

Cross-cultural measurement invariance of the Revised Child Anxiety and Depression Scale across 11 world-wide societies

D. Stevanovic^{1*}, Z. Bagheri², O. Atilola³, P. Vostanis⁴, D. Stupar¹, P. Moreira⁵, T. Franic⁶, N. Davidovic⁶, R. Knez⁷, A. Nikšić⁸, K. Dodig-Ćurković⁹, M. Avicenna¹⁰, I. Multazam Noor¹¹, L. Nussbaum¹², A. Deljkovic¹³, A. Aziz Thabet¹⁴, P. Petrov¹⁵, D. Ubalde¹⁶, L. A. Monteiro¹⁷ and R. Ribas^{18†}

¹ Clinic for Neurology and Psychiatry for Children and Youth, Belgrade, Serbia

² Department of Biostatistics, Shiraz University of Medical Sciences, Shiraz, Iran

³ Lagos State University, Lagos, Nigeria

⁴ School of Psychology, Leicester University, Leicester, UK

⁵ University Lusitana, Lisboa, Portugal

⁶ Child and Adolescent Psychiatry, School of Medicine, University of Split, Split, Croatia

⁷ Medical School, University of Rijeka, Croatia

⁸ Department of Psychology, Faculty of Humanities and Social Sciences in Rijeka, Croatia

⁹ Medical Faculty Osijek, University Health Center Osijek, Croatia

¹⁰ Faculty of Psychology, State Islamic University Syarif Hidayatullah, Jakarta, Indonesia

¹¹ Dr Soeharto Heerdjan Mental Hospital, Jakarta, Indonesia

¹² Department of Child and Adolescent Psychiatry, University of Medicine and Pharmacy 'Victor Babes', Timisoara, Romania

¹³ Mental Health Center Pljevlja, Montenegro

¹⁴ Al Quds University, Gaza

¹⁵ Department of Child and Adolescent Psychiatry, University Hospital St. Marina, Varna, Bulgaria

¹⁶ Department of Psychology, St. Dominic College of Asia, City of Bacoor, Philippines

¹⁷ Universidade Estacio de Sá in Rio de Janeiro, Brazil

¹⁸ Federal University of Rio de Janeiro, Brazil

Background. In order to compare estimates by one assessment scale across various cultures/ethnic groups, an important aspect that needs to be demonstrated is that its construct across these groups is invariant when measured using a similar and simultaneous approach (i.e., demonstrated cross-cultural measurement invariance). One of the methods for evaluating measurement invariance is testing for differential item functioning (DIF), which assesses whether different groups respond differently to particular items. The aim of this study was to evaluate the cross-cultural measurement invariance of the Revised Child Anxiety and Depression Scale (RCADS) in societies with different socioeconomic, cultural, and religious backgrounds.

Methods. The study was organised by the International Child Mental Health Study Group. Self-reported data were collected from adolescents residing in 11 countries: Brazil, Bulgaria, Croatia, Indonesia, Montenegro, Nigeria, Palestinian Territories, the Philippines, Portugal, Romania and Serbia. The multiple-indicators multiple-causes model was used to test the RCADS items for DIF across the countries.

Results. Ten items exhibited DIF considering all cross-country comparisons. Only one or two items were flagged with DIF in the head-to-head comparisons, while there were three to five items flagged with DIF, when one country was compared with the others. Even with all cross-culturally non-invariant items removed from nine language versions tested, the original factor model representing six anxiety and depressive symptoms subscales was not significantly violated.

Conclusions. There is clear evidence that relatively small number of the RCADS items is non-invariant, especially when comparing two different cultural/ethnic groups, which indicates on its sound cross-cultural validity and suitability for cross-cultural comparisons in adolescent anxiety and depressive symptoms.

Received 2 March 2016; Accepted 12 May 2016; First published online 29 June 2016

Key words: Adolescents, anxiety, children, depression, instrument.

* Address for correspondence: D. Stevanovic, Clinic for Neurology and Psychiatry for Children and Youth, Belgrade, Serbia.

(Email: dejanstevanovic@eunet.rs)

† There was an error of omission in the title that has now been corrected and a notice has been published providing details.

Introduction

Studying psychopathology in multinational contexts is not only important for estimating prevalence rates, but also in evaluating the whole range of interplay of

genetic, neurobiological and environmental correlates, as well as various treatment options, while eliciting the same aspects of certain psychopathology across different cultures. From a psychometrical point of view, the latter is only possible if assessment methods operate in the same way and underlying constructs have the same theoretical structure for two or more cultural groups, in order to prevent biased estimations (Dimitrov, 2010). In other words, cross-cultural validity of an assessment tool is assumed if its assessment methods operate in the same way and its underlying constructs have the same theoretical structure for two or more cultural groups (Dimitrov, 2010). The trend had been for researchers to assume that the replication of the theoretical construct of an assessment scale developed in one language/culture in another guarantees its cross-cultural equivalence and suitability of such tool for cross-cultural comparative estimation (Byrne & Watkins, 2003). This assumption is being increasingly debunked as it has been established that a prerequisite for establishing measurement equivalence of an assessment scale is that the theoretical construct is measured in each culture in the same way and tested simultaneously across the cultural/ethnic groups (He & van de Vijver, 2012). Therefore, in order to compare estimates by one assessment scale across various cultures/ethnic groups, an important aspect that needs to be demonstrated is that its factorial structure across different ethnic/cultural groups is invariant when measured using a similar and simultaneous approach (i.e., demonstrated cross-cultural measurement invariance; Byrne & Watkins, 2003; Gregorich, 2006; Dimitrov, 2010; Milfont & Fisher, 2010).

