2017; Noordhof et al. 2015; Patalay et al. 2015; Tackett et al. 2013) are particularly relevant for the present study. A robust finding in past studies of the bi-factor model of psychopathology in children and adolescents is the support for the bi-factor model (Carragher et al. 2015; Castellanos-Ryan et al. 2016; Laceulle et al. 2015; Lahey et al. 2015; Martel et al. 2017; Noordhof et al. 2015; Patalay et al. 2015; Tackett et al. 2013). Independent of whether these studies included two group factors (Castellanos-Ryan et al. 2016; Laceulle et al. 2015; Lahey et al. 2015; Patalay et al. 2015; Tackett et al. 2013), or three (Carragher et al. 2015; Martel et al. 2017) or four group factors (Noordhof et al. 2015), they have consistently found that the bi-factor model fitted better than the corresponding two or three or four first-order oblique models (Carragher et al. 2015; Laceulle et al. 2015; Lahey et al. 2015; Martel et al. 2017; Patalay et al. 2015). Additionally, the P-factor in the bi-factor model has shown acceptable external validity. As examples, Martel et al. (2017) reported that the P-factor (but not the group factors) was significantly associated with global executive functioning, and Patalay et al. (2015) reported that the P-factor best predicted future psychopathology and academic attainment. To date, only two studies have examined reliability for the bi-factor model (i.e., Martel et al. 2017; Murray et al. 2016), and the findings have been mixed. Based on ω_h and ω_t values of at least 0.75 for meaningful interpretation of a scale, Martel et al. (2017) reported acceptable reliability for the P-factor (ω_h = 0.898). In contrast, Murray et al. (2016) reported unacceptable reliability for the P-factor (ω_h values ranging from 0.53 to 0.64 for eight age groups, ranging from 7 to 15 years).

Although there have been numerous studies of the bi-factor model of psychopathology in children and adolescents, there are limitations and omissions in the existing literature. First, to date, all relevant studies have been on community samples and have used dimensional scores (derived from questionnaires and rating scales) of children's and adolescents' problems. No study has examined clinic-referred samples, based on clinical diagnosis, to model their broad dimensions (such as internalizing and externalizing). Thus, it is uncertain how these findings are directly applicable to samples of children and adolescent referred to clinical settings and who are given clinical diagnoses. Indeed, it is possible that there could exist qualitatively different features in clinical and non-clinical levels for some psychopathologies, thereby raising concerns over measuring and using clinical traits in non-clinical population to understand the clinical level (Murray et al. 2016; Reise and Waller 2009). Second, there have been only two studies involving children (Lahey et al. 2015; Martel et al. 2017) and a third study combining children with adolescents (Tackett et al. 2013). Therefore, the data in this age group

are presently limited. Third, the reliability data for the bi-factor model of psychopathology is also limited, and they have all reported only ω_h and ω_t scores, and none of the other indices (such as, ECV, PUC, and *H*) that allow for a more sophisticated and accurate interpretation of the dimensionality of the general factor (in the present study's case, the P-factor) in the bi-factor model. Fourth, although Martel et al. (2017) examined the convergent and discriminant validity of the P-factor and group factors (fear, distress, externalizing) by investigating how these factors correlated with similar factors for maternal psychopathology, none of the other studies involving children and adolescents provided similar information. Thus, there is need for studies to include such an evaluation to reinforce or question these initial (though innovative) findings.

Given the aforementioned limitations and omissions on the bi-factor model of psychopathology in children and adolescents, the first aim of the current study was to use CFA to simultaneously examine the structure of the major childhood internalizing and externalizing disorders, based on interviews of parents of clinic-referred children in Australia. For Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM-IV) (also DSM-IV-TR), the common internalizing disorders for children and adolescents included in the present study were separation anxiety disorder (SAD), social phobia (SOP), specific phobia (SPP), panic disorder (PD), agoraphobia (AG), generalized anxiety disorder (GAD), obsessive–compulsive disorder (OCD), post-traumatic stress disorder (PTSD), dysthymia (DYTH), and major depressive disorder (MDD). The externalizing disorders included were ODD, CD, and ADHD. The present study focused on parent interviews for reasons explained in the "Methods" section below. Five different measurement models were examined. They were one-factor (model 1); two-factor oblique with primary factors for the internalizing disorders and the externalizing disorders (model 2); three-factor oblique model with primary factors for distress disorders, fear disorders and externalizing disorders (model 3); bi-factor model with orthogonal factors for a P-factor and group factors for the internalizing disorders and externalizing disorders (model 4); and bi-factor model with orthogonal factors for a P-factor and group factors for distress disorders, fear disorders, and externalizing disorders (model 5). All five models are shown in Figs. 1, 2, 3, 4, and 5. To allow examination of the robustness of the findings, all the models were examined in children and adolescents separately. A second aim of the study was to examine model-based reliabilities of the factors(s) in the optimum bi-factor model for these groups, contingent on such a model being supported. More specifically, the focus was to examine ω_h and ω_t, as well as ECV, PUC, and *H* reliability indices. A third aim of the study was to test the convergent and divergent validities of the factors in the optimum bi-factor model, contingent on such a model being supported. Based on the recent findings reported by Martel et al. (2017), support was expected (in terms of fit) for a bi-factor model (either with a P-factor and three group factors [fear, distress, externalizing] or a P-factor and two group factors [internalizing and externalizing]), and relatively good support in terms of reliability and validity for the P-factor is such a model.

Methods

Participants

The data for all participants were collected from archival files from the Academic Child Psychiatry Unit (ACPU) of the Royal Children's Hospital, Melbourne, Australia. The ACPU is

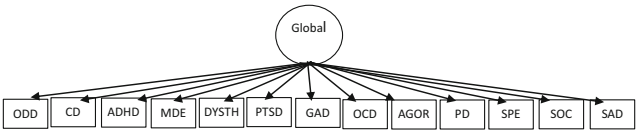

Fig. 1 1-factor (Model 1). ADISC-IV Disorders: *SAD* separation anxiety disorder, *SOC* social phobia, *SPE* specific Phobia, *PD* panic disorder, *AGOR* agoraphobia, *GAD* generalized anxiety disorder, OCD obsessive compulsive disorder, *PTSD* post traumatic stress disorder, *DYSTH* dysthymia, *MDE* major depressive disorder, *ADHD* attention seficit/hyperactivity disorder, *CD* conduct disorder, *ODD* oppositional defiant disorder

an outpatient psychiatric unit that provides services for children and adolescents with behavioral, emotional, and/or learning problems. Referrals are generally from other medical services, schools, and social and welfare organizations. The present study used the records of children, aged between 6 and 18 years, referred between 2004 and 2016, who had been interviewed for clinical diagnosis. In total, there were 2099 children, comprising 1504 males (71.8%) and 592 females (28.2%). The overall mean age of participants was 11.22 years ($SD = 3.10$). The frequencies of children (< 12 years) and adolescents were (≥ 12) were 1233 and 866, respectively (see Table 1).

Table 1 provides the sociodemographic characteristics and clinical diagnoses of participants in the study. As shown, most fathers were employed, and most mothers were mainly employed or involved in home duties. About two thirds of participants had mothers and fathers who had attended at least secondary schools, and most were from families with income less than $50,000 (AUS) per year. These figures correspond close to the Australian population. In terms of parental relationship, approximately 48.68% were living together, 43.6% were separated or divorced, and the remainder were single for other reasons (e.g., death of partner).

In relation to clinical disorders, externalizing disorders were highly prevalent, with around 75.3% and 66.8% of the participants having ADHD and ODD/CD, respectively (see Table 1). Among the internalizing disorders, GAD, SPP, DYTH, and SOP were more prevalent. Approximately 44.6%, 32.30%, 39.5%, and 32.2% of the participants were diagnosed with GAD, SPP, DYTH, and SOP, respectively. PD, PTSD, and AG were relatively rare. Table 2 shows the frequencies of different levels of comorbidity for the sample used in the present study. As shown of those with a clinical diagnosis, only 9% had no comorbidity. Approximately, 69.3% had one to five other disorders. Although details are not presented, for those with an anxiety disorder, 57% had a depressive disorder, and for those with a depressive disorder, 87.2% had an anxiety disorder. A total of 75% of children were comorbid for at least one externalizing and one internalizing disorder.

