Psychometric properties of the Strength and Difficulties Questionnaire from five European countries

CECILIA A. ESSAU,[1] BEATRIZ OLAYA,[2,3] XENIA ANASTASSIOU-HADJICHARALAMBOUS,[4] GINA PAULI,[1] CATHERINE GILVARRY,[1] DIANE BRAY,[1] JEAN O'CALLAGHAN[1] & THOMAS H. OLLENDICK[5]

1 Department of Psychology, Roehampton University, London, UK
2 CIBER en Salud Mental (CIBERSAM), Madrid, Spain
3 Parc Sanitari Sant Joan de Déu, Sant Boi de Llobregat, Barcelona, Spain
4 Department of Psychology, University of Nicosia, Nicosia, Cyprus
5 Child Study Centre, Department of Psychology, Virginia Tech, Blacksburg, VA, USA

Correspondence
Cecilia A. Essau, Professor of Developmental Psychopathology & Director, Centre for Applied Research and Assessment in Child and Adolescent Wellbeing (CARACAW), Department of Psychology, Roehampton University, Whitelands College, Holybourne Avenue, London SW15 4JD, UK.
Telephone (+44) (0) 20 8392 3647
Fax (+44) (0) 20 8392 3527
Email: c.essau@roehampton.ac.uk

Abstract

The Strength and Difficulties Questionnaire (SDQ) is a brief screening instrument for assessing emotional and behavioural problems in children and adolescents. This study examined the factor structure and validity of the self-report original English version of the SDQ and four of its many translations (German, Cypriot Greek, Swedish, and Italian). A total of 2418 adolescents from five European countries (Germany, Cyprus, England, Sweden, Italy), ages 12 to 17 years, participated. The sample was drawn from general (school) populations. In addition to the SDQ, all participants completed the Spence Children's Anxiety Scale (SCAS), a measure of anxiety symptoms. The internal consistency and validity of the SDQ total difficulties were good for most countries. Confirmatory factor analysis showed that both five- and three-factor models provided good fit for the whole sample; however, the three-factor model fit somewhat better than the five-factor model. The factor structure differed across countries, with the three-factor model showing better fit indices in Cyprus, whereas the five-factor model fitted better in Germany. Fit indices for the UK, Sweden, and Italy were poor for both models. When the reversed items were removed, the goodness-of-fit improved significantly for the total sample and in each country. It is therefore recommended that the reversed items be removed or re-worded in future studies. *Copyright © 2012 John Wiley & Sons, Ltd.*

Introduction

The Strengths and Difficulties Questionnaire (SDQ: Goodman, 1997) is a brief screening instrument which has been designed to assess emotional and behavioural problems in children and adolescents. The SDQ has four subscales to measure major difficulties commonly experienced by children and adolescents (conduct problems, hyperactivity-inattention, emotional symptoms, peer problems) and one subscale to assess strengths (prosocial behaviour). There are two versions of the SDQ: (1) The informant-rated version of the SDQ can be completed by either the parents or teachers of children and adolescents aged 4 to 16 years; (2) a self-report version of the SDQ can be completed by young adolescents aged 11 to 16 years. The present study was based on the self-report version of the SDQ.

Since its publication in 1997, the SDQ has been used in both clinical and community settings throughout the world. It has been translated into more than 40 languages. The popularity of the SDQ has been attributed to its brevity, simplicity, and good reliability and validity. Furthermore, unlike many other self-report questionnaires such as the Youth Self-report (Achenbach, 1991), the SDQ is publicly available, can be downloaded from the internet (http://www.sdqinfo.org), and can be used without incurring any expense.

Three major types of studies have been conducted using the SDQ. The first series of studies evaluated the psychometric properties of the scale. In the original paper describing the development of the SDQ (Goodman, 2001), the internal consistency and the retest reliability were generally satisfactory with a mean Cronbach's α for the total difficulty score of 0.73 and the retest reliability (four to six months) of 0.62. Numerous other studies conducted in various countries have reported α coefficients for the SDQ to similarly range from 0.70 to 0.81 (Capron *et al.*, 2007; Hawes and Dadds, 2004; Muris *et al.*, 2003, 2004; Riso *et al.*, 2010; van Widenfelt *et al.*, 2003; Yao *et al.*, 2009). Within the SDQ subscales, acceptable α coefficients have also been reported for emotional symptoms (0.63–0.78), hyperactivity-inattention (0.66–0.73), and prosocial behaviour (0.59–0.87); however, relatively low α coefficients have been reported for conduct problems (0.41–0.67) and peer problems (0.27–0.52) (Hawes and Dadds, 2004; Muris *et al.*, 2003, 2004; van Widenfelt *et al.*, 2003; Yao *et al.*, 2009).The test–retest reliability coefficient (two months) of the Chinese SDQ was 0.71 (Yao *et al.*, 2009). Collectively, these findings have provided a firm basis for the reliability and internal consistency of the SDQ as a whole but have called into question some of the subscales.

The second type of studies has examined the validity of the SDQ. The specificity and sensitivity of the SDQ for identifying psychiatric diagnoses as identified by structured diagnostic interview was 94.6% and 63.3%, respectively (Goodman *et al.*, 2000). Furthermore, the SDQ successfully identified over 70% of children and adolescents with conduct, hyperactivity, depressive and most anxiety disorders; however, identification was poor for specific phobia, separation anxiety, eating disorders, and panic disorder/agoraphobia, with sensitivity ranging from 30 to 50%. The convergent validity of the SDQ has been tested by computing correlations between the SDQ and other measures that purport to assess psychopathological symptoms such as the Youth Self-report (Achenbach, 1991) and the Revised Children's Manifest Anxiety Scale (RCMAS: Reynolds and Richmond, 1978). For example, in a recent study by Yao *et al.* (2009), the SDQ was found to correlate significantly with the Youth Self-report. Within the SDQ subscales, the emotional symptoms correlated significantly with the following Youth Self-report subscales: the internalizing problems, anxiety/depressed, withdraw/depressed, and somatic complaints. The conduct problems correlated significantly with the Youth Self-report subscales of rule-breaking, aggressive behaviours, and externalizing problems. These results support the validity of the SDQ. The SDQ also correlated significantly with anxiety symptoms of the parent version of the RCMAS (Muris *et al.*, 2003), with the strongest correlation being found between RCMAS and the emotional symptom subscale.