Cross-cultural studies over the past two decades have showed that adolescents from various cultures report different levels of anxiety and depressive symptoms (e.g., Charman & Pervova, 1996; Ollendick *et al.* 1996; Stewart *et al.* 2004; Essau *et al.* 2008, 2011). However, to which extent are these estimations bias-free is questionable, because a very few scales for anxiety and depressive symptoms have data for cross-cultural measurement invariance (Stevanovic *et al.* 2016). There is some evidence that scales such as the Fear Survey Schedule for Children-Revised (Ollendick, 1983), Multidimensional Anxiety Scale for Children (March *et al.* 1997), Revised Child Anxiety and Depression Scale (RCADS) (Chorpita *et al.* 2000), Revised Children's Manifest Anxiety Scale (Reynolds & Richmond, 2000), and Mood and Feelings Questionnaire (Angold *et al.* 1995), might provide invariant measurements across different ethnic/cultural groups, but more work is needed to claim that these scales are suitable for cross-cultural comparisons.

The RCADS (Chorpita *et al.* 2000) has drawn significant research and clinical attention recently. The

RCADS was developed with an idea to capture various anxiety and depressive symptoms at once considering significant co-morbidity between the two types of psychopathology. The scale includes the main features of separation anxiety disorder (SAD), social phobia (SP), generalised anxiety disorder (GAD), panic disorder (PD), obsessive-compulsive disorder (OCD), and major depressive disorder (MDD), represented as a theoretical six-factor construct (Chorpita *et al.* 2000). Considering either the self- or parent report of the scale in clinical and community samples of children and adolescents, the six-factor structure was replicated to a fair extent in various psychometric studies (i.e., de Ross *et al.* 2002; Chorpita *et al.* 2005; Ebesutani *et al.* 2010, 2011; Brown *et al.* 2013), including several language versions, such as Danish (Esbjörn *et al.* 2012), Dutch (Mathyssek *et al.* 2013; Kösters *et al.* 2015), French (Bouvard *et al.* 2015), or Spanish (Sandín *et al.* 2009; Park *et al.* 2015). Two studies are also available reporting the cross-cultural measurement invariance data (Latzman *et al.* 2011; Trent *et al.* 2012). Both studies used the original English version to evaluate measurement invariance across African and White American children and adolescents. The results supported the six-factor structure in both ethnicity groups, with similar item cluster patterns mapping well onto the six-factor structure of the RCADS and the same general item-factor relationships in both studies. The latter study identified five non-invariant items across the two groups (Trent *et al.* 2012).

The cross-cultural measurement invariance of the RCADS is yet to be established beyond among ethnic populations in the USA including other cultural groups with different socioeconomic, cultural and religious backgrounds to claim it is suitable for cross-cultural comparisons. As part of a larger project assessing psychometric properties of several adolescent psychopathology rating scales and trauma correlates under the auspices of the Phase II project of the International Child Mental Health Study Group (ICMH-SG, for details see www.icmhsg.org; Atilola *et al.* 2013; Stevanovic *et al.* 2014), we report data on the cross-cultural measurement invariance of the RCADS self-report from adolescents sampled in 11 countries world-wide, namely Brazil, Bulgaria, Croatia, Indonesia, Montenegro, Nigeria, Palestinian Territories, the Philippines, Portugal, Romania and Serbia. Our research design addressed two important aspects for cross-cultural invariance assessments. First, more than two different cultures/ethnic groups were included to establish cross-cultural invariance, because with more groups included, there would be greater variability in the measuring construct among the groups to be detected for cross-cultural invariance.

Second, several language versions were included, because demonstrating cross-cultural measurement invariance for the construct of one scale in the country of its origin does not imply that its construct is transferred into its translations. Thus, the underlying construct represented by the scale needs to be simultaneously tested and confirmed using different language versions for documenting its cross-cultural measurement invariance.

Methods

Participants

This study was conducted simultaneously across the included countries from October 2014 to June 2015. Participants represented a sample of convenience from rural and urban communities based on the location of the researchers. The same recruitment procedure was followed across all locations. Samples from each location were drawn from at least five schools within a frame of randomly selected schools within a well-defined local, political, or administrative zone that is convenient for the participating researcher. At least one school is however drawn from both a rural and an urban setting. It was estimated that valid data from at least 250 participants were needed, but all researchers from particular country in the project were encouraged to include more participants if financially and technically possible. Ethical approvals were obtained in all countries from the appropriate local authorities and/or ethical committees. The adolescents were contacted by school psychologists/counsellors/teachers and were informed about the study. Of all contacted, only those who agreed to participate and returned the written self/parental consents (depending on age) were included.

Instruments

RCADS

The RCADS is a 47-item scale about anxiety and depressive symptoms in youths (Chorpita *et al.* 2000, 2005). The six subscales correspond to SAD (7 items), SP (9 items), GAD (6 items), PD (9 items), OCD (6 items) and MDD (10 items). The respondent indicates on a 4-point scale how often each symptom is present ranging from 'never' (0 point) to 'always' (3 points). The sum of all answered items in a subscale is the total raw score of that subscale. The RCADS also yields a Total Anxiety scale (sum of all scores from anxiety subscales) and a Total Internalizing scale (sum of all subscales' scores). The RCADS was professionally translated and culturally adapted into

Arabic, Bulgarian, Bahasa Indonesia, Croatian, Montenegrin, Portuguese for Portugal, Portuguese for Brazil, Romanian, Serbian and Yoruba languages following the same procedure in each country: two forward translations, a single form development, a single back-translation and the pre-testing (i.e., cognitive debriefing).