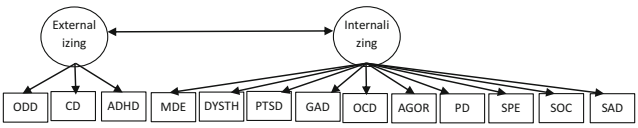

Fig. 2 2-factor oblique: primary factors = internalizing & externalizing disorders (Model 2). ADISC-IV Disorders: *SAD* separation anxiety disorder, *SOC* social phobia, *SPE* specific Phobia, *PD* panic disorder, *AGOR* agoraphobia, *GAD* generalized anxiety disorder, OCD obsessive compulsive disorder, *PTSD* post traumatic stress disorder, *DYSTH* dysthymia, *MDE* major depressive disorder, *ADHD* attention seficit/hyperactivity disorder, *CD* conduct disorder, *ODD* oppositional defiant disorder

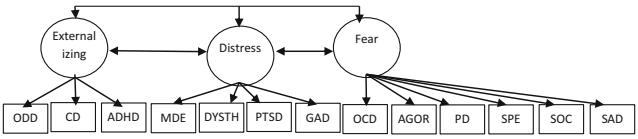

Fig. 3 3-factor oblique: primary factors = distress, fear & externalizing disorders (Model 3). ADISC-IV Disorders: *SAD* separation anxiety disorder, *SOC* social phobia, *SPE* specific Phobia, *PD* panic disorder, *AGOR* agoraphobia, *GAD* generalized anxiety disorder, OCD obsessive compulsive disorder, *PTSD* post traumatic stress disorder, *DYSTH* dysthymia, *MDE* major depressive disorder, *ADHD* attention seficit/hyperactivity disorder, *CD* conduct disorder, *ODD* oppositional defiant disorder

Measures

Anxiety Disorders Interview Schedule for Children Clinical diagnosis was based on the Anxiety Disorders Interview Schedule for Children (ADISC-IV; Silverman and Albano 1996). The ADISC-IV is a semi-structured interview, based on the DSM-IV/DSM-IV-TR diagnostic system (APA 2000). Although ADISC-IV has been designed primarily to facilitate the diagnosis of the major childhood internalizing disorders, it can also be used for diagnosing other major childhood disorders. ADISC-IV diagnoses do not take into account the hierarchical, exclusionary rules outlined by the DSM-IV for making diagnoses. The ADISC-IV guideline for diagnosis is that the child be given diagnosis of all disorders meeting the diagnostic criteria. There are different ADISC-IV versions for parent interview and for child interview, and clinical diagnosis can be based either on parent or child interview or on both interviews considered together. All diagnoses reported in this study were based on parent interviews, as the child interview version does not allow for the diagnoses of CD and ODD, both of which were part of the externalizing disorders modeled in the psychopathology measurement models evaluated in the present study. Additionally, it should be noted that there is evidence of poor levels of agreement for diagnosis between information across the child and parent versions of the ADISC-IV (Grills and Ollendick 2003) and that clinical interviews of children can lead to unreliable diagnosis (Jensen et al. 1999). The parent version of ADISC-IV has robust psychometric properties (Silverman et al. 2001). Test–retest reliability for the ADISC-IV scores over a 7- to 14-day interval has shown good-to-

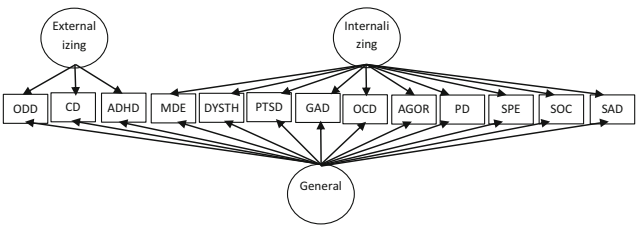

Fig. 4 Bifactor, general factor and internalizing and externalizing group factors (Model 4). ADISC-IV Disorders: *SAD* separation anxiety disorder, *SOC* social phobia, *SPE* specific Phobia, *PD* panic disorder, *AGOR* agoraphobia, *GAD* generalized anxiety disorder, OCD obsessive compulsive disorder, *PTSD* post traumatic stress disorder, *DYSTH* dysthymia, *MDE* major depressive disorder, *ADHD* attention seficit/hyperactivity disorder, *CD* conduct disorder, *ODD* oppositional defiant disorder

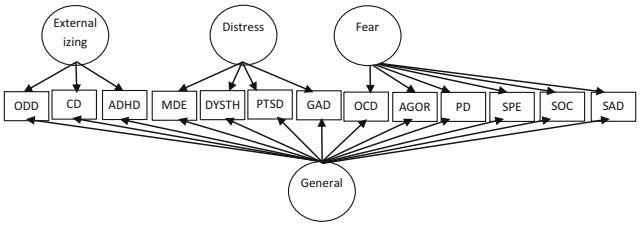

Fig. 5 Bifactor, one general factor and distress, fear and externalizing group factors (Model 5). ADISC-IV Disorders: *SAD* separation anxiety disorder, *SOC* social phobia, *SPE* specific Phobia, *PD* panic disorder, *AGOR* agoraphobia, *GAD* generalized anxiety disorder, OCD obsessive compulsive disorder, *PTSD* post traumatic stress disorder, *DYSTH* dysthymia, *MDE* major depressive disorder, *ADHD* attention seficit/hyperactivity disorder, *CD* conduct disorder, *ODD* oppositional defiant disorder

excellent reliability. Kappa values between interviewers for this version range from 0.65 to 1.00 (Silverman et al. 2001).

Child Behavior Checklist/6–18 The Child Behavior Checklist/6–18 (CBCL) is a measure in the Achenbach System of Empirically Based Assessment (Achenbach and Rescorla 2001). Completed by parents, it has 113 items and is used to rate children between 4 and 18 years of age. Respondents indicate the degree or frequency of each behavior described in the item on a scale of 0 (not true), 1 (somewhat or sometimes true), or 2 (very true or often true). The standard rating period for the CBCL is 6 months. The CBCL has excellent psychometric properties and includes scales for various behavior and emotional problems (Achenbach and Rescorla 2001). In addition, it provides two broad scores for internalizing behavior problems and externalizing behavior problems. These broad scores were used in the present study to examine the external validities of the factors in the optimum model.

Procedure

The study was approved by the RCH ethics committee as part of ACPU's comprehensive examination of children and adolescents referred for psychological problems. Each legal guardian and participant provided informed written consent for any data provided by them to be used in future research studies. This is a standard part of the ACPU assessment procedure. All children and their parents participated in separate interviews and testing sessions with breaks over 2 days. Information was also obtained from teachers using various checklists and questionnaires. In all cases, parental consent forms were completed prior to the assessment. The data collected covered a comprehensive demographic, medical, educational, psychological, familial, and social assessment of the child and child's family. All psychological data were collected by research assistants, who were students in clinical psychology, and under the supervision of registered psychologists. The research assistants were provided with extensive supervised training by their supervisors prior to them collecting data.

This training for the ADISC-IV included observations of it being administered by the psychologists. The research assistants commenced administering the ADISC-IV only after they attained competence in its administration, as assessed by their supervisors. There was adequate inter-rater reliability for the diagnoses made between the research assistants and the psychologists ($\kappa = 0.88$). Standard procedures were used for the administration of all measures.