The third type of studies has examined the factor structure of the SDQ. Compared to the two other types of studies, considerable controversy remains regarding the factor structure of this instrument. Support for the five-factor solution that corresponds with the hypothesized subscales of hyperactivity-inattention, emotional symptoms, peer problems, conduct problems, and prosocial behaviour have been reported in several factor analytic studies conducted in community samples in Australia (Hawes and Dadds, 2004), England (Goodman, 2001), Germany (Klasen *et al.*, 2003), the Netherlands (Muris *et al.*, 2003), and Sweden (Smedje *et al.*, 1999). Confirmatory factor analysis (CFA) of the Chinese translation of the SDQ (Yao *et al.*, 2009) similarly supported the five-factor structure. However, the five-factor structure has not always fit the data based on other translated versions of the SDQ. For example, an American study (Dickey and Blumberg, 2004) and a Finnish study (Koskelainen *et al.*, 2001) conducted in community samples found the best-fitting factor solution involved three dimensions (externalizing, internalizing, and a prosocial subscale). A recent Italian community study (Riso *et al.*, 2010) using CFA also supported a three factors comprised of externalizing, internalizing, and prosocial behaviour. In sum, it remains unsettled at this time whether the factor structure of the

SDQ as determined by Goodman (1997) can be applied to adolescents from other cultural backgrounds.

The main aim of the present study therefore was to explore the factor structure of the SDQ and its convergent validity when used among adolescents in five European countries. The more specific aims were: (1) to investigate the psychometric properties of the four translations (German, Swedish, Italian and Cypriot Greek) and the English version of the SDQ; (2) to examine whether the same factor structure fits across the five European countries; (3) to determine if the SDQ scores correlate in the expected direction with a measure of emotional symptoms; (4) to compare levels of emotional and behavioural problems across the five European countries.

Method

Sample

A total of 2418 adolescents who attended urban schools in Germany ($N = 495$), Cyprus ($N = 471$), England ($N = 469$), Sweden ($N = 484$), and in Italy ($N = 499$) participated in the study. Schools were randomly selected from a list of state schools that was provided by the local education authority. Participants' ages ranged between 12 and 17 years [mean = 14.58 years; standard deviation (SD) = 1.6]. There was no significant differences among the five samples in terms of age distribution [$F(4, 2417) = 2.22$, non-significant (n.s.)]. However, a gender difference was found, with significantly more girls than boys participating in the study [$\chi^2(1) = 33.04$, $p < 0.001$].

The British sample was recruited from schools in southwest London. The sample consisted of 218 boys (46.5%) and 251 girls (53.5%), with a mean age of 14.4 years (SD = 1.7). The German sample was recruited from schools in Northrhein Westfalia. The sample consisted of 216 boys (43.6%) and 279 girls (56.4%); the mean age was 14.6 years (SD = 1.6). The Cypriot sample was recruited from schools in the non-occupied territory of the island, and consisted of 174 boys (36.9%) and 297 girls (63.1%) and the mean age of the respondents was 14.6 years (SD = 1.7). The Swedish sample was recruited from schools in the southern part of Stockholm. Two hundred and thirty (47.5%) of the sample were boys, and 254 (52.5%) were girls; the mean age was 14.5 years (SD = 1.6). The Italian sample was recruited from schools in the northern part of Milan. More than half of the Italian sample were girls ($N = 336$; 67.3%); 163 (32.7%) were boys. Their mean age was 14.7 years (SD = 1.5).

Except for participants in the UK, almost all (ranging from 92 to 97%) the adolescents in the other four European countries were Caucasian and the ethnic composition of the sample reflected the composition of the ethnic distribution of each country. The ethnic composition found in the UK was somewhat more diverse but still largely Caucasian: 76% were White, 15% were Asian/Asian British, 6% were Black/Black British, 3% were Chinese or other ethnic groups. All of these adolescents were born and raised in the UK and were either second or third generation British citizens. This ethnic composition is representative for that in the UK. In all countries, the socio-economic status of participants varied greatly, with parents working as unskilled labourers as well as professionals and managers.

School approval and parental written informed consent were obtained before participation in the study. The response rate ranged from 78% in Germany to 94% in Sweden; common reason for non-participation was related to adolescents forgetting to get the consent form signed by their parent. Adolescent's participation was voluntary and no incentives were given for their participation. The adolescents completed questionnaires in their classroom and two research assistants were available to provide assistance if necessary and to ensure independent responding.

Instruments

The Strengths and Difficulties Questionnaire (SDQ: Goodman, 1997)

Each subscale of the SDQ (conduct problems, hyperactivity-inattention, emotional symptoms, peer problems) contains five items which are rated on a three-point Likert-type scale (0 = not true, 1 = somewhat true, or 2 = certainly true). Each subscale is calculated by adding scores on the relevant items (after reversing indicated items). A total difficulties score can be calculated by adding the scores of the four difficulties subscales, with higher scores reflecting greater difficulties. Higher scores on the prosocial behaviour subscale reflect more strength. The SDQ also contains an extended set of items measuring the impact of mental health problem (called "impact supplement") on everyday life.

The Spence Children's Anxiety Scale (SCAS: Spence, 1997)

The Spence Children's Anxiety Scale (SCAS) is a 38-item measure of anxiety symptoms in children and adolescents, which reflect symptoms of the main DSM-IV (Fourth Edition of the Diagnostic and Statistical Manual of Mental Disorders: American Psychiatric Association, 1994) anxiety disorders. The SCAS consists of six subscales: separation anxiety, social phobia, obsessive-compulsive disorder, panic/agoraphobia, physical injury fears, and generalized anxiety disorder. Each item is rated on a four-point scale in terms of its frequency from "never" (0) to "always" (3). Internal consistency and test–retest reliability of the SCAS have been reported as

satisfactory, with Cronbach's α generally well above 0.70 and a test–retest correlation coefficient of 0.60 (Spence, 1997). The scale has also been shown to possess acceptable convergent validity as demonstrated by a significant correlation with the RCMAS (Reynolds and Richmond, 1978). Furthermore, clinically anxious children have been shown to have significantly higher SCAS scores and marked particularly high scores on the subscales that reflect the anxiety disorders they suffered from. In the present study, the Cronbach's α for the total SCAS was 0.97 for the UK, 0.89 for Germany, 0.91 for Italy, 0.93 for Sweden, and 0.91 for Cyprus.

Results

Psychometric properties of SDQ in five European countries

Table 1 presents Cronbach's α values of the SDQ total difficulties, the five subscales reported by Goodman (2001), and the three factors proposed by other studies (Dickey and Blumberg, 2004; Koskelainen *et al.*, 2001; Riso *et al.*, 2010) in each country. The internal consistencies of the SDQ total difficulties ranged from satisfactory to good, with values ranging from 0.51 for Italy to 0.74 for Germany. Cronbach's α for the five SDQ subscales were much lower, indicating that some items do not fit with the rest of the items on that subscale. Of all the subscales, conduct problems had the lowest internal consistency coefficients (0.02–0.45), followed by peer problems (0.15–0.54). From

the three theoretical factors, prosocial behaviour had the lowest α value, ranging from 0.10 to 0.34. Because the reversed items have been reported to confound the SDQ factors (Palmieri and Smith, 2007; Riso *et al.*, 2010), the internal consistency of the SDQ total difficulties score, the five subscales and the three factors were analysed after the exclusion of the reversed items (i.e. items 7, 11, 14, 21, and 25) (Table 1). Internal consistency values of the total difficulties score, conduct problems, and hyperactivity-inattention subscales improved considerably after removing the reversed items in all countries whereas peer problems' values did not improve after items 11 and 14 were removed. Regarding the three-factor model, the Cronbach's α value of prosocial behaviour and externalizing problems subscale also improved considerably with the exclusion of reversed items.