Family affluence scale (FAS)

Information on socioeconomic status (SES) was collected using the FAS, a valid indicator of adolescents' material circumstances (Boyce *et al.* 2006). The FAS is a self-report questionnaire that provides information about indicators of familial wealth using four parameters, including family car ownership, adolescent's own bedroom, family ownership of a computer and family holiday in the previous 12 months. The FAS was also professionally translated and culturally adapted into the above mentioned languages following the same procedures. A composite FAS score was calculated for each participant based on the responses to these items, with possible scores ranging between 0 and 9. The greater the obtained score, the higher the SES.

Data analysis

Categorical confirmatory factor analysis (CCFA) was used to examine whether the six-factor model fitted the data well in each country. The reliability of the questionnaire in each country was assessed by Cronbach's alpha coefficient. Cross-cultural measurement invariance of the RCADS was assessed through the multiple-indicators multiple-causes (MIMIC) model (Joreskog & Goldberger, 1975). From a technical point of view, the MIMIC is an extension of the confirmatory factor analysis model that includes covariates, thus the effects of confounding variables can be controlled while measurement invariance is examined across the groups of interest. The MIMIC consists of two basic components, namely measurement and structural components. The measurement component of the model relates one or more latent variables to its indicators; here measurement component modelled the relationship between six latent variables related to RCADS subscales and items. In the structural component, the association between latent variables and covariates such as age and gender was modelled. To test items for differential item functioning (DIF) as an indicator of measurement non-invariance across different countries, all items were regressed on each country, and a significant direct effect from the country of interest to an item indicated measurement non-invariance.

Our primary analysis showed that there were significant differences in terms of age, gender, types of religion and the FAS score across different countries. Therefore, assessing the measurement invariance across the countries was conducted, while controlling the effects of these covariates.

In order to perform all pairwise comparisons across all countries, we conducted MIMIC analysis nine times, where each country was considered in turn as the reference group, while the other nine countries were considered as the focal groups by defining nine indicator variables for each country. To detect items with DIF, we estimated a MIMIC model in which all possible direct effects from items to covariates were constrained to zero. If this model fits the data well, it would suggest that there is no DIF item; however, a model with poor fit would imply source of non-invariance. Large magnitudes of modification indices would suggest that which items are likely to show DIF (Bye *et al.* 1985). Afterwards, the paths from a probable DIF item to the indicator variables of interest were freely estimated, one at a time (in an iterative process); and in a case of significant direct effect, the items were distinguished as DIF.

The fit of invariant and non-invariant models was evaluated using several criteria: chi-square statistics, root mean square error of approximation (RMSEA), Tucker-Lewis index (TLI), and comparative fit index (CFI). However, because chi-square statistics detect even trivial differences in a large sample size (Cheung & Rensvold, 2002), we used the other above-mentioned indices. The CFI and TLI values >0.90 indicate acceptable and >0.95 excellent model fit, while the RMSEA values <0.08 indicate acceptable and <0.06 excellent model fit. We applied the Mean and Variance-adjusted Weighted Least Square estimation procedure being proposed for ordinal indicators in the Mplus 6.1 software to fit CCFA and MIMIC models (Muthén & Muthén, 1998–2010).

Results

Complete data for the present study were available from 3,908 adolescents aged 13–18 years from 11 countries: Brazil ($n=289$), Bulgaria ($n=263$), Croatia ($n=548$), Indonesia ($n=452$), Montenegro ($n=328$), Nigeria ($n=365$), Palestinian Territories ($n=319$), Philippines ($n=284$), Portugal ($n=628$), Romania ($n=329$), and Serbia ($n=387$; Table 1). There were significant differences across the countries in gender ($\chi^2(10)=81.06, p<0.001$), age ($F(10)=136.63, p<0.001$), socioeconomic status – FAS score ($F(10)=64.62, p<0.001$), and religion distribution ($\chi^2(40)=5646.42, p<0.001$, Appendix 1).

CCFA

The value of Cronbach's alpha coefficient for the whole scale in all countries was >0.88 implying on adequate reliability (Table 1). Table 2 shows the results of CCFA for each country. The CFI, TLI and RMSEA values indicated an acceptable fit to the data for Bulgaria, Croatia, Indonesia, Montenegro, Serbia, Palestinian Territories and Portugal. For Brazil, Nigeria and Romania the CFI and/or TLI values were slightly below acceptable values (≤ 0.87), while the RMSEA values were all below 0.08, indicating an acceptable fit to the data. However, all three fit indexes indicated a poor fit to the data for the Philippines, even after considering modification indexes, thus this country was omitted from the subsequent analyses.

Cross-cultural measurement invariance

Considering that the assessment of measurement invariance by the MIMIC model is an iterative process, only the results of the parameter estimates of the items showing DIF in the final step are presented in Table 3. As indicated, each country was in turn considered as the reference country and, due to space limitation, simply the focal countries in which DIF items were detected across those and the reference country are presented. Overall, only ten items exhibited DIF considering all cross-country comparisons; namely items 1, 9, 10, 12, 13, 19, 32, 42, 46 and 47. These items were from all subscales except the PD subscale, which had none DIF item.