Table 1 Demographic characteristics of participants

	Overall		Mother		Father	
	N	*N%*	*N*	*N%*	*N*	*N%*
Gender						
Male (frequency, percentage)	1504	71.8				
Female (frequency, percentage)	592	28.2				
Age (mean, *SD*)	11.22	3.10				
Age group						
Child (frequency, percentage)	1233	58.7				
Adolescence (frequency, percentage)	866	41.3				
Employment (frequency, percentage)						
Employed			1011	48.2	1466	69.8
Home duties			738	35.2	44	2.1
Pensioner			135	6.4	89	4.2
Unemployment			67	3.2	181	8.6
Others			110	5.2	97	4.6
Total			2061	98.2	1877	89.4
Missing			38	1.9	222	10.6
Education (frequency, percentage)						
Primary			32	1.5	59	2.8
Secondary			1160	55.3	1006	47.9
Technical/trade			329	15.7	392	18.7
Tertiary			522	24.9	401	19.1
			2043	97.3	1858	88.5
Missing			56	2.7	241	11.5
Family income (frequency, percentage)						
$0–$30,000	730	34.8				
$30,001–$40,000	233	11.1				
$40,000–$50,000	185	8.8				
> $50,000	784	37.4				
Total	1932	92.0				
Missing	169	8.0				
Parental relationship (frequency, percentage)						
Living together	1024	48.8				
Separated	543	25.9				
Divorced	371	17.7				
Death of one/both parents	64	3.1				
Others	60	2.9				
Total	2062	98.2				
Missing	37	1.8				
Diagnosis (frequency, percentage)						
Separation anxiety disorder	492	23.4				
Social phobia	679	32.3				
Specific phobia	684	32.6				
Panic disorder	229	10.9				
Agoraphobia	121	5.8				
Generalized anxiety disorder	936	44.6				
Obsessive–compulsive disorder	459	21.9				
Post-traumatic stress disorder	321	15.3				
Dysthymia	829	39.5				
Major depressive disorder	383	18.2				
Attention deficit/hyperactivity disorder	1581	75.3				
Oppositional defiant disorder	1402	66.8				
Conduct disorder	775	36.9				
No diagnosis	71	3.4				
Missing	0	0				

Table 2 Frequencies of different levels of comorbidity

Number of additional disorders	Number	*N*%	Cumulative *N*%
No disorder	71	3.4	3.4
0	188	9.0	12.3
1	296	14.1	26.4
2	379	18.1	44.5
3	312	14.9	59.4
4	257	12.2	71.6
5	209	10.0	81.6
6	136	6.5	88.0
7	114	5.4	93.5
8	80	3.8	97.3
9	33	1.6	98.9
10	18	.9	99.7
11	4	.2	99.9
12	2	.1	100.0

However, where necessary, researchers read the items to participants (approximately 5% of the sample). Approximately 95% of the parent ADISC-IV interviews involved mothers only, and the remainder involved fathers only or both fathers and mothers together. Using the categorical data from the parent ADISC-IV, clinical diagnosis was also determined by a consultant child psychiatrist, who independently reviewed these data. The inter-rater reliability for diagnoses (for 10% of the parent interviews) of the initial diagnosis and the consultant child psychiatrist was high (kappa values of 0.90).

Data Analysis

Software All the CFA models in the study were conducted using Mplus (version 7) software (Muthén and Muthén 2013).

Extraction As clinical diagnosis for each disorder resulted in binary scores (disorder present that was coded 1, and disorder not present that was coded 0), the mean and variance-adjusted weighted least squares (WLSMV) extraction was used for all the CFA analyses (Rhemtulla et al. 2012). This is a robust estimator, recommended for CFA with ordered-categorical scores, including dichotomous scores. The WLSMV estimator does not assume normally distributed variables. According to measurement experts, relative to other estimators, the WLSMV estimator provides the best option for modeling categorical data, including dichotomous data (Beauducel and Herzberg 2006; Lubke and Muthén 2004; Millsap and Yun-Tein 2004).

Model Fit For the CFA models, at the statistical level, model fit can be examined using χ^2 values (WLSMVχ^2 values in the current case). As all types of χ^2 values, including WLSMVχ^2, are inflated by large sample sizes, the fit of the models is generally interpreted by researchers using approximate fit indices, such as the root mean squared error of approximation (RMSEA), the comparative fit index (CFI), Tucker–Lewis Index (TLI), and the weighted root mean square residual (WRMR). For models based on maximum likelihood estimation, the guidelines suggested by Hu and Bentler (1998) are that RMSEA values close to 0.06 or below can be taken as good fit, close to 0.07 to < 0.08 as moderate fit, close to 0.08 to 0.10 as marginal fit, and close to > 0.10 as poor fit. For the CFI and TLI, values of 0.95 or

above are taken as indicating good model-data fit, and values of 0.90 and < 0.95 are taken as acceptable fit. The cutoff score for good fit suggested for WRMR is less than 0.90 (Yu and Muthen 2002). For the present study, these appropriate fit indices, rather than the χ^2 statistic, were used as evidence of model fit. However, it is worth noting that despite the widespread use of these indices and fit values, a simulation study by Nye and Drasgow (2011) concluded that appropriate indices cutoff values for WLSMV estimation can vary across conditions.

Reliability In relation to reliability, ω_h and ω_t, ECV, PUC, and H values were computed using the program developed by Watkins (2013) and cross-checked using the program developed by Dueber (2016). In a bi-factor model, the ECV of the general factor will be high and the ECV of the group factors will be low whenever there is little common variance beyond that of the general factor. High values for general factor indicate the presence of a strong general factor dimension (unidimensionality) in the bi-factor model (Reise et al. 2013a). In contrast, low ECV values for the general factor do not indicate support for the presence of a strong general factor dimension (unidimensionality) but support for a multidimensionality model.

A model-based internal consistency reliability that is analogous to alpha coefficient that is especially useful for bi-factor models is omega hierarchical when referring to the general factor (ω_h; Zinbarg et al. 2006) and the omega subscale (ω_t) when referring to the group factors. The ω_h can be interpreted as an estimator of how much variance in summed (standardized) scores can be attributed to the general factor (Brunner et al. 2012). The values for ω_h and ω_t range from 0 to 1, with 0 indicating zero reliability and 1 reflecting perfect reliability. For a bi-factor model, the percentage of uncontaminated correlations (PUC; Bonifay et al. 2015) indicates the bias that could result from forcing multidimensional data into a unidimensional model (Bonifay et al. 2015, p. 507). The ECV, the PUC, and ω_h and ω_t values can be examined concurrently to decide if the indicators in a bi-factor model can be interpreted as having sufficient reliability to view them as essentially unidimensional or if they should be considered multidimensional. According to Reise et al. (2013a, b), if PUC > 0.80, then the indicators in the bi-factor model, if supported, can be interpreted as primarily unidimensional. When the PUC < 0.80, then such an interpretation requires ECV > 0.60 and ω_h > 0.70. Failure to meet these criteria would mean that a multidimensional interpretation is warranted for the set of indicators, even if the bi-factor measurement model shows good fit. H is an index of construct reliability or replicability to estimate the reliability of the underlying group and general factors (Hancock and Mueller 2001). According to Rodriguez et al. (2016), H values > 0.80 are indicative of a stable well-defined latent variable, and values < 0.70 are indicative of a factor that is not worth specifying.

Convergent and Divergent Validities To test the convergent and divergent validities of the factors in the optimum model, the broad CBCL internalizing and externalizing scores were regressed on the factors of the optimum model.

Results

Missing Values and Fit of the Null Model

There were no missing values for the clinical cases used in the present study.