The validity of the SDQ was examined by correlating it with the SCAS (Table 2). Divergence between externalizing (SDQ conduct problems and hyperactivity-inattention scales) and internalizing symptoms on the SCAS was expected. However, Table 2 shows that in some countries, both conduct problems and hyperactivity-inattention subscales on the SDQ correlated significantly with the anxiety scores as measured using the SCAS.

In order to determine whether correlations between SDQ emotional problems and SCAS anxiety scores were significantly stronger (i.e. convergent validity) than correlations between SDQ conduct and hyperactivity-inattention subscales (i.e. divergent validity), the Steiger's Z (Steiger, 1980)

Table 1 Cronbach's α coefficients for the five-factor and the three-factor model in the five countries with and without reversed items

	Germany	Cyprus	UK	Sweden	Italy
Five factors model					
Total difficulties	0.74	0.61	0.71	0.61	0.52
Without reversed items	0.73	0.84	0.79	0.67	0.60
Emotional symptoms	0.70	0.80	0.58	0.68	0.60
Conduct problems	0.45	0.12	0.26	0.02	0.22
Without item 7	0.49	0.60	0.43	0.17	0.33
Hyperactivity-inattention	0.62	0.07	0.49	0.39	0.27
Without items 21 and 25	0.62	0.64	0.62	0.39	0.44
Peer problems	0.54	0.15	0.34	0.29	0.50
Without items 11 and 14	0.35	0.29	0.33	0.14	0.21
Prosocial behaviour	0.64	0.83	0.60	0.71	0.68
Three factors model					
Prosocial problems	0.25	0.10	0.17	0.34	0.12
Without items 7, 11, and 14	0.64	0.83	0.60	0.71	0.68
Internalizing problems	0.66	0.70	0.62	0.68	0.54
Externalizing problems	0.65	0.44	0.58	0.42	0.33
Without items 21 and 25	0.65	0.78	0.68	0.42	0.48

Table 2 Correlations between SDQ and SCAS in the five countries

SDQ	Total score	Separation anxiety	Social phobia	SCAS Obsessive-compulsive disorder	Panic-agoraphobia	Fears of physical injury	Generalized anxiety
Total score							
Germany	0.53**	0.36**	0.46**	0.46**	0.36**	0.36**	0.34**
Cyprus	0.43**	0.35**	0.37**	0.27**	0.49**	0.31**	0.28**
UK	0.63**	0.62**	0.51**	0.57**	0.55**	0.51**	0.56**
Sweden	0.19**	−0.02	0.16**	0.24**	0.20**	0.13**	0.19**
Italy	0.26**	0.27**	0.17**	0.28**	0.22**	0.11*	0.16**
Emotional symptoms							
Germany	0.59**	0.40**	0.54**	0.46**	0.42**	0.32**	0.46**
Cyprus	0.47**	0.36**	0.35**	0.28**	0.43**	0.33**	0.48**
UK	0.57**	0.57**	0.47**	0.57**	0.67**	0.44**	0.51**
Sweden	0.34**	0.12**	0.29**	0.29**	0.31**	0.20**	0.42**
Italy	0.60**	0.45**	0.43**	0.51**	0.48**	0.37**	0.49**
Conduct problems							
Germany	0.22**	0.14**	0.11*	0.30**	0.12**	0.17**	0.15**
Cyprus	0.20**	0.17**	0.12**	0.12*	0.25**	0.18**	0.14**
UK	0.28**	0.28**	0.20**	0.26**	0.29**	0.21**	0.23**
Sweden	0.01	−0.03	0.003	−0.01	−0.04	0.10*	0.01
Italy	−0.004	0.05	−0.06	0.12**	0.01	−0.15**	0.02
Hyperactivity-inattention							
Germany	0.30**	0.19**	0.28**	0.24**	0.20**	0.24**	0.16**
Cyprus	0.19**	0.09	0.20**	0.13**	0.23**	0.13**	0.11*
UK	0.46**	0.46**	0.35**	0.44**	0.37**	0.39**	0.44**
Sweden	−0.05	−0.21**	−0.05	0.07	0.08	−0.07	−0.05
Italy	−0.08	−0.05	−0.04	−0.13**	−0.07	−0.12**	0.02
Peer relationship problems							
Germany	0.20**	0.18**	0.19**	0.17**	0.13**	0.17**	0.05
Cyprus	0.23**	0.23**	0.23**	0.15**	0.27**	0.14**	−0.02
UK	0.41**	0.39**	0.38**	0.29**	0.41**	0.37**	0.35**
Sweden	0.15**	0.05	0.15**	0.22**	0.12**	0.12**	0.06
Italy	0.01	0.11*	−0.02	0.09*	0.03	0.08	−0.19**

(*Continues*)

Table 2 (Continued)

SDQ	SCAS						
	Total score	Separation anxiety	Social phobia	Obsessive-compulsive disorder	Panic-agoraphobia	Fears of physical injury	Generalized anxiety
Prosocial behaviour							
Germany	0.07	0.07	0.07	−0.02	0.06	−0.02	0.17**
Cyprus	0.11*	0.07	0.03	0.11*	0.08	0.10*	0.22**
UK	−0.04	0.03	0.01	−0.10*	−0.09	−0.002	−0.02
Sweden	0.31**	0.31**	0.34**	0.20**	0.10*	0.19**	0.35**
Italy	0.22**	0.05	0.19**	0.14**	0.19**	0.23**	0.19**
Impact score							
Germany	0.44**	0.18**	0.42**	0.37**	0.26**	0.33**	0.33**
Cyprus	0.23**	0.11	0.21**	0.20**	0.27**	0.11	0.24**
UK	0.42**	0.55**	0.32**	0.31**	0.18**	0.31**	0.37**
Sweden	0.20**	0.03	0.13	0.27**	0.15	0.04	0.23**
Italy	0.30**	0.21**	0.23**	0.31**	0.27**	0.03	0.26**

Note: SCAS, Spence Children's Anxiety Scale; SDQ, Strengths and Difficulties Questionnaire.
*Significant ($p < 0.05$).
**Significant ($p < 0.01$).

test was conducted, according to the method for correlated correlations proposed by Meng *et al.* (1992: two-tailed Z-critical $= 1.96$ for $p < 0.05$ and 2.58 for $p < 0.01$). Z Scores comparing SDQ emotional problems and conduct problems ranged from 6.2 for Germany and Cyprus to 11.5 for Italy. Z Scores comparing SDQ emotional symptoms and hyperactivity-inattention ranged from 3.3 for the UK to 11.4 for Italy, thus indicating that correlations between SCAS anxiety scores and SDQ emotional symptoms were significantly greater than the relationship between SCAS anxiety scores and SDQ conduct problems or hyperactivity-inattention problems. These results support, on the one hand, the convergent validity between the SDQ emotional problems subscale and another emotional symptoms measure, and on the other hand, the divergent validity with externalizing problems.