As the parameter estimates suggested, Serbian adolescents reported lower scores (indicated by negative values of β coefficients) for items 19 and 47 in the MDD subscale compared with adolescents from Indonesia and Croatia, respectively. In contrast, Serbians reported higher scores (indicated by positive values of β coefficients) for items 9 and 46 in the SAD subscale and item 42 in the OCD subscale as opposed to Croatians, Indonesians and Romanians, respectively. When Nigeria was considered as the reference country, lower scores were observed for items 19 and 47 in the MDD subscale, items 9 and 46 in the SAD subscale, and item 1 in the GAD subscale compared with Indonesians, Croatians, Serbian and Brazilian, respectively. However, adolescents from Nigeria reported higher scores for item 42 in the OCD subscale compared with adolescents from Romania. In addition, adolescents from Indonesia reported lower scores for item 9 compared with adolescents from Serbia and Montenegro, but higher scores for items 1 and 42 than adolescents from Croatia and Romania. In the case of Bulgaria as the reference group, these adolescents also reported lower scores for items 19

Table 1. Distribution of Participants by Gender, Age and Socioeconomic status across countries, with Cronbach's alpha coefficient for the RCADS total score

Country	Gender, Male/female <i>n</i> (%)*	Age, M (s.d.) years**	FAS score, M (s.d.) years***	Cronbach's alpha coefficient
Brazil, <i>n</i> = 289	123 (42.6)/166 (57.4)	13.13 (1.24)	4.74 (2.07)	0.95
Bulgaria, <i>n</i> = 263	120 (45.6)/143 (54.4)	14.88 (1.41)	5.67 (1.85)	0.93
Croatia, <i>n</i> = 548	165 (30.1)/383 (69.9)	16.45 (1.01)	5.25 (1.77)	0.94
Indonesia, <i>n</i> = 452	172 (38.1)/280 (61.9)	15.35 (1.33)	3.97 (2.06)	0.92
Montenegro, <i>n</i> = 328	139 (42.4)/189 (57.6)	15.66 (1.51)	5.36 (1.99)	0.95
Nigeria, <i>n</i> = 365	159 (43.6)/206 (56.4)	14.56 (1.43)	3.06 (2.11)	0.91
Palestine, <i>n</i> = 319	163 (51.1)/156 (48.9)	14.98 (2.01)	4.25 (1.86)	0.92
Philippines, <i>n</i> = 284	156 (54.9)/128 (45.1)	16.35 (0.68)	4.65 (1.89)	0.88
Portugal, <i>n</i> = 628	303 (48.2)/325 (51.8)	15.67 (1.84)	5.04 (2.12)	0.96
Romania, <i>n</i> = 329	163 (49.5)/166 (50.5)	15.28 (1.47)	5.39 (1.87)	0.92
Serbia, <i>n</i> = 387	156 (40.3)/231 (59.7)	15.57 (1.51)	4.55 (1.89)	0.94

* χ^2 (10) = 81.06, $p < 0.001$; ** F (10) = 136.63, $p < 0.001$; *** F (10) = 64.62, $p < 0.001$.

and 47 as compared with adolescents from Indonesia and Croatia; while reporting higher scores for item 42 than adolescents from Romania. When Croatia was the reference country, the negative values of β coefficients for item 12 in the SP subscale, item 19 in the MDD and item 1 in the GAD indicated that these adolescents reported lower scores than adolescents from Nigeria and Indonesia. The positive values of β coefficients indicated that they reported higher scores for items 9 and 42 as opposed to adolescents from Indonesia and Romania. When Palestine was the reference country, items 12 and 19 had lower scores (indicated by negative values of β coefficient) as

compared with Nigeria and Indonesia, while items 9, 1 and 42 were with higher scores compared with Indonesia, Croatia and Romania. In the case of Montenegro as the reference group, items 9 and 19 exhibited uniform DIF across this country and Indonesia, item 47 across Montenegro and Croatia and item 42 across Montenegro and Romania.

Furthermore, the positive values of β coefficients indicated that adolescents from Romania reported higher scores in all the items with DIF than adolescents from Serbia (item 1), Indonesia (item 32), Bulgaria (items 10 and 47), Montenegro (item 1), Portugal (item 1) and Brazil (item 1) except for the item 13 in the GAD subscale, where they reported lower scores than adolescents from Indonesia. Three items were also detected as uniform DIF across Portugal and Nigeria (item 12), Portugal and Croatia (item 47) as well as Portugal and Romania (item 42). Finally, Brazilians also reported lower scores than Nigerians in item 12, Croatians in item 47 and Portuguese in item 12; however, they had higher scores for items 9 and 42 as compared with adolescents from Indonesia and Romania, respectively.

The further analyses considered the RCADS after removing the DIF items. From the SAD subscale, it were removed items 9 and 46, from the SP items 12 and 32, from the GAD items 1 and 13, from the OCD items 10 and 42, and from the MDD items 19 and 47, while in the PD all items remained whereas no DIF was detected. We rerun the CCFA on the remaining 37 items represented by the original six-factor model (Table 2). The fit indexes changed only slightly, and their values again indicated on an acceptable fit to the data for Bulgaria, Croatia, Indonesia, Montenegro, Serbia, Palestinian territories and Portugal. For Brazil,

Table 2. Fit statistics for the six-factor model across included countries

	RCADS 47 items			RCADS 37 items		
	CFI	TLI	RMSEA	CFI	TLI	RMSEA
Brazil	0.90	0.88	0.07	0.88	0.89	0.07
Bulgaria	0.93	0.92	0.04	0.93	0.92	0.04
Croatia	0.90	0.90	0.05	0.90	0.90	0.05
Indonesia	0.91	0.90	0.04	0.92	0.91	0.04
Montenegro	0.92	0.91	0.04	0.93	0.92	0.04
Nigeria	0.88	0.87	0.04	0.88	0.87	0.04
Palestine	0.92	0.91	0.04	0.93	0.92	0.40
Philippines	0.41	0.37	0.17	-	-	-
Portugal	0.93	0.92	0.06	0.92	0.91	0.06
Romania	0.88	0.87	0.05	0.87	0.86	0.06
Serbia	0.93	0.92	0.05	0.93	0.92	0.05

RMSEA, root mean square error of approximation; TLI, Tucker-Lewis index; CFI, comparative fit index.