Fit of the Models Tested in the Present Study

Table 3 shows the results of all the CFA models tested for children and adolescents separately. Based on guidelines proposed by Hu and Bentler (1998), all fit indices for the one-factor model (M1s) in both groups showed poor fit. For the two-factor (M2s) and the three-factor (M3s) models, for both groups, the RMSEA indicated good fit, whereas the CFI and TFI for both groups indicated close to acceptable or acceptable fit. The exception was that the CFI value for the three-factor model in the adolescent group indicated good fit. The WRMR values for the one-, two-, and three-factor model for both groups were noticeably above 0.90. The fit values for both the bi-factor models (i.e., with two [M4s] and three [M5s] group factors) showed good fit in terms of the RMSEA, CFI, and TLI values. For both groups, the WRMR values for the bi-factor model with two group factors were lower than that for the bi-factor model with three group factors. Indeed, these values for the bi-factor model with two group factors were either just below or close to 0.90. Overall, there was reasonable support for both the bi-factor models. As the two bi-factor models were not nested, it was not possible to compare these models using the χ^2 (or WLSMVχ^2) difference test or the approximate fit indices based on χ^2 for the models. However, given that the WRMR values in the bi-factor model with two group factors were close to 0.90, and also because the bi-factor model with two group factors is more parsimonious than the bi-factor model with three group factors, the bi-factor model with two group factors was interpreted as the optimum model for both the groups in the present study and was used in subsequent reliability and validity analyses.

Factor Loadings for the Factors in the Optimum Bi-factor Model

Table 4 presents the standardized factor loadings of the 13 disorders on their respective latent factors in the optimum bi-factor model. For the adolescent P-factor, all disorders except GAD, ADHD, CD, and ODD had salient loadings, based on Thurstone's (1947) classical criterion for "salience" as standardized loading ≥ 0.3. For the internalizing group factor for adolescents, all disorders except OCD, PTSD, and DYTH had salient loadings, and for the externalizing group factor, all three externalizing disorders (ADHD, CD and ODD) had salient loadings on it. The loadings for SPE, GAD, ADHD, CD, and ODD were much higher (relatively) on their group factors (0.40, 0.50, 0.55, 0.88, and 0.76, respectively) than the P-factor (0.34, 0.29, 17, 20 and 0.24, respectively). Thus, much of the variance in the P-factor can be attributed to the internalizing disorders, with the externalizing disorders contributing negligible to low amounts of variances. Indeed, at the disorder level, with the exception of SPE, GAD, ECV values (which indicate the amount of common variance contributed by a disorder to the P-factor), the other internalizing disorders were high, ranging from 0.53 to 0.99. These values were very low for the externalizing disorders (ADHD = 0.09, ODD = 0.05, and CD = 0.09). Taken together, these findings can be interpreted as indicating questionable support for a P-factor since it was saturated with mostly variances from the internalizing disorders and negligible to low variances from the externalizing disorders. Additionally, because there was low amount of variance left in the group factor for internalizing disorders, the internalizing factors may be substantively of less use. In contrast, because there was high amount of variance left in the group factor for externalizing, it can be interpreted that there is support for the externalizing group factor, even after removing the variances allocated to the general factor.

For the child P-factor, all disorders except SPE, AGRO, GAD, ADHD, CD, and ODD had salient loadings. For the internalizing group factor in this group, all disorders except DYTH

Table 3 Fit of the childhood disorders models for psychopathology tested in the study

Models (M)	Model fit (WLSMV)					
	χ^2	*df*	RMSEA (90% CI)	CFI	TLI	WRMR
Adolescents ($N = 866$)						
M1. 1 primary factor (global)	647.126***	65	0.102 (0.095–0.109)	0.680	0.616	2.540
M2. 2 primary factors (I and E)	180.928***	64	0.046 (0.038–0.054)	0.936	0.922	1.315
M3. 3 primary factors (D, F, E)	141.078***	62	0.038 (0.030–0.047)	0.957	0.945	1.137
M4. Bi-factor (with P, and I & E specific factors)	89.416**	52	0.029 (0.018–0.039)	0.979	0.969	0.853
M5. Bi-factor (with P, and D, F, & E specific factors)	96.958***	52	0.032 (0.022–0.041)	0.975	0.963	0.923
Children ($N = 1233$)						
M1. 1 primary factor (global)	801.746***	65	0.096 (0.090–0.102)	0.723	0.668	2.809
M2. 2 primary factors (I & E)	264.474***	64	0.050 (0.044–0.057)	0.925	0.908	1.606
M3. 3 primary factors (D, F, E)	234.150***	62	0.047 (0.041–0.054)	0.935	0.919	1.493
M4. Bi-factor (with P, and I & E specific factors)	105.665***	52	0.029 (0.021–0.037)	0.980	0.970	0.940
M5. Bi-factor (with P, and D, F, & E specific factors)	144.411***	52	0.038 (0.031–0.045)	0.965	0.948	1.137

CFI comparative fit index, *CI* confidence interval, *RMSEA* root mean square error of approximation, *TLI* Tucker–Lewis Index, *WRMR* weighted root mean square residual, *I* internalizing disorders, *E* externalizing disorders, *D* distress disorders, *F* fear disorders, *P* P-factor

$^{**}p < .01$; $^{***}p < .001$

and MDD had salient loadings. For the externalizing group factor, all three externalizing disorders (ADHD, CD, and ODD) had salient loadings. The loadings for ten disorders (seven internalizing disorders [SAD, SOP, SPE, PD, AROG, GAD, and OCD], and three

Table 4 Factor loadings and reliabilities indices for the factors in the two-group (internalizing and externalizing) bi-factor model

Disorders	Adolescents				Children			
	P	I	E	ECV	P	I	E	ECV
Separation anxiety disorder	0.48	0.45		0.53	0.41	0.60		0.32
Social phobia	0.57	0.39		0.68	0.46	0.49		0.47
Specific phobia	0.34	0.40		0.42	0.18	0.51		0.11
Panic disorder	0.53	0.34		0.71	0.41	0.60		0.32
Agoraphobia	0.59	0.48		0.60	0.28	0.62		0.17
Generalized anxiety disorder	0.29	0.50		0.25	0.21	0.44		0.19
Obsessive–compulsive disorder	0.63	0.26		0.85	0.50	0.57		0.44
Posttraumatic stress disorder	0.52	0.04		0.99	0.34	0.32		0.53
Dysthymia	0.71	−0.12		0.97	0.87	−0.01		0.99
Major depressive disorder	0.76	−0.34		0.83	0.82	0.01		0.99
Attention deficit hyperactivity disorder	0.17		0.55	0.09	0.02		0.53	0.00
Conduct disorder	0.20		0.88	0.05	0.25		0.74	0.10
Oppositional defiant disorder	0.24		0.76	0.09	0.30		0.83	0.12
Explained common variance (ECV)	0.52	0.21	0.27		0.42	0.35	0.24	
Omega hierarchical (ω_h)	0.65				0.48			
Omega hierarchical subscale (ω_t)		0.14	0.75			0.41	0.73	
% of uncontaminated correlations(PUC)	0.53				0.39			
Index of construct reliability (*H*)	0.85	0.62	0.84		0.87	0.77	0.79	

P P-factor, *I* internalizing group factor, *E* externalizing group factor

externalizing disorders [ADHD, CD, and ODD]) were relatively higher on their group factors than the P-factor. Thus, relatively more of the variance in the P-factor can be attributed to the internalizing disorders, with the externalizing disorders contributing negligible to low amounts of variances. At the disorder level, with the exception of PTSD, DYTH, and MDD, ECV values of the other internalizing disorders were low, ranging from 0.00 to 0.47. These values were especially low for the externalizing disorders (ADHD = 0.00, ODD = 0.10, and CD = 0.12). Taken together, these moderate amounts of variances for the internalizing disorders and negligible to low variances from the externalizing disorders on the P-factor can be interpreted as indicating support for a weak P-factor. Additionally, as there was a moderate amount of variance left in the group factor for internalizing, the internalizing factors may be substantively meaningful. In contrast, because there was high amount of variance left in the group factor for externalizing, it can be interpreted that there is support for the externalizing group factor, even after taking out the variances allocated to the general factor.