Factor analysis of SDQ across five countries

Given the relatively low internal consistencies of some of the SDQ subscales, CFA was conducted to compare the factor structure of the SDQ in the five countries. Analyses were carried out using the EQS 6.1 program (Multivariate Software, 2007). Since the observed variables included in

the models were ordinal, a maximum likelihood robust method based on polichoric correlations was used (Satorra and Bentler, 1994). This method is thought to perform as other robust methods for ordinal variables when sample size is adequately large for the model (Lei, 2009).

The five-correlated factor model (Figure 1) reported by Goodman (2001) and the three-correlated factor model (Figure 2) reported in other studies (Dickey and Blumberg, 2004; Koskelainen *et al.*, 2001; Riso *et al.*, 2010) were compared. In this study, the Satorra–Bentler scaled chi-square, the root mean square error of approximation (RMSEA), the standardized root mean square residual (SRMR), the Normed Fit Index (NFI), the Non- Normed Fit Index (NNFI), and the Comparative Fit Index (CFI) were used as goodness-of-fit indices. Values of NFI, NNFI and CFI higher than 0.90 indicate adequate fit (Marsh and Hau, 1996). RMSEA of 0.05 or lower is considered as indicative of an adequate fit (Brown and Cudeck, 1993). Robust versions of all fit statistics were computed except for the SRMR, which has no robust counterpart but which summarizes the fit in a way the other (robust) indices do not. A favourable value of the SRMR is less than 0.10 (Hu and Bentler, 1999). Additionally, the Akaike Information Criterion (AIC), the conditional AIC and the Expected Cross-validation Index
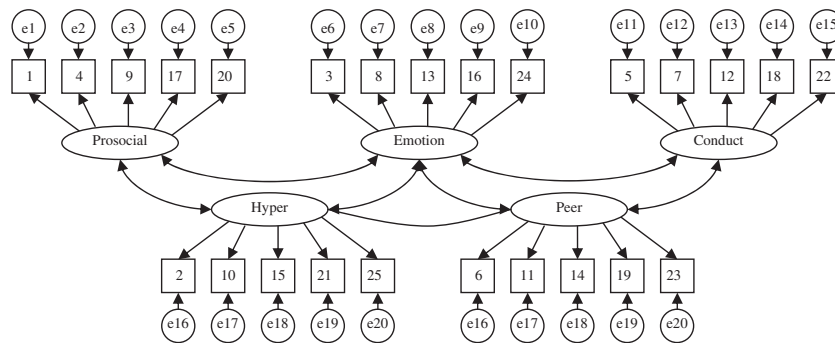

Figure 1 The five-correlated factor model of the SDQ questionnaire. Note: Hyper, hyperactivity-inattention; Peer, peer problems; Prosocial, prosocial behaviour; Emotion, emotional symptoms; Conduct, conduct problems.
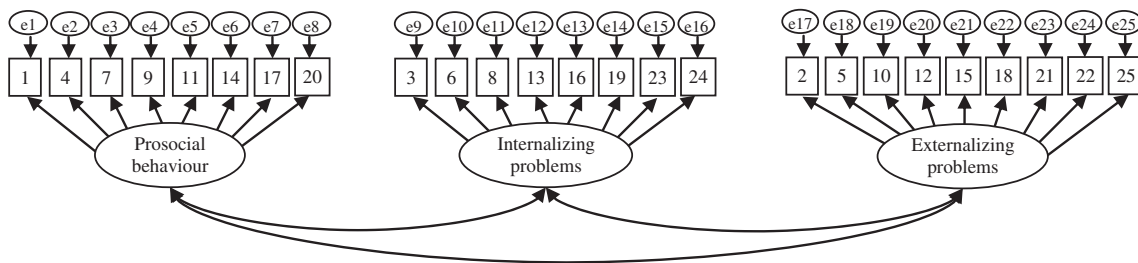

Figure 2 The three-correlated factor model of the SDQ questionnaire.

(ECVI) were used as parsimony indices: smaller model AIC, conditional AIC and ECVI than the comparison model indicates better fit.

Intercorrelation values between the five factors in the whole sample range from 0.69 (prosocial behaviour–emotional symptoms) to 1.00 (hyperactivity-inattention–peer problems). Intercorrelation values for the three factors in the whole sample were 0.68 (prosocial behaviour–internalizing problems), 0.70 (prosocial behaviour–externalizing problems) and 0.86 (internalizing–externalizing problems). The first two columns of Table 3 present the goodness-of-fit indices comparing both models. Fit indices from the five-factor and the three-factor models suggest that both models are acceptable. However, the five-factor model fitted less well than the three-factor model. Comparing the parsimony indices, the AIC, conditional AIC and EVIC indices from the three-factor model are smaller than the comparison model five-factor model, suggesting again that the three-factor model fits better for the overall sample. Factor loadings of the five-factor model in the whole sample ranged from 0.77 to 0.84 (prosocial behaviour), 0.66 to 0.89 (emotional symptoms), 0.46 to −0.74 (conduct problems), 0.51 to −0.75 (hyperactivity-inattention), and 0.49 to −0.79 (peer problems). Regarding the three-factor model, factor loadings in the whole sample ranged from 0.73 to 0.86 (prosocial behaviour), 0.54 to 0.86 (internalizing problems), and 0.56 to 0.74 (externalizing problems). Factor loadings from the CFA in each country showed values ranging from 0.25 to 0.81 for the prosocial problems, with the lowest value corresponding to item 20 ("I often volunteer to help others") in the UK; from 0.14 to 0.84 for the emotional symptoms, with the lowest value corresponding to item 24 ("I have many fears, I am easily scared") in Sweden; from −0.02 to 0.82 for conduct problems, with the lowest value corresponding to item 7 ("I usually do as I am told") in Italy; 0.01 to 0.87 for hyperactivity-inattention, with the lowest value corresponding to item 25 ("I finish my work I'm doing. My attention is good") in the UK; and −0.02 to 0.88 for peer problems with the lowest value corresponding to item 11 ("I have one good friend or more") in the UK.

Another CFA was conducted for the whole sample and separately for each country without the reversed items (Table 3). Comparison between models with and without the reversed items was calculated with the difference value of the Satorra–Bentler χ^2 [calculated with the computation process from Satorra and Bentler (2001)]. When the five-factor model was considered, the goodness-of-fit improved significantly in the overall sample and in all countries. The difference values of the Satorra–Bentler χ^2 were also significant in the overall sample and in all countries ($p < 0.001$). When the three-factor model was considered,

again the goodness-of-fit also improved and the difference values of the Satorra–Bentler χ^2 were significant in the overall sample and in all countries ($p < 0.001$). Intercorrelations between the five factors in the whole sample once the reversed items were removed ranged from 0.49 for correlation between prosocial problems and hyperactivity-inattention to 0.86 for correlation between conduct problems and hyperactivity-inattention. Intercorrelations between the three factors in the whole sample without the reversed items were 0.69 (prosocial behaviour–internalizing problems), 0.54 (prosocial behaviour–externalizing problems), and 0.83 (internalizing–externalizing problems).