Table 3. Parameters estimate and its standard error of the items showing DIF when each country was considered as the reference group

Reference country	Focal countries	Social phobia		Major depression		Separation anxiety		Generalised anxiety		Obsessive-compulsive	
		Item 12 β (s.e.)	Item 32 β (s.e.)	Item 19 β (s.e.)	Item 47 β (s.e.)	Item 9 β (s.e.)	Item 46 β (s.e.)	Item 1 β (s.e.)	Item 13 β (s.e.)	Item 10 β (s.e.)	Item 42 β (s.e.)
Serbia	Indonesia	-	-	-0.703(0.07)	-	-	0.843(0.067)	-	-	-	-
	Croatia	-	-	-	-0.626(0.05)	0.853(0.059)	-	-	-	-	-
Nigeria	Romania	-	-	-	-	-	-	-	-	-	0.303(0.018)
	Serbia	-	-	-	-	-0.397(0.06)	-	-	-	-	-
	Indonesia	-	-	-0.661(0.068)	-	-	-	-0.645 (0.064)	-	-	-
	Croatia	-	-	-	-0.636 (0.052)	-	-	-	-	-	0.292(0.021)
Indonesia	Romania	-	-	-	-	-	-	-	-	-	-
	Brazil	-	-	-	-	-	-0.226(0.016)	-	-	-	-
	Serbia	-	-	-	-	-0.634 (0.054)	-	-	-	-	-
	Croatia	-	-	-	-	-	-	0.536(0.05)	-	-	-
Bulgaria	Montenegro	-	-	-	-	-0.256 (0.025)	-	-	-	-	-
	Romania	-	-	-	-	-	-	-	-	-	0.264(0.024)
	Indonesia	-	-	-0.647(0.046)	-	-	-	-	-	-	-
Croatia	Croatia	-	-	-	-0.568 (0.051)	-	-	-	-	-	-
	Romania	-	-	-	-	-	-	-	-	-	0.272(0.025)
	Nigeria	-0.968 (0.085)	-	-	-	-	-	-	-	-	-
Palestine-Gaza	Indonesia	-	-	-0.765(0.067)	-	0.787(0.06)	-	-0.798 (0.064)	-	-	-
	Romania	-	-	-	-	-	-	-	-	-	0.293(0.023)
	Nigeria	-0.869 (0.084)	-	-	-	-	-	-	-	-	-
	Indonesia	-	-	-0.691(0.062)	-	0.618(0.059)	-	-	-	-	-
Romaine	Croatia	-	-	-	-	-	-	0.744(0.053)	-	-	-
	Romaine	-	-	-	-	-	-	-	-	-	0.275(0.024)

Continued

Table 4. Distribution of the RCADS scores for included countries

Country	SP score	PD score	SAD score	GAD score	OCD score	Total Anxiety score	MDD score
Brazil	8.27 (3.62)	1.94 (3.68)	2.92 (2.38)	5.74 (3.28)	2.51 (3.20)	21.38 (14.06)	2.58 (3.87)
Bulgaria	10.26 (5.16)	5.22 (4.06)	2.22 (2.44)	5.83 (3.36)	5.28 (3.31)	28.82 (14.38)	6.90 (4.43)
Croatia	9.82 (5.40)	5.96 (4.70)	3.00 (2.86)	6.79 (3.68)	5.19 (3.46)	30.77 (16.73)	7.23 (4.98)
Indonesia	8.89 (4.12)	5.04 (3.58)	5.24 (3.27)	6.48 (3.51)	5.78 (3.03)	31.43 (13.90)	6.36 (3.73)
Montenegro	7.06 (4.63)	5.34 (4.88)	2.48 (3.08)	4.94 (3.64)	3.75 (3.47)	23.57 (15.97)	5.82 (5.09)
Nigeria	9.21 (4.47)	7.23 (4.56)	6.62 (3.67)	5.04 (3.31)	7.37 (3.82)	35.48 (15.36)	7.04 (4.44)
Palestine	9.74 (5.40)	6.67 (4.44)	6.56 (4.39)	6.92 (3.83)	6.33 (3.60)	36.23 (17.54)	7.20 (4.53)
Portugal	9.70 (5.75)	6.37 (5.81)	3.51 (3.73)	8.32 (4.55)	4.88 (3.89)	32.78 (20.57)	7.17 (5.95)
Romania	9.68 (4.97)	5.24 (3.78)	2.05 (2.13)	5.37 (2.91)	5.11 (3.18)	27.46 (13.77)	6.70 (3.95)
Serbia	9.22 (5.02)	5.63 (4.71)	2.53 (2.82)	6.37 (3.72)	4.11 (3.25)	27.86 (15.95)	6.65 (4.75)
F (df)*	22.26 (12)**	36.65 (12)*	97.76 (12)**	41.73 (12)**	42.52 (12)**	39.26 (12)**	33.46 (12)**

SP, social phobia; PD, panic disorder; SAD, separation anxiety disorder; GAD, generalised anxiety disorder; OCD, obsessive-compulsive disorder; MDD, major depressive disorder,

*Adjusted for age, gender and religion type; ** $p < 0.001$.