Reliability for the Factors in the Optimum Bi-factor Model

Table 4 also presents the model-based reliability indices for all the factors in the optimum bi-factor models for children and adolescents separately. As shown, for both groups, the ECV values of the P-factor were much higher than the internalizing and externalizing group factors, with the values for the P-factor being higher for adolescents than children (for adolescents, P-factor = 0.52, internalizing group factor = 0.21, externalizing group factor = 0.27; for children, P-factor = 0.42, internalizing group factor = 0.35, externalizing group factor = 0.24). For adolescents and children, the ω_h values for the P-factor were 0.65 and 0.48, respectively. The ω_t values for the internalizing and externalizing group factors were 0.14 and 0.75, respectively, for adolescents, and 0.41 and 0.73, respectively, for children. Thus, although much of the reliable variance were attributed to the P-factor, there were still high levels of variances for the externalizing group factors left in both groups and moderate levels of variances for the internalizing group factor in children. The PUC values of the P-factor in adolescents and children were 0.53 and 0.39 respectively. As noted earlier, Reise et al. (2013a, b) have proposed that if PUC > 0.80, or if the PUC < 0.80, ECV > 0.60, and $\omega_h > 0.70$, then the bi-factor model can be interpreted as primarily unidimensional. As the findings failed to meet either of these criteria, it is not appropriate to interpret the findings of the present study as supporting a unidimensional model for the childhood disorders in the optimum bi-factor model in both age groups.

The *H* values for the adolescent P-factor and internalizing and externalizing group factors were 0.85, 0.62, and 0.84, respectively. They were 0.87, 0.77, and 0.79, respectively, for children. According to Rodriguez et al. (2016), *H* values > 0.80 are indicative of a stable well-defined latent variable, and values < 0.70 are indicative of a factor that is not worth specifying. Based on this guideline, the P-factor and the externalizing group factor for adolescents can be considered well-defined stable factors for this group and the P-factor and the externalizing and, to a lesser degree, the internalizing group factors for children can be considered well-defined stable factors accordingly.

Convergent and Divergent Validities of the Factors in the Two-Group (Internalizing and Externalizing) Bi-factor Model

Table 5 shows the findings of the predictions of broad CBCL externalizing and internalizing scores by the factors in the two group (internalizing and externalizing) bi-factor model for

children and adolescents. As shown, for children, the P-factor predicted both the CBCL internalizing and externalizing scores positively. The internalizing group factor predicted CBCL internalizing positively, and the externalizing group factor predicted CBCL externalizing positively. For adolescents, the P-factor also predicted both the CBCL internalizing and externalizing scores positively. The externalizing group factor predicted CBCL externalizing positively, and the internalizing group factor did not predict CBCL either externalizing or internalizing. Taken together, these findings can be interpreted as support for the convergent and divergent validities of the factors in the two-group (internalizing and externalizing) bi-factor model for both children and adolescents.

Discussion

Based on interviews of parents of clinic-referred (ACPU) children and adolescents, the first aim of the present study was to simultaneously examine the structure of the major DSM-IV/DSM-IV TR childhood internalizing disorders (SAD, SOP, SPP, PD, AG, GAD, OCD, PTSD, DYTH, and MDD) and externalizing disorders (ADHD, CD, and ODD). Five models were compared: (i) one-factor, (ii) two-factor oblique with primary factors for internalizing and externalizing disorders, (iii) three-factor oblique model with primary factors for distress, fear and externalizing disorders, (iv) bi-factor model with orthogonal P-factor and internalizing and externalizing group factors, and (v) bi-factor model with orthogonal P-factor and group factors for distress, fear, and externalizing. For both adolescents and children, the two- and three-factor models (but not the one-factor model) showed adequate fit. Also, for both groups, both bi-factor models showed good fit and were supported. Between these models, the bi-factor model with two group factors is more parsimonious and showed slightly better fit in terms of their WRMR values. Thus, the bi-factor model with the internalizing and externalizing group factors was interpreted as the optimum model. For this model, there was support for the convergent and divergent validities of the factors. For children, the P-factor predicted both the broad CBCL internalizing and externalizing scores positively. The internalizing group factor predicted CBCL internalizing positively, and the externalizing group factor predicted CBCL externalizing positively. For adolescents, the P-factor predicted both the CBCL internalizing and externalizing scores positively. The externalizing group factor predicted CBCL externalizing positively, and the internalizing group factor did not predict either CBCL externalizing or internalizing.

The relatively good and better support for the bi-factor models over the first-order oblique models found in the present study is consistent with past studies involving children (Lahey et al. 2015; Martel et al. 2017; Tackett et al. 2013) and adolescents (Carragher et al. 2016;

Table 5 Standardized path coefficients for the regression of the CBCL internalizing and externalizing scores on the factors in the two-group (internalizing and externalizing) bi-factor model

	Child			Adolescent		
	P	I	E	P	I	E
Internalizing total score	0.53***	0.39***	0.04	0.72***	0.16	− 0.05
Externalizing total score	0.21***	− 0.02	0.89***	0.20***	− 0.06	0.78***

P P-factor, *I* internalizing group factor, *E* externalizing group factor

*$p < .05$; **$p < .01$; ***$p < .001$

Castellanos-Ryan et al. 2016; Laceulle et al. 2015; Noordhof et al. 2015; Patalay et al. 2015; Tackett et al. 2013), as well as adults (Caspi et al. 2014; Lahey et al. 2012; Stochl et al. 2015). Despite this, the findings also extend existing data in this area. The present study used clinic-referred children provided with specific clinical diagnoses. All past studies in this area involving children and adolescents have utilized community samples and have used dimensional scores (derived from questionnaires and rating scales) of children's and adolescents' problems, to model their broad dimensions, such as internalizing and externalizing.

Although it is worth noting that the support for the bi-factor model in this study corresponds to past findings, the P-factor and group factors in the present study differed in important ways from those of past studies. Most past studies have generally found that both the internalizing and externalizing disorders contributed high variances to the P-factor. In contrast, in the present study, for both age groups, there were relatively low variances for the externalizing disorders on the P-factors, with the P-factor being saturated with mostly variances from the internalizing disorders (similar to studies by Laceulle et al. (2015)). Consequently, in both age groups, and unlike previous studies, there was high amount of variance in the externalizing group factors. It is possible that the differences found across the present study compared to those of past studies may be explained by the fact that unlike past studies that used community samples and measured psychopathology dimensionally using rating scales, the present study used a clinic-referred sample and measured psychopathology in terms of categorical clinical diagnosis.

The support for a strong P-factor and externalizing group factor for adolescents were also reinforced in terms of the reliability of these factors. The ECV value of the P-factor was much higher than the internalizing and externalizing group factors, and the ω_h value for the P-factor was 0.65, while the ω_t values for the internalizing and externalizing group factors were 0.14 and 0.75, respectively. Similarly, a moderately strong P-factor, a moderately strong internalizing group factor, and a strong externalizing group factor for children were supported in terms of the reliability of these factors. The ECV value of the P-factor was much higher than the internalizing and externalizing group factors for this group, and the ω_h value for the P-factor was only 0.48, and the ω_t values for the internalizing and externalizing group factors were 0.41 and 0.73, respectively. Additionally, the findings here indicated that the PUC values of the P-factor in adolescents and children were 0.53 and 0.39 respectively. Reise et al. (2013a, b) have proposed that if PUC > 0.80, or if the PUC < 0.80, ECV > 0.60, and $\omega_h > 0.70$, then the bi-factor model can be interpreted as primarily unidimensional. As the findings of the present study failed to meet either of these criteria, it can be argued that they are not supportive of unidimensional P-factors for the childhood internalizing and externalizing disorders in both children and adolescents. Also, the *H* values for the adolescent P-factor and internalizing and externalizing group factors were 0.85, 0.62, and 0.84, respectively. The same factors were 0.87, 0.77, and 0.79, respectively, for children. According to Rodriguez et al. (2016), *H* values > 0.80 are indicative of a stable well-defined latent variable, and values < 0.70 are indicative of a factor that is not worth specifying. Based on this guideline, and consistent with the argument presented, the P-factor and the externalizing group factor for adolescents can be considered well-defined stable factors for this group. Furthermore, the P-factor and the externalizing (and to a lesser degree, the internalizing group) factors for children can be considered well-defined stable factors in this group. Based on the combined findings for children and adolescents, the present authors' speculate that the P and all group factors in children and adolescent are meaningful, and they need to be considered when examining substantive issues, such as the validity and external correlates of the factors. In support of this, as shown in the present study,

for children, the internalizing group factor still predicted CBCL internalizing positively, and the externalizing group factor still predicted CBCL externalizing positively, even after removing the variances for the P-factor. For adolescents, the externalizing group factor still predicted CBCL externalizing positively, even after removing the variance for the P-factor. It may be worth noting that as there has been limited evaluation of the reliability of the factors in the bi-factor model (Martel et al. 2017; Murray et al. 2016), the present reliability findings extend existing data. Furthermore, ω_h and ω_t values and ECV, PUC, and *H* reliability indices (and not just ω_h and ω_t values, as in past studies) were combined and applied here, providing a more sophisticated and accurate interpretation of the dimensionality of the P-factor in the bi-factor model compared to that of previous studies.