Table 3 shows the goodness-of-fit indices comparing the five-factor and three-factor models in the overall sample and across the five countries without the reversed items. The goodness-of-fit indices of the five-factor model was better than the three-factor model in Germany, the UK, Sweden, and Italy, whereas in Cyprus both models showed equal indices. The parsimony indices (i.e. AIC, conditional AIC and ECVI) were smaller in the five-factor model in all countries. Compared with the goodness-of-fit indices of the overall sample including all items, the goodness-of-fit indices of the five-factor model were better than the three-factor model when the reversed items were excluded. Factor loading values in each country once the reversed items were excluded ranged from 0.12 to 0.88 for prosocial behaviour, with the lowest factor loading corresponding to item 20 ("I often volunteer to help others") in the UK; 0.17 to 0.86 for emotional symptoms, with the lowest value corresponding to item 13 ("I am often unhappy, down-hearted or tearful") in the UK; 0.04 to 0.85 for conduct problems, with the lowest value corresponding to item 18 ("I am often accused of lying or cheating") in Italy; 0.08 to 0.92, with the lowest factor loading corresponding to item 10 ("I am constantly fidgeting or squirming") in Sweden; and 0.13 to 0.96, with the lowest value corresponding to item 23 ("I get on better with adults than with people my own age") in the UK. Regarding the three-factor model, factor loading values ranged from 0.12 to 0.88 for prosocial behaviour, with the lowest value corresponding to item 20 ("I often volunteer to help other people") in the UK; 0.05 to 0.87 for internalizing problems, with the lowest factor loading corresponding to item 19 ("Other children or young people pick on me or bully me") in the UK; and from 0.01 to 0.81 for externalizing problems, with the lowest value corresponding to item 22 ("I take things that are not mine from home, school or elsewhere") in the UK.

We also aimed at testing for the equivalence of the SDQ structure across the five countries. To do this, the multi-group invariance procedure described by Byrne (2006)

Table 3 Goodness-of-fit indices for five-factor model and three-factor model in the overall sample and in each country (with and without the reversed items)

| | Reversed items included | | Reversed items excluded | | | | | | | | | | | | |
| | Total sample (N = 2418) | | Total sample (N = 2418) | | Germany (N = 495) | | UK (N = 469) | | Cyprus (N = 471) | | Sweden (N = 484) | | Italy (N = 499) | |
Fit indices	Five-factor	Three-factor	Five-factor	Three-factor	Five-factor	Three-factor	Five-factor	Three-factor	Five-factor	Three-factor	Five-factor	Three-factor	Five-factor	Three-factor
Satorra–Bentler scales chi-square (df)	4048.07 (265)	2906.60 (272)	711.07 (160)	835.62 (167)	434.69 (160)	568.83 (167)	1037.58 (160)	1167.47 (167)	213.77 (160)	236.62 (167)	964.19 (160)	1030.71 (167)	1284.61 (160)	1411.84 (167)
NFI	0.93	0.95	0.98	0.98	0.83	0.77	0.70	0.66	0.98	0.98	0.56	0.53	0.50	0.45
NNFI	0.93	0.95	0.98	0.98	0.86	0.80	0.68	0.65	0.99	0.99	0.52	0.51	0.44	0.41
CFI	0.94	0.96	0.98	0.98	0.88	0.82	0.73	0.69	0.99	0.99	0.60	0.57	0.54	0.48
RMSEA	0.07	0.06	0.04	0.04	0.06	0.06	0.10	0.11	0.03	0.03	0.10	0.10	0.11	0.12
SRMR	0.10	0.10	0.06	0.06	0.09	0.11	0.10	0.10	0.06	0.06	0.10	0.10	0.10	0.11
Model AIC	3518.97	2362.60	391.07	501.62	114.69	234.83	717.58	833.47	−106.22	−97.38	644.19	696.71	964.60	1077.84
Model conditional AIC	1723.19	519.38	−693.30	−630.20	−717.06	−633.31	−106.52	−26.68	−921.56	−948.38	−184.28	−168.0	130.91	207.67
ECVI	1.71	1.23	0.33	0.38	1.08	1.36	2.48	2.74	0.66	0.68	2.75	2.90	2.78	2.88

was followed. The first step was testing for configural invariance, in which no equality constraints are imposed on the parameters and the number of factors and factor-loading pattern are the same across the five groups. The fit of this configural model provides the baseline model against which all subsequent models are compared. The second step requires testing for measurement invariance; that is, testing for the invariance of factor loadings. For this, constraints to equality for all freely estimated factor loadings are specified for the five countries. Finally, a test for invariance related to the structural model was conducted, in which invariance is tested for factor loadings and factor covariances. For this, constraints equal to zero are specified for both factor loadings and covariances between factors.

Table 4 presents results on the test for invariance for the five-factor and three-factor models without reversed items, which includes goodness-of-fit indices like the Satorra–Bentler χ^2, the robust CFI and RMSEA, the SRMR, and the comparison between models including constraints and the configural model (i.e. the model against which all remaining models are compared to determine evidence of invariance). Regarding the test for invariance across countries of the five-factor model, the configural model (with no constraints) showed an adequate well-fitting multi-group model. This suggests that the structure of the SDQ is optimally represented as a five-factor model across the five countries. When factor loading constraints were specified, the goodness-of-fit indices were still acceptable, although the difference value of the Satorra–Bentler χ^2 [calculated with the computation process from Satorra and Bentler (2001)] was significant, thus indicating that the multi-group model underwent some deterioration in model fit. The Lagrange Multiplier test revealed that factor loadings that seemed to be more non-invariant (i.e. non-equivalent) across countries were those from item 4 to prosocial problems ("I usually share with others"), item 13 to emotional symptoms ("I am often unhappy, down-hearted or tearful"), and item 19 to peers problems ("Other children or young people pick on me or bully me"). Test of invariance for factor loadings and covariances between the five factors revealed still acceptable fit indices, although the difference value of the Satorra–Bentler χ^2 was significant, thus indicating deterioration in model fit. The parameters that appeared to contribute to the non-equivalence between countries were factor loading from item 4 to prosocial behaviour ("I usually share with others"), item 8 ("I worry a lot") and item 24 ("I have many fears, I am easily scared") to emotional symptoms, item 12 ("I fight a lot. I can make other people do what I want") and item 18 ("I am often accused of lying or cheating") to conduct problems, item 15 ("I am easily distracted, I find it difficult to concentrate")

Table 4 Goodness-of-fit statistics of tests for invariance of the SDQ five-factor and three-factor structure[a]