Nigeria and Romania the CFI and/or TLI values were also slightly below acceptable values (≤ 0.86), while the RMSEA values were all below 0.08, thus indicating an acceptable fit to the data.

Table 4 shows the RCADS scores across the countries. For all scores, there were statistically significant differences among the countries considering the controlling variables. In order to estimate whether the DIF items contributed to the differences, it was analysed would a change occur if a DIF item would have been removed. Based on Table 2, in seven (14.58%) out of 48 head-to-head comparisons removing the DIF items resulted with significant changes; the difference in the scores went from significant to non-significant or vice versa (details available on request). This pattern was observed for items 1, 10, 12, 42 and 47.

Discussion

This is the only study designed to evaluate the cross-cultural measurement invariance of the RCADS adolescent self-report across several countries with different socioeconomic, cultural and religious backgrounds.

Before discussing the main findings, it is worth mentioning that the present study confirmed the original six-factor model for the RCADS representing six anxiety and depression subscales (Chorpita *et al.* 2005; Ebesutani *et al.* 2010), for ten new language versions, namely Arabic, Bulgarian, Bahasa Indonesia, Croatian, Montenegrin, Portuguese for Portugal, Portuguese for Brazil, Romanian, Serbian and Yoruba. The original six-factor model of the Portuguese version for Brazil, the Yoruba version for Nigeria and the

Romanian version, had the CFI and/or TLI values slightly below acceptable ranges, with the RMSEA values being acceptable. The CFI and TLI also had slightly lower values previously reported for the Spanish version (Sandín *et al.* 2009), the French version (Bouvard *et al.* 2015) and the original English as well (de Ross *et al.* 2002). These lower values of the two indexes may be an effect of the sample itself or their sizes (Fan *et al.* 1999), rather than because of significant violations in the construct itself, but this finding deserves a replication. Nevertheless, the six-factor model was not confirmed in the present study for the original English version used in the Philippines. Besides that the items might not measure the intended underlying construct for these adolescents, a possible explanation is also that the English version might not capture anxiety and depressive symptoms as it would be if the original language was used. In the Philippines several languages are spoken besides the official English and Filipino languages.

Turning to the main findings, the MIMIC model analyses showed that in overall ten out of 47 RCADS items exhibited DIF considering all comparisons. These items belong to the separation anxiety disorder, social phobia, generalised anxiety disorder, obsessive-compulsive disorder and major depressive disorder subscale, two items per each subscale. None of the items from the panic disorder subscale exhibited DIF. However, only one or two items were flagged with DIF in the head-to-head comparisons, while there were three to five items flagged with DIF when one country was compared with the other nine. The items, which mostly exhibited DIF were item 9 ('I worry about being away from my parents'), item 19 ('I have no energy for things'), item 42 ('I have to

do some things over and over again') and item 47 ('I feel restless'). Using the original English version to test measurement invariance across African and White American children and adolescents, five non-invariant items were identified (Trent *et al.* 2012). These include RCADS items 6, 8, 23, 42 and 44. Except item 42 ('I have to do some things over and over again'), the others were not identified in the present study. Taken together the previous study and our findings, no items from the RCADS subscales cluster together following some specific pattern, such as content or symptoms, to claim that DIF items are more likely to occur in one than another anxiety subscale or even a depression subscale considering cultural context. In other words, no anxiety or depressive symptoms are more prone to cultural or language bias and there are possibly other factors contributing to DIF in the RCADS. Furthermore, the original six-factor model was not violated even if all DIF items were removed. After removing all ten DIF items and re-running the CFA on the remaining 37 items, the fit indexes changed only slightly for the six-factor model. This finding on its theoretical factor structure confirms that the RCADS is indeed represented by items in six subscales, five about different anxiety and one about depressive symptoms, even with a smaller number.

These results have two important research implications. At first place, when studying cross-cultural invariance, with more groups included across different countries and using different language versions, greater variability and more sensitive to cultural influences among the groups would be detected for items measuring anxiety and depressive symptoms. Second, in every cross-cultural/ethnic comparisons with the RCADS, its non-invariant items should be taken into account in order to demonstrate that differences in anxiety or depressive symptoms are less likely bias estimations due to non-invariant items in the composite scores (Trent *et al.* 2012). The final analysis in our study showed that DIF items, in particular RCADS item 1, 10, 12, 42 and 47, might have some effects on cross-group comparisons. Similar recommendations have been made based on finding of cross-cultural measurement non-invariance in other widely used measures of childhood psychopathology (Stevanovic *et al.* 2014).

The current study had an important strength. The study utilised a large sample size, using multiple language versions of the RCADS, employing a cross-country design rather than in-country and including adolescents from socioeconomically, culturally and religiously diverse nations. All these elements allowed us to cover greater variability among adolescents and countries, and to broaden the generalisability of