As already illustrated, the findings in the present study also showed differences in the bi-factor model across children and adolescents. While there were negligible to low variances for the internalizing group factor in adolescents, there was moderate amount of variance in this factor in children. Thus, for adolescents, there was support for the P-factor and the externalizing group factor. For children, there was support for the P-factor and both the externalizing and internalizing group factors. Also, because the reliability was relatively higher for the P-factor in adolescents than children, it can be argued that the adolescent P-factor is relatively stronger than the child P-factor.

Given that the P-factor can be interpreted as strength of comorbidity of the internalizing and externalizing disorders (Murray et al. 2016), the present findings appear to propose that comorbidity (in particular considering the internalizing disorders, as the P-factor was saturated with variances from the internalizing disorders) may manifest differently at different developmental stages and that there is stronger comorbidity in adolescents than children. Indeed, consistent with this view, Castellanos-Ryan et al. (2016) have suggested that the factor loading on the P-factor could vary developmentally, with internalizing symptoms becoming stronger with increasing age. Caspi et al. (2014) has proposed a dynamic mutualism process hypothesis to account for this. According to this hypothesis, symptoms both across and within domains can reinforce one another through local (bi-directional associations between different symptom typologies) interactions such that, over time, these local interactions can lead to an increase in symptom inter-correlations. However, it should be noted that Murray et al. (2016) found no support for this hypothesis and therefore further research is recommended.

The findings of the present study have implications for (i) taxonomy in relation to children and adolescents, (ii) understanding the comorbidity of the internalizing disorders and externalizing disorders, (iii) trans-diagnostic assessment and diagnosis and treatment, and (iv) research on bi-factor models of psychopathology. In relation to taxonomy, support for the P-factor and broad internalizing and externalizing factors is inconsistent with the DSM approach that separates anxiety, depressive, and externalizing disorders into different diagnostic groups. It is also inconsistent with how the relevant disorders are organized in DSM-5, which suggests four different groups for the internalizing disorders. One group comprises SAD, SOP, SPP, PD, and AG, whereas GAD, OCD, PTSD, and MDD combined with DTYH are each in three different groups.

The support for the bi-factor model implies that current taxonomy that considers the different types of these disorders as discrete diagnostic categories may need reconsideration to recognize the high degree of comorbidity among them. Instead, the support for the bi-factor model in the present study suggests that for children and adolescents, all the common childhood psychopathologies could be grouped under an overall group called childhood psychopathology and separated into subgroups of internalizing disorder and externalizing

disorders. The list of symptoms in the internalizing disorder group could be the key non-overlapping symptoms for the different anxiety and depressive disorders, and the list of symptoms in the externalizing disorder group could be the key non-overlapping symptoms for ADHD, CD, and ODD. To recognize that the major specific symptoms present in an individual, appropriate descriptors could be added to the diagnosis of internalizing disorder or externalizing disorder. For example, when an individual has panic and social phobia, the diagnosis could be internalizing disorder with panic and social phobia. Despite this proposal, it needs to be stressed that further research and replication of the findings in the present study is needed before such changes could be adopted.

In relation to understanding the comorbidity of the internalizing and externalizing disorders, the support found in the present study for the P-factor, with salient loadings for virtually all internalizing disorders, can be inferred as indication of the strength of the associations of these disorders with the underlying latent factor and by extension the comorbidity of the disorders. Consistent with past studies (for a meta-analysis study, see Angold et al. 1999; Krueger and Markon 2006), these findings suggest high comorbidity among the internalizing disorders on one hand, and the externalizing disorders on the other, and between the internalizing and externalizing disorders.

In relation to assessment, diagnosis, and treatment, the close associations between the internalizing and externalizing disorders, via the P-factor, found in the present study highlight the need for a comprehensive evaluation of all the internalizing and externalizing disorders for a better understanding of a child's or an adolescent's psychopathology. The findings also imply that treatment may have to focus on general risk factors for psychopathology. In this respect, recently developed trans-diagnostic treatment approaches, such as for anxiety and depression disorders in children and adolescents (Ehrenreich-May and Bilek 2012), would be valuable. In brief, trans-diagnostic approaches focus on common factors that produce symptoms in related classes of disorders, thereby addressing multiple concerns or disorders within an individual (McEvoy et al. 2009).

In relation to research on the bi-factor model of psychopathology, the present findings suggest that researchers not only need to demonstrate good fit of such models (as has been the case with most of the studies in this area) but also need to examine the reliability of the P and group factors in the bi-factor model, as carried out in the present study. This enabled the present work to demonstrate that although the bi-factor model with two group factors was supported, all group factors in children and adolescents were also meaningful, and they need to be considered in structural models of psychopathology across these groups.

There are several strengths to the present study. First, it involved a large clinical sample, with clinical diagnoses of the major internalizing and externalizing disorders, derived via structured clinical interviews. Therefore, these findings appear more useful from a clinical viewpoint. Second, to increase the creditability of our findings, structural models were conducted separately for children and adolescents. The close comparability of the findings across these groups attest to the robustness of the findings reported here. Third, recent methodological developments were applied to evaluate the reliability of the factors in a bi-factor model (e.g. ECV, PUC, and *H*).

Despite these significant strengths, there are limitations in this study that need to be considered when interpreting the findings. First, not all disorders relevant to children and adolescents, such as eating disorders, autism spectrum conditions, psychotic symptoms, and substance abuse disorders, were included in the analyzed sample. It cannot be ruled out that the inclusion of these disorders may have produced different results. Second, the present study

examined the factor structure of the common DSM-IV anxiety and depressive disorders at the diagnostic level. Thus, the findings may not reflect the factor structure of anxiety and depressive disorders at the level of symptoms. As noted by Seeley et al. (2011), when analyses are focused at the disorder level, the underlying dimensionality associated with diagnostic criteria for specific disorders is ignored, and consequently, associations among symptoms might show different patterns of associations than those obtained on the basis of diagnostic categories/classifications. Third, approximately 75.3% and 66.8% of the participants had ADHD and ODD/CD respectively. Thus, it is not known if this exerted any influence on the findings. Fourth, as this study examined clinic-referred children, the findings here may not be applicable to internalizing and externalizing disorders in children and adolescents from the general community. Indeed, there is evidence that in clinical samples, comorbidity rates are usually higher than among general population samples (Angold et al. 1999) due to a methodological problem called Berkson's bias (i.e., people with multiple disorders are more likely to be referred to clinics than are people with single disorders). Fifth, all the participants in this study were from the same clinic. It is possible that this may constitute an additional bias. Sixth, the present study used a predominantly male sample, and this may have added some gender-related bias to the findings. Seventh, it is important to keep in mind that although children were the target of analysis, the information about these disorders was derived from interviews of parents and not children themselves. It is therefore possible that this may have also influenced parameter estimates. Finally, the present data (in line with previous P-factor literature) refer to a Westernized population of developing individuals. Therefore, generalization to different cultural groups should be treated with caution (i.e., potential cultural bias of the findings needs to be considered).