	S–B χ^2	df	CFI	SRMR	RMSEA	RMSEA 90% CI	Model comparison	Δ S–B χ^2	Δ df	Δ *CFI
Five-factor model										
Model 1 Configural (no constraints)	2709.60	800	0.928	0.096	0.071	0.068–0.074	—	—	—	—
Model 2 Measurement Model invariant (factor loadings)	3623.43	860	0.882	0.114	0.082	0.079–0.085	2 versus 1	1181.53	60	0.046
Model 3 Structural Model invariant (factor loadings + factor covariances)	4941.97	900	0.827	0.191	0.097	0.094–0.10	3 versus 1	1501.17	100	0.101
Three-factor model										
Model 1 Configural (no constraints)	3056.34	835	0.905	0.102	0.075	0.072–0.078	—	—	—	—
Model 2 Measurement Model invariant (factor loadings)	4086.94	899	0.863	0.124	0.086	0.084–0.089	2 versus 1	695.66	64	0.043
Model 3 Structural Model invariant (factor loadings + factor covariances)	4820.61	911	0.832	0.185	0.095	0.092–0.097	3 versus 1	701.999	76	0.073

[a]CFI and RMSEA are based on a Robust estimating method. S–B χ^2, Satorra–Bentler scales chi-square.

to hyperactivity-inattention, and item 19 ("Other children or young people pick on me or bully me") and item 23 ("I get on better with adults than with people my own age") to peer problems. Among covariables between the five factors, those that appeared as with having most non-invariance were prosocial behaviour–emotional symptoms, prosocial problems–hyperactivity-inattention, emotional symptoms–conduct problems, emotional symptoms–hyperactivity-inattention, emotional symptoms–peer problems, and conduct problems–hyperactivity-inattention.

Test of invariance for the three-factor model revealed a good multi-group fit, although goodness-of-fit indices were slightly lower than those reported for the five-factor multi-group test. When factor loading constraints were specified, the difference value of the Satorra–Bentler χ^2 was significant, indicating a decrement in the multi-group fit. However, as can be seen in Table 4, the goodness-of-fit indices were still acceptable. Factor loadings that appeared to be the most non-equivalent between countries were those from item 4 ("I usually share with others") to prosocial problems, item 18 ("I am often accused of lying or cheating") to externalizing problems, and item 19 ("Other children or young people pick on me or bully me") to internalizing problems. Test of invariance including constraints equal to zero for factor loadings and covariance between the three factors showed also acceptable multi-group fit, although fit indices were poorer than the three-factor configural model and model including only factor loading constraints. The value of the Satorra–Bentler χ^2 was significant, thus indicating again poorer fitting than the three-factor configural model. Factor loadings from item 4 to prosocial behaviour, item 6 ("I play on my own. I generally play alone or keep to myself") to internalizing subscale, and item 18 to externalizing subscale were found as the most non-equivalent factor loadings across the five countries. Covariance between internalizing and externalizing subscales appeared also as not operating equivalently across countries.

Following the traditional approach to report evidence of invariance based on the difference value of the Satorra–Bentler χ^2 (Jöreskog, 1971), it can be concluded that some of the factor loadings and the covariance between factors are non-equivalent across the five countries, both in the five-factor and the three-factor models. If a practical perspective is taken into account, those Δ *CFI values less than 0.01 would be indicative of multi-group invariance (Cheung and Rensvold, 2002). Both the five-factor and the three-factor models displayed Δ *CFI greater than 0.01, thus indicating that some factor loadings and covariance between factors were not equivalent across the groups. When comparing Δ Satorra–Bentler χ^2 and Δ *CFI of the five-factor and the three-factor structure, the lower values were obtained in the three-factor model, suggesting that the three-factor structure might be more equivalent across the five countries than the five factor.

Emotional and behavioural problems in five European countries

The means and standard deviations for the total difficulties, as well as their subscales are presented in Table 5. *Post hoc* tests revealed that the adolescents in Germany and the UK reported significantly higher total difficulties than adolescents in the other three countries. Within the SDQ subscales, adolescents in Germany were significantly higher on emotional symptoms and conduct problems, and on prosocial behaviour compared to adolescents in other countries. Hyperactivity-inattention was significantly higher in the UK than in other countries. Adolescents in Sweden and in the UK scored significantly higher on peer problems than adolescents in the other three countries.

Discussion

The present study examined the internal consistency, convergent validity, and factor structure of the Cypriot

Table 5 Mean and SD values of the SDQ and its subscales in the five countries

	Germany Mean (SD)	Cyprus Mean (SD)	UK Mean (SD)	Sweden Mean (SD)	Italy Mean (SD)	Effects
SDQ total scores	10.93 (4.9)	9.87 (3.8)	10.87 (2.9)	8.99 (2.3)	10.26 (2.3)	Country: $F = 26.15$, $p < 0.001$
Emotional symptoms	3.25 (2.2)	1.13 (1.7)	0.88 (1.1)	0.48 (0.97)	0.83 (1.1)	Country: $F = 252.14$, $p < 0.001$
Conduct problems	2.77 (1.6)	2.41 (1.1)	2.32 (0.84)	2.09 (0.62)	2.39 (0.79)	Country: $F = 26.59$, $p < 0.001$
Hyperactivity-inattention	3.24 (1.9)	3.48 (1.3)	4.58 (1.1)	3.37 (0.98)	4.04 (1.0)	Country: $F = 85.41$, $p < 0.001$
Peer problems	1.68 (1.7)	2.88 (1.7)	3.09 (0.96)	3.06 (0.93)	3.01 (1.1)	Country: $F = 105.18$, $p < 0.001$
Prosocial behaviour	7.03 (1.9)	3.57 (2.7)	2.05 (1.4)	2.51 (1.7)	2.40 (1.7)	Country: $F = 565.38$, $p < 0.001$

Greek, Swedish, Italian and German translations of the SDQ and its original English version. To our knowledge, the present study is the first to compare and contrast simultaneously the SDQ among adolescents in more than two countries. Our findings can be summarized as follows: In all five countries, the internal consistency of the total SDQ ranged from satisfactory to good, replicating several previous studies, with the exception of the one found in Italy where the internal consistency was not satisfactory (e.g. Capron *et al.*, 2007; Hawes and Dadds, 2004; Muris *et al.*, 2003, 2004; Riso *et al.*, 2010; van Widenfelt *et al.*, 2003; Yao *et al.*, 2009). However, similar to those reported in previous studies (Riso *et al.*, 2010), the α coefficients of some SDQ subscales in the Cypriot and Swedish data were below acceptable limits. Consistent with findings by Riso *et al.* (2010) in Italy and Muris *et al.* (2004) in the Netherlands, these subscales with poor internal consistencies are subscales those with reversed items (i.e. conduct and hyperactive-inattention problems, peer problems). Similar to findings reported by Riso *et al.* (2010), in our study, removal of the reversed items ("I usually do as I am told"; "I have one good friend or more"; "Other people my age generally like me"; "I think before I do things"; "I finish the work I'm doing. My attention is good") led to an increase of the Cronbach's α. Furthermore, given the low internal consistency for conduct problem subscale, some authors have suggested that this subscale should be improved in the construction, semantics and the wording of the various items (Rønning *et al.*, 2004).