findings to a multicultural context. There were, however, some limitations that need to be taken into consideration when interpreting our findings. First, participants were sampled from regions of convenience. Using a sample of convenience could limit generalisability of the findings to adolescents from other regions in the same country. Second, the sample size varied across the countries and possibly not the same levels of the variability in expressions of anxiety and depression symptoms was captured, which could further limit the generalisability. In addition, it has been argued that a wide range in sample size could bias the results when conducting multi-group confirmatory factor analyses (MG-CFA; Brown, 2006, p. 279), which might be possible with our study due to a varied sample size across the included countries. Third, considering that we only collected self-reports, there might be more non-invariance if child- and/or parent-reports were used, which should be explored in future studies. Fourth, since the main disadvantage of MIMIC models is that they only test uniform DIF, but not non-uniform DIF, it might be more fruitful to use IRT methods as alternatives to test measurement invariance. Another alternative is MG-CFA technique, which assesses other types of invariance including metric and strong invariance, equal to testing non-uniform and uniform DIF. Nevertheless, it should be noted that neither IRT nor MG-CFA can control the effects of additional variables when assessing DIF. Finally, in the present study, the effect of clustered data was not taken into account through multilevel structural equation modelling, because at least 20–50 clusters are needed to ensure stable parameter estimates and convergence in a model fitting (Stegmueller, 2013). Since the present study included only ten countries, we could not utilise the advantage of multilevel structure equation modelling such as multilevel MIMIC or CFA.

In conclusion, the present study has established that with larger number and higher diversity of comparison across cultural-groups, the cross-cultural measurement invariance of the RCADS still holds firmly. There is clear evidence that relatively small number of the RCADS items is non-invariant, especially when considering two different cultural/ethnic groups, which indicates on sound cross-cultural validity. Even with all cross-culturally non-invariant items removed from nine language versions included, the original factor model representing six subscales was not violated. It is indicated that the RCADS self-report has fair suitability for cross-cultural research, but in every cross-cultural/ethnic comparisons, RCADS non-invariant items should be taken into account for the composite scores in order to demonstrate that differences in anxiety or depressive symptoms are less likely bias estimations due to non-invariant items.

Supplementary material

The supplementary material for this article can be found at <http://dx.doi.org/10.1017/S204579601600038X>.

Acknowledgements

We would like to thank to Bautista MAA from Dept of Psychology, St. Dominic College of Asia, City of Bacoor, Philippines, Cabral S and Barroso J, Brazil, Oladimeji O from Zankli Medical Center/Liverpool School of Tropical Medicine, Nigeria, Yatan Pal Singh Balhara from Department of Psychiatry, All India Institute of Medical Sciences, Ansari Nagar, New Delhi, India.

Financial support

This research received no specific grant from any funding agency, commercial or not-for-profit sectors.

Conflict of Interest

None.

Ethical Standard

The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

Availability of Data and Materials

The data and material supporting the findings could be obtained from Dejan Stevanovic, email stevanovic.dejan79@gmail.com.

References

- Angold A, Costello EJ, Messer SC, Pickles A (1995). Development of a short questionnaire for use in epidemiological studies of depression in children and adolescents. *International Journal of Methods in Psychiatric Research* **5**, 237–249.
- Atilola O, Balhara YPS, Stevanovic D, Avicenna M, Kandemir H (2013). Self-reported mental health problems among adolescents in developing countries: results from an international pilot sample. *Journal of Developmental Behavioral Pediatrics* **34**, 129–137.
- Bouvard M, Denis A, Roulin JL (2015). The French version of the revised child anxiety and depression scale (RCADS)

in a nonclinical sample. *Swiss Journal of Psychology* **74**, 119–127.

- Boyce W, Torsheim T, Currie C, Zambon A (2006). The family affluence scale as a measure of national wealth: validation of an adolescent self-report measure. *Social Indicators Research* **78**, 473–487.
- Brown RC, Yaroslavsky I, Quinoy AM, Friedman AD, Brookman RR, Southam-Gerow MA (2013). Factor structure of measures of anxiety and depression symptoms in African American youth. *Child Psychiatry Human Development* **44**, 525–536.
- Brown TA (2006). *Confirmatory Factor Analysis for Applied Research*. Guilford Press: New York, NY.
- Bye BV, Gallicchio SJ, Dykacz JM (1985). Multiple-indicator, multiple-cause models for a single latent variable with ordinal indicators. *Sociological Methods Research* **13**, 487–509.
- Byrne BM, Watkins D (2003). The issue of measurement invariance revisited. *Journal of Cross-Cultural Psychology* **34**, 155–175.
- Charman T, Pervova I (1996). Self-reported depressed mood in Russian and UK schoolchildren. A research note. *Journal of Child Psychology and Psychiatry* **37**, 879–883.
- Cheung GW, Rensvold RB (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling* **9**, 233–255.
- Chorpita BF, Yim L, Moffitt C, Umemoto LA, Francis SE (2000). Assessment of symptoms of DSM-IV anxiety and depression in children: a revised child anxiety and depression scale. *Behavior Research and Therapy* **38**, 835–855.
- Chorpita BF, Moffitt CE, Gray J (2005). Psychometric properties of the revised child anxiety and depression scale in a clinical sample. *Behaviour Research and Therapy* **43**, 309–322.
- de Ross RL, Gullone E, Chorpita BF (2002). The revised child anxiety and depression scale: a psychometric investigation with Australian youth. *Behaviour Change* **19**, 90–101.
- Dimitrov DM (2010). Testing for factorial invariance in the context of construct validation. *Measurement and Evaluation in Counseling and Development* **43**, 121–149.
- Ebesutani C, Bernstein A, Nakamura BJ, Chorpita BF, Weisz JR, Research Network on Youth Mental Health (2010). A psychometric analysis of the revised child anxiety and depression scale – parent version in a clinical sample. *Journal of Abnormal Child Psychology* **38**, 249–260.
- Ebesutani C, Chorpita B, Higa-McMillan C, Nakamura B, Regan J, Lynch R (2011). A psychometric analysis of the revised child anxiety and depression scales – parent version in a school sample. *Journal of Abnormal Child Psychology* **39**, 173–185.
- Esbjörn BH, Sømhøvd MJ, Turnstedt C, Reinholdt-Dunne ML (2012). Assessing the revised child anxiety and depression scale (RCADS) in a national sample of Danish youth aged 8–16 years. *PLoS ONE* **7**, e37339.
- Essau CA, Leung PW, Conradt J, Cheng H, Wong T (2008). Anxiety symptoms in Chinese and German adolescents: their relationship with early learning experiences, perfectionism, and learning motivation. *Depression and Anxiety* **25**, 801–810.