In conclusion, although, the present study found that the bi-factor model with the internalizing and externalizing disorders showed good fit and was the optimum model. However, the fact that there was (i) support for a strong externalizing group factor for adolescents, (ii) a moderately strong internalizing group factor, and (iii) a strong externalizing group factor for children weakens the support for a dominant P-factor in both these groups. Thus, the conclusion of the present study differs from the prevailing general consensus for a robust dominant P-factor for psychopathology (Castellanos-Ryan et al. 2016; Laceulle et al. 2015) because it calls for the inclusion of the group factors in structural models of psychopathology. Nevertheless, this view is not without support from previous studies (Laceulle et al. 2015). Like the findings in the present study, these studies also showed that there was substantial variance in the group factors even after removing the variances for the general factor. Given the discrepancy of the findings in the present study and that of past studies, more studies are needed that take into consideration the aforementioned limitations. Such studies will be valuable as they could have implications for the understanding of the development and the course of childhood psychopathology across different age and cultural groups (prioritizing those referring to relatively under-researched populations), which in turn can have implications for more developmentally and culturally responsive diagnosis and treatment of childhood disorders.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

Ethical Approval All procedures performed in this study involving human participants were in accordance with the ethical standards of University's Research Ethics Board and with the 1975 Helsinki Declaration.

Informed Consent Informed consent was obtained from all participants.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

Achenbach, T. M. (1966). The classification of children's psychiatric symptoms: a factor-analytic study. *Psychological Monographs: General and Applied, 80*(7), 1–37. https://doi.org/10.1037/h0093906.

Achenbach, T. M., & Edelbrock, C. S. (1978). The classification of child psychopathology: a review and analysis of empirical efforts. *Psychological Bulletin, 85*(6), 1275–1301. https://doi.org/10.1037/0033-2909.85.6.1275.

Achenbach, T. M., & Rescorla, L. (2001). *ASEBA school-age forms & profiles*. Burlington: University of Vermont, Research Center for Children, Youth, & Families.

Angold, A., & Costello, E. J. (2009). Nosology and measurement in child and adolescent psychiatry. *Journal of Child Psychology and Psychiatry, 50*(1–2), 9–15. https://doi.org/10.1111/j.1469-7610.2008.01981.x.

Angold, A., Costello, E. J., & Erkanli, A. (1999). Comorbidity. *Journal of Child Psychology and Psychiatry, 40*(1), 57–87. https://doi.org/10.1111/1469-7610.00424.

Beauducel, A., & Herzberg, P. Y. (2006). On the performance of maximum likelihood versus means and variance adjusted weighted least squares estimation in CFA. *Structural Equation Modeling, 13*(2), 186–203. https://doi.org/10.1207/s15328007sem1302_2.

Bonifay, W. E., Reise, S. P., Scheines, R., & Meijer, R. R. (2015). When are multidimensional data unidimensional enough for structural equation modeling? An evaluation of the DETECT multidimensionality index. *Structural Equation Modeling, 22*(4), 504–516. https://doi.org/10.1080/10705511.2014.938596.

Brunner, M., Nagy, G., & Wilhelm, O. (2012). A tutorial on hierarchically structured constructs. *Journal of Personality, 80*(4), 796–846. https://doi.org/10.1111/j.1467-6494.2011.00749.x.

Carragher, N., Krueger, R. F., EatonN. R., & Slade, T. (2015). Disorders without borders: Current and future direction in the meta-structure of mental disorders. *Social Psychiatry and Psychiatric Epidemiology, 50*(3), 339–350

Carragher, N., Sunderland, M., Batterham, P. J., Calear, A. L., Elhai, J. D., Chapman, C., & Mills, K. (2016). Discriminant validity and gender differences in DSM-5 posttraumatic stress disorder symptoms. *Journal of Affective Disorders, 190*, 56–67. https://doi.org/10.1016/j.jad.2015.09.071.

Caspi, A., Houts, R. M., Belsky, D. W., Goldman-Mellor, S. J., Harrington, H., Israel, S., Meier, M. H., Ramrakha, S., Shalev, I., Poulton, R., & Moffitt, T. E. (2014). The p factor: one general psychopathology factor in the structure of psychiatric disorders? *Clinical Psychological Science, 2*(2), 119–137. https://doi.org/10.1177/2167702613497473.

Castellanos-Ryan, N., Brière, F. N., O'Leary-Barrett, M., Banaschewski, T., Bokde, A., Bromberg, U., et al. (2016). The structure of psychopathology in adolescence and its common personality and cognitive correlates. *Journal of Abnormal Psychology, 125*(8), 1039–1052. https://doi.org/10.1037/abn0000193.

DSM-IV-TR, A. P. A., (2000). *Diagnostic and Statistical Manual of Mental Disorders, text revision*. Washington, DC: American Psychiatric Association.

Dueber, D. M. (2016). Bifactor indices calculator: a Microsoft Excel-based tool to calculate various indices relevant to bifactor CFA models. Retrieved August 7, 2018, from: http://sites.education.uky.edu/apslab/resources

Ehrenreich-May, J., & Bilek, E. L. (2012). The development of a transdiagnostic, cognitive behavioral group intervention for childhood anxiety disorders and co-occurring depression symptoms. *Cognitive and Behavioral Practice, 19*(1), 41–55. https://doi.org/10.1016/j.cbpra.2011.02.003.

Grills, A. E., & Ollendick, T. H. (2003). Multiple informant agreement and the anxiety disorders interview schedule for parents and children. *Journal of the American Academy of Child & Adolescent Psychiatry, 42*(1), 30–40. https://doi.org/10.1097/00004583-200301000-00008.

Hancock, G. R., & Mueller, R. O. (2001). Rethinking construct reliability within latent variable systems. In R. Cudeck, S. Du Toit, & D. Sörbom (Eds.), *Structural equation modeling: present and future. A festschrift in honor of Karl Jöreskog* (pp. 195–216). Scientific Software International: Lincolnwood.

Hu, L. T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: sensitivity to underparameterized model misspecification. *Psychological Methods, 3*(4), 424–453. https://doi.org/10.1037/1082-989X.3.4.424.

Jensen, P. S., Rubio-Stipec, M., Canino, G., Bird, H. R., Dulcan, M. K., Schwab-Stone, M. E., & Lahey, B. B. (1999). Parent and child contributions to diagnosis of mental disorder: are both informants always necessary? *Journal of the American Academy of Child & Adolescent Psychiatry, 38*(12), 1569–1579. https://doi.org/10.1097/00004583-199912000-00019.

Krueger, R. F. (1999). The structure of common mental disorders. *Archives of General Psychiatry, 56*(10), 921–926. https://doi.org/10.1001/archpsyc.56.10.921.

Krueger, R. F., & Finger, M. S. (2001). Using item response theory to understand comorbidity among anxiety and unipolar mood disorders. *Psychological Assessment, 13*(1), 140–151. https://doi.org/10.1037/1040-3590.13.1.140.

Krueger, R. F., & Markon, K. E. (2006). Reinterpreting comorbidity: a model-based approach to understanding and classifying psychopathology. *Annual Review of Clinical Psychology, 2*, 111–133. https://doi.org/10.1146/annurev.clinpsy.2.022305.095213.

Laceulle, O. M., Vollebergh, W. A., & Ormel, J. (2015). The structure of psychopathology in adolescence replication of a general psychopathology factor in the TRAILS Study. *Clinical Psychological Science, 3*(6), 850–860. https://doi.org/10.1037/0021-843X.113.3.358.

Lahey, B. B., Applegate, B., Waldman, I. D., Loft, J. D., Hankin, B. L., & Rick, J. (2004). The structure of child and adolescent psychopathology: generating new hypotheses. *Journal of Abnormal Psychology, 113*(3), 358–385. https://doi.org/10.1037/0021-843X.113.3.358.