To examine the convergent validity of the SDQ, correlation coefficients were calculated between the SDQ and the SCAS. Within the SDQ subscales, the strongest correlation was between emotional symptoms and total anxiety symptoms. In line with previous studies (Ishikawa *et al.*, 2009; Muris *et al.*, 2004; Spence, 1997), our results showed moderate to strong correlations between SDQ emotional symptoms and SCAS anxiety symptoms. Our findings also showed significant positive correlations between conduct problems and hyperactivity-inattention problems and the total anxiety scores in Germany, Cyprus, and the UK. This result could be interpreted as being in line with the high comorbidity between anxiety, conduct disorder, and attention deficit hyperactivity disorder (ADHD) (Cunningham and Ollendick, 2010).

Adolescents in Germany and the UK reported significantly more total difficulties than adolescents in the other three European countries, with the lowest SDQ scores found in Sweden. The mean scores of the total difficulties in the UK and Germany however were comparable to those found previously in the Netherlands (Muris *et al.*, 2004) and China (Yao *et al.*, 2009), but lower than those reported in Italy (Riso *et al.*, 2010). The reason for the self-report of a high level of emotional and behavioural problems in Germany and the UK are unclear. The higher total difficulties scores found in the UK seemed to be consistent with previous studies that showed high prevalence of mental health problems among British young people. Specifically, 10% of the 5–16 year olds were reported to have a clinically diagnosed mental disorder, including emotional disorders (4%), conduct disorder (6%), and hyperkinetic disorder (2%) (Meltzer *et al.*, 2000). Furthermore, in another study, British parents tended to report less marital satisfaction and they considered being a parent as more stressful compared to parents in Belgium, Denmark, and Sweden (Nekkebroeck *et al.*, 2010); these authors interpreted this finding as related to societal attitudes towards the value of being a parent, and to employers' attitudes regarding the needs of parents to be flexible with their working hours. Additionally, the UK fares less well in child-care provision for early childhood (Department for Education and Skills, 2006), which may add to feelings of stress and more complicated family lives. It could be speculated that these stressful situations at home lead to emotional and behavioural problems. The high total difficulties scores in Germany are also in line with the high prevalence of mental disorders among German adolescents. As reported in several epidemiological studies in Germany, up to 20% of adolescents meet the diagnosis of anxiety and depressive disorders some time in their early lives (Essau, 2002; Essau *et al.*, 2000; Wittchen *et al.*, 1998). Some of the most consistent risk factors for the development of mental health disorders among German adolescents include school stress and changes in family situations (e.g. parental divorce and unemployment), and parental psychopathology (Essau, 2002).

There has been considerable controversy in previous studies regarding the generalizability of the five-factor model proposed by Goodman (1999). Consistent with previous studies, the fit indices indicated that both the five-factor and the three-factor models provided good fit for the whole sample; however, in comparing both models, the three-factor model seemed to fit better. Still, the factor structure differed across countries. Specifically, the three-factor model which includes prosocial behaviour, internalizing and externalizing problems showed better fit indices in Cyprus, whereas the five-factor model fitted better in Germany. Fit indices for the UK, Sweden, and Italy were relatively poor for both models.

By removing the reversed items, the goodness-of-fit improved significantly for the total sample and in each country, and the five-factor structure showed the best-fitting indices in the whole sample and in each country. This finding supports previous findings that the reversed items tend to confound the SDQ factors (Palmieri and Smith, 2007; Riso *et al.*, 2010). It is therefore important to consider removing or re-wording the reversed items, and keep the five-factor structure of the SDQ.

Test of invariance showed that the SDQ structure might be different across the five countries. An alternative factor structure could be analysed in future studies to reach a best-fitting model specific for each country. For example, it might be possible that a different number of factors appear or that cross-factor loadings should be added to obtain a model that fits adequately for a specific population.

There are several limitations to the present study, which need to be taken into consideration when interpreting our findings. First, our participants were 12- to 17-year-olds who were recruited from urban schools. Although a convenience sample and seemingly representative of urban youth in those countries, use of such samples may have implications for the generalizability of our findings to other youth in these and other European countries. Second, the data were based solely on the adolescents' self-report and no behavioural observations or clinical indices were used to confirm this self-report measure. Further test of convergent and discriminant validity should be investigated via parent or teacher report. Unfortunately, in the present study, we were unable to receive ethical approval to include parents and teachers as informants. Third, emotional and behavioural symptoms were assessed by means of questionnaires. Future studies might employ clinical diagnostic interviews in order to examine how impairing the reported symptoms are and to further establish the validity of child and adolescent psychopathology among children and adolescents in various European countries. Fourth, our findings may be limited by issues which plague all cross-cultural research such as difference in response styles. However, we are not aware of research that supports differences in response styles across the countries examined in this study. Finally, although the SDQ was administered in each participant's mother tongue, the content of the items might have different meanings across cultures (Berry *et al.*, 2002). These limitations notwithstanding, our findings support the usefulness of the SDQ as an efficient way of screening for emotional and behavioural problems in children and adolescents in most of these five European countries.

Declaration of interest statement

The authors have no competing interests.

References

Achenbach T.M. (1991) *Manual for the Youth Self-report and 1991 Profile*, Burlington: VT, University of Vermont.

American Psychiatric Association. (1994) *Diagnostic and Statistical Manual of Mental Disorders, 4th Edition (DSM-IV)*, Washington, DC: American Psychiatric Association.

Berry J.W., Poortinga Y.H., Segall M.H., Dasen P.R. (2002) Methodological concerns. In Berry J.W., Poortinga Y.H., Segall M.H., Dasen P.R. (eds) *Cross-cultural Psychology: Research and Applications*, pp. 286–316, Cambridge: Cambridge University Press.

Brown M., Cudeck R. (1993) Alternative ways of assessing model fit. In Bollen K., Long J. (eds) *Testing Structural Equation Models*, pp. 136–162, London: Sage.

Byrne B.M. (2006) *Structural equation modeling with EQS: Basic concepts, applications and programming (2nd edition)*, Mahwah, NJ: Erlbaum.

Capron C., Therond C., Duyme M. (2007) Psychometric properties of the French version of the self-report and teacher Strengths and Difficulties Questionnaire (SDQ). *European Journal of Psychological Assessment*, **23**, 79–88.

Cheung G.W., Rensvold R.B. (2002) Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, **10**, 238–262.

Cunningham N.R., Ollendick T.H. (2010) Comorbidity of anxiety and conduct problems in children: implications for clinical research and practice. *Clinical Child and Family Psychology Review*, **13**, 333–347.