- Essau CA, Ishikawa SI, Sasagawa S, Sato H, Okajima I, Otsui K, Georgiou GA, O'Callaghan J, Michie F (2011). Anxiety symptoms among adolescents in Japan and England: their relationship with self-construals and social support. *Depression and Anxiety* **28**, 509–518.
- Fan X, Thompson B, Wang L (1999). Effects of sample size, estimation methods, and model specification on structural equation modeling fit indexes. *Structural Equation Modeling* **6**, 56–83.
- Gregorich SE (2006). Do self-report instruments allow meaningful comparisons across diverse population groups? Testing measurement invariance using the confirmatory factor analysis framework. *Medical Care* **44**, S78–S94.
- He J, van de Vijver F (2012). Bias and equivalence in cross-cultural research. *Online Readings in Psychology and Culture* **2**, 8.
- Joreskog KG, Goldberger AS (1975). Estimation of a model with multiple indicators and multiple causes of a single latent variable. *Journal of the American Statistical Association* **70**, 631–639.
- Kösters MP, Chinapaw MJ, Zwaanswijk M, van der Wal MF, Koot HM (2015). Structure, reliability, and validity of the revised child anxiety and depression scale (RCADS) in a multi-ethnic urban sample of Dutch children. *BMC Psychiatry* **15**, 1.
- Latzman RD, Naifeh JA, Watson D, Vaidya JG, Heiden LJ, Damon JD, Hight TL, Young J (2011). Racial differences in symptoms of anxiety and depression among three cohorts of students in the southern United States. *Psychiatry: Interpersonal Biological Processes* **74**, 332–348.
- March JS, Parker JDA, Sullivan K, Stallings P, Conners K (1997). The multi-dimensional anxiety scale for children (MASC): factor structure, reliability, and validity. *Journal of the American Academy of Child and Adolescent Psychiatry* **36**, 554–565.
- Mathyssek CM, Olino TM, Hartman CA, Ormel J, Verhulst FC, Van Oort FV (2013). Does the revised child anxiety and depression scale (RCADS) measure anxiety symptoms consistently across adolescence? The TRAILS study. *International Journal of Methods in Psychiatric Research* **22**, 27–35.
- Milfont TL, Fisher R (2010). Testing measurement invariance across groups: applications for cross-cultural research. *International Journal of Psychological Research* **3**, 111–121.
- Muthén LK, Muthén BO (1998–2010). *Mplus User's Guide*. 6th edn. Muthén & Muthén: Los Angeles, CA.
- Ollendick TH (1983). Reliability and validity of the revised fear survey schedule for children (FSSC-R). *Behaviour Research and Therapy* **21**, 395–399.
- Ollendick TH, Yang B, King NJ, Dong Q, Akande A (1996). Fears in American, Australian, Chinese, and Nigerian children and adolescents: a cross-cultural study. *Journal of Child Psychology and Psychiatry* **37**, 213–220.
- Park AL, Ebesutani CK, Bose D, Chorpita BF (2015). Psychometric properties of a Spanish translation of the revised child anxiety and depression scale—parent version. *Journal of Psychopathology and Behavioral Assessment* 1–13. doi: 10.1007/s10862-015-9517-7.
- Reynolds CR, Richmond BO (2000). *Revised Children's Manifest Anxiety Scale: Manual*. Western Psychological Services: Los Angeles, CA.
- Sandín B, Valiente RM, Chorot P (2009). RCADS: evaluación de los síntomas de los trastornos de ansiedad y depresión en niños y adolescentes. *Revista de Psicopatología y Psicología Clínica* **14**, 193–206.
- Stegmuller D (2013). How many countries for multilevel modeling? A comparison of frequentist and Bayesian approaches. *American Journal of Political Science* **57**, 748–761.
- Stevanovic D, Urbán R, Atilola O, Vostanis P, Singh Balhara YP, Avicenna M, Knez R, Franic T, Petrov P (2014). Does the strengths and difficulties questionnaire—self report yield invariant measurements across different nations? Data from the International Child Mental Health Study Group. *Epidemiology and Psychiatric Sciences* **24**, 323–334.
- Stevanovic D, Jafari P, Knez R, Franic T, Atilola O, Davidovic N, Bagheri Z, Lakic A (2016). Can we really use available scales for child and adolescent psychopathology across cultures? A systematic review of cross-cultural measurement invariance data. *Transcultural Psychiatry*, in press.
- Stewart SM, Kennard BD, Lee PW, Hughes CW, Mayes TL, Emslie GJ, Lewinsohn PM (2004). A cross-cultural investigation of cognitions and depressive symptoms in adolescents. *Journal of Abnormal Psychology* **113**, 248.
- Trent LR, Buchanan E, Ebesutani C, Ale CM, Heiden L, Hight TL, Damon JD, Young J (2012). A measurement invariance examination of the revised child anxiety and depression scale in a southern sample: differential item functioning between African American and Caucasian youth. *Assessment* **20**, 175–187.