Lahey, B. B., Applegate, B., Hakes, J. K., Zald, D. H., Hariri, A. R., & Rathouz, P. J. (2012). Is there a general factor of prevalent psychopathology during adulthood? *Journal of Abnormal Psychology, 121*(4), 971–977. https://doi.org/10.1037/a0028355.

Lahey, B. B., Rathouz, P. J., Keenan, K., Stepp, S. D., Loeber, R., & Hipwell, A. E. (2015). Criterion validity of the general factor of psychopathology in a prospective study of girls. *Journal of Child Psychology and Psychiatry, 56*(4), 415–422. https://doi.org/10.1111/jcpp.12300.

Lubke, G. H., & Muthén, B. O. (2004). Applying multigroup confirmatory factor models for continuous outcomes to Likert scale data complicates meaningful group comparisons. *Structural Equation Modeling, 11*(4), 514–534. https://doi.org/10.1207/s15328007sem1104_2.

Martel, M. M., Pan, P. M., Hoffmann, M. S., Gadelha, A., do Rosário, M. C., Mari, J. J., et al. (2017). A general psychopathology factor (P factor) in children: structural model analysis and external validation through familial risk and child global executive function. *Journal of Abnormal Psychology, 126*(1), 137–148. https://doi.org/10.1037/abn0000205.

McEvoy, P. M., Nathan, P., & Norton, P. J. (2009). Efficacy of transdiagnostic treatments: a review of published outcome studies and future research directions. *Journal of Cognitive Psychotherapy, 23*(1), 20–33. https://doi.org/10.1891/0889-8391.23.1.20.

Millsap, R. E., & Yun-Tein, J. (2004). Assessing factorial invariance in ordered-categorical measures. *Multivariate Behavioral Research, 39*(3), 479–515. https://doi.org/10.1207/S15327906MBR3903_4.

Murray, A. L., Eisner, M., & Ribeaud, D. (2016). The development of the general factor of psychopathology 'p factor' through childhood and adolescence. *Journal of Abnormal Child Psychology, 44*(8), 1573–1586. https://doi.org/10.1007/s10802-016-0132-1.

Muthén, L. K., & Muthén, B. O. (2013). *Mplus 7.11*. Los Angeles: Muthén & Muthén.

Noordhof, A., Krueger, R. F., Ormel, J., Oldehinkel, A. J., & Hartman, C. A. (2015). Integrating autism-related symptoms into the dimensional internalizing and externalizing model of psychopathology. The TRAILS Study. *Journal of Abnormal Child Psychology, 43*(3), 577–587. https://doi.org/10.1007/s10802-014-9923-4.

Nye, C. D., & Drasgow, F. (2011). Assessing goodness of fit: simple rules of thumb simply do not work. *Organizational Research Methods, 14*(3), 548–570. https://doi.org/10.1177/1094428110368562.

Patalay, P., Fonagy, P., Deighton, J., Belsky, J., Vostanis, P., & Wolpert, M. (2015). A general psychopathology factor in early adolescence. *British Journal of Psychiatry, 207*(1), 15–22. https://doi.org/10.1192/bjp.bp.114.149591.

Reise, S. P. (2012). The rediscovery of bi-factor measurement models. *Multivariate Behavioral Research, 47*(5), 667–696. https://doi.org/10.1080/00273171.2012.715555.

Reise, S. P., & Waller, N. G. (2009). Item response theory and clinical measurement. *Annual Review of Clinical Psychology, 5*, 27–48.

Reise, S. P., Bonifay, W. E., & Haviland, M. G. (2013a). Scoring and modeling psychological measures in the presence of multidimensionality. *Journal of Personality Assessment, 95*(2), 129–140. https://doi.org/10.1080/00223891.2012.725437.

Reise, S. P., Scheines, R., Widaman, K. F., & Haviland, M. G. (2013b). Multidimensionality and structural coefficient bias in structural equation modeling: a bi-factor perspective. *Educational and Psychological Measurement, 73*(1), 5–26. https://doi.org/10.1177/0013164412449831.

Rhemtulla, M., Brosseau-Liard, P. E, & Savalei, V. (2012). When can categorical variables be terated as continuous? A comparison of robust continuous and categorical SEM estimation methods under suboptimal conditions. *Psychological Methods, 17*(3), 354.

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Evaluating bifactor models: Callculating and interpreting statistical indices. *Psychological Methods, 21*(2), 137.

Seeley, J. R., Kostly, D. B., Farmer, R. F., & Lewinsohn, P. M. (2011). The modelling of internalizing disorders on the basis of patterns of lifetime comorbidity: associations with psychosocial functioning and psychiatric disorders among first-degree relatives. *Journal of Abnormal Psychology, 120*(2), 308.

Silverman, W. K., & Albano, A. M. (1996). *Anxiety Disorders Interview Schedule for DSM-IV: child interview schedule* (Vol. 2). Boulder: Graywind Publications.

Silverman, W. K., Saavedra, L. M., & Pina, A. A. (2001). Test-retest reliability of anxiety symptoms and diagnoses with the Anxiety Disorders Interview Schedule for DSM-IV: child and parent versions. *Journal of the American Academy of Child & Adolescent Psychiatry, 40*(8), 937–944. https://doi.org/10.1097/00004583-200108000-00016.

Slade, T. I. M., & Watson, D. (2006). The structure of common DSM-IV and ICD-10 mental disorders in the Australian general population. *Psychological Medicine, 36*(11), 1593–1600. https://doi.org/10.1017/S0033291706008452.

Stochl, J., Khandaker, G. M., Lewis, G., Perez, J., Goodyer, I. M., Zammit, S., Sullivan, S., Croudace, T. J., & Jones, P. B. (2015). Mood, anxiety and psychotic phenomena measure a common psychopathological factor. *Psychological Medicine, 45*(07), 1483–1493. https://doi.org/10.1017/S003329171400261X.

Tackett, J. L., Lahey, B. B., Van Hulle, C., Waldman, I., Krueger, R. F., & Rathouz, P. J. (2013). Common genetic influences on negative emotionality and a general psychopathology factor in childhood and adolescence. *Journal of Abnormal Psychology, 122*(4), 1142–1153. https://doi.org/10.1037/a0034151.

Thurstone, L. L. (1947). *Multiple factor analysis*. Chicago: University of Chicago.

Vollebergh, W. A., Iedema, J., Bijl, R. V., de Graaf, R., Smit, F., & Ormel, J. (2001). The structure and stability of common mental disorders: the NEMESIS study. *Archives of General Psychiatry, 58*(6), 597–603. https://doi.org/10.1001/archpsyc.58.6.597.

Watkins, M. W. (2013). *Omega [Computer software]*. Phoenix: Ed & Psych Associates.

Wright, A. G., Krueger, R. F., Hobbs, M. J., Markon, K. E., Eaton, N. R., & Slade, T. (2013). The structure of psychopathology: toward an expanded quantitative empirical model. *Journal of Abnormal Psychology, 122*(1), 281–294. https://doi.org/10.1037/a0030133.

Yu, C. Y., & Muthen, B. (2002). *Evaluation of model fit indices for latent variable models with categorical and continuous outcomes*. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans.

Zinbarg, R. E., Revelle, W., Yovel, I., & Li, W. (2005). Cronbach's α, Revelle's β, and McDonald's ωH: their relations with each other and two alternative conceptualizations of reliability. *Psychometrika, 70*(1), 123–133. https://doi.org/10.1007/s11336-003-0974-7.

Zinbarg, R. E., Yovel, I., Revelle, W., & McDonald, R. P. (2006). Estimating generalizability to a latent variable common to all of a scale's indicators: a comparison of estimators for ωh. *Applied Psychological Measurement, 30*(2), 121–144. https://doi.org/10.1177/0146621605278814.

Underlying research materials related to the current work (for example data, samples or models) can be accessed through contacting Professor Rapson Gomez.