Department for Education and Skills. (2006) *Choice for Parents. The Best Start for Children: Making it Happen. An Action Plan for the Ten Year Strategy*, London: The Stationery Office.

Dickey W.C., Blumberg S.J. (2004) Revisiting the factor structure of the Strengths and Difficulties Questionnaire: United States, 2001. *Journal of the American Academy of Child and Adolescent Psychiatry*, **43**, 1159–1167.

Essau C.A. (2002) *Depression bei Kindern und Jugendlichen*, München: Ernst Reinhardt Verlag.

Essau C.A., Conradt J., Petermann F. (2000) Frequency, comorbidity, and psychosocial impairment of anxiety disorders in adolescents. *Journal of Anxiety Disorders*, **14**, 263–279.

Goodman R. (1997) The Strengths and Difficulties Questionnaire: a research note. *Journal of Child Psychology and Psychiatry*, **38**, 581–586.

Goodman R. (1999) The extended version of the Strengths and Difficulties Questionnaire as a guide to child psychiatric caseness and consequent burden. *Journal of Child Psychology and Psychiatry*, **40**, 791–799.

Goodman R. (2001) Psychometric properties of the Strengths and Difficulties Questionnaire. *Journal of the American Academy of Child and Adolescent Psychiatry*, **40**, 1337–1345.

Goodman R., Ford T., Richards H., Gatward R., Meltzer H. (2000) The development and well-being assessment: description and initial validation of an integrated assessment of child and adolescent psychopathology. *Journal of Child Psychology and Psychiatry*, **41**, 645–656.

Hawes D., Dadds M.R. (2004) Australian data and psychometric properties of the Strengths and Difficulties Questionnaire. *The Australian and New Zealand Journal of Psychiatry*, **38**, 644–651.

Hu L., Bentler P.M. (1999) Cutoff criteria for fit indexes in covariance structure analysis. Conventional criteria versus new alternatives. *Structural Equation Modeling*, **6**, 1–55.

Ishikawa S., Sato H., Sasagawa S. (2009) Anxiety disorder symptoms in Japanese children and adolescents. *Journal of Anxiety Disorders*, **23**, 104–111.

Jöreskog K.G. (1971) Simultaneous factor analysis in several populations. *Psychometrika*, **36**, 409–426.

Klasen H., Woerner W., Rothenberger A., Goodman R. (2003) German version of the Strength and Difficulties Questionnaire (SDQ-German) – overview and evaluation of initial validation and normative results. *Praxis der Kinderpsychologie und Kinderpsychiatrie*, **52**, 491–502.

Koskelainen M., Sourander A., Vauras M. (2001) Self-reported strengths and difficulties in a community sample of Finnish adolescents. *European Child & Adolescent Psychiatry*, **10**, 180–185.

Lei P.W. (2009) Evaluating estimation methods for ordinal data in structural equation modeling. *Quality and Quantity*, **45**, 495–507.

Marsh H.W., Hau K.T. (1996) Assessing goodness-of-fit: when parsimony is undesirable. *The Journal of Experimental Education*, **64**, 364–390.

Meltzer H., Gatward R., Goodman R., Ford T. (2000) *Mental Health of Children and Adolescents in Great Britain*, London: Stationery Office.

Meng X.L., Roesnthal R., Rubin D.B. (1992) Comparing correlated correlation coefficients. *Psychological Bulletin*, **111**, 172–175.

Multivariate Software. (2007) *EQS – Structural Equation Modeling Software*, Encino, CA: Multivariate Software.

Muris P., Meesters C., Eijkelenboom A., Vincken M. (2004) The self-report version of the Strengths and Difficulties Questionnaire: its psychometric properties in 8- to 13-year-old non-clinical children. *British Journal of Clinical Psychology*, **43**, 437–448.

Muris P., Meesters C., van den Berg F. (2003) The Strengths and Difficulties Questionnaire (SDQ) – further evidence for its reliability and validity in a community sample of Dutch children and adolescents. *European Child & Adolescent Psychiatry*, **12**, 1–8.

Nekkebroeck J., Barnes J., Bonduelle M., Wennerholm U.-B., Ponjaert-Kristoffersen I., Loft A., Sutcliffe A.G. (2010) International comparison of parenting styles in ICSI, IVF and natural conception families: results from a European study. *The European Journal of Developmental Psychology*, **7**(3), 329–349. DOI: 10.1080/17405620802217547

Palmieri P., Smith G. (2007) Examining the structural validity of the Strengths and Difficulties Questionnaire (SDQ) in a US sample of custodial grandmothers. *Psychological Assessment*, **19**, 189–198.

Reynolds C.R., Richmond B.O. (1978) What I think and feel: a revised measure of children's manifest anxiety. *Journal of Abnormal Child Psychology*, **6**, 271–280.

Riso D.D., Salcuni S., Chessa D., Raudino A., Lis A., Altoè G. (2010) The Strengths and Difficulties Questionnaire (SDQ). Early evidence of its reliability and validity in a community sample of Italian children. *Personality and Individual Differences*, **49**, 570–575.

Rønning J.A., Handegaard B.H., Sourander A., Mørch W.-T. (2004) The Strengths and Difficulties Self-report Questionnaire as a screening instrument in Norwegian community samples. *European Child & Adolescent Psychiatry*, **13**, 73–82.

Satorra A., Bentler P.M. (1994) Corrections to test statistics and standard errors in covariance structure analysis. In von Eye A., Clogg C.C. (eds) *Latent Variables Analysis: Applications for Developmental Research*, pp. 399–419, Thousand Oaks, CA: Sage.

Satorra A., Bentler P.M. (2001) A scaled difference chi-square test statistic for moment structure analysis. *Psychometrika*, **66**, 507–514.

Smedje H., Broman J.-E., Hetta J., von Knorring A.-L. (1999) Psychometric properties of a Swedish version of the "Strengths and Difficulties Questionnaire". *European Child & Adolescent Psychiatry*, **8**, 63–70.

Spence S.H. (1997) The structure of anxiety symptoms among children: a confirmatory factor analytic study. *Journal of Abnormal Psychology*, **106**, 280–297.

Steiger J.H. (1980) Test for comparing elements of correlations matrix. *Psychological Bulletin*, **87**, 245–281.

van Widenfelt B.M., Goedhart A.W., Treffers P.D., Goodman R. (2003) Dutch version of the Strengths and Difficulties Questionnaire (SDQ). *European Child & Adolescent Psychiatry*, **12**, 281–289.

Wittchen H.-U., Nelson C.B., Lachner G. (1998) Prevalence of mental disorders and psychosocial impairments in adolescents and young adults. *Psychological Medicine*, **28**, 109–126.

Yao S., Zhang C., Zhu X., Jing X., McWhinnie C. M., Abela R.R.Z. (2009) Measuring adolescent psychopathology: psychometric properties of the self-report Strengths and Difficulties Questionnaire in a sample of Chinese adolescents. *Journal of Adolescent Health*, **45**, 55–62.