Published in final edited form as:

Int J Soc Psychiatry. 2012 July ; 58(4): 386–392. doi:10.1177/0020764011400999.

Differential Item Functioning due to Gender between Depression and Anxiety Items among Chilean Adolescents

Cristina B. Bares, Ph.D., M.S.W, Fernando Andrade, M.A.S, Jorge Delva, Ph.D., M.S.W, Andrew Grogan-Kaylor, Ph.D., M.S.W, and Akihito Kamata, Ph.D.

Cristina B. Bares: cbb@umich.edu; Fernando Andrade: fandrade@umich.edu; Jorge Delva: jdelva@umich.edu; Andrew Grogan-Kaylor: agrogan@umich.edu; Akihito Kamata: kamata@uoregon.edu

Abstract

Although much is known about the higher prevalence of anxiety and depressive disorders among adolescent females, less is known about the differential item endorsement due to gender in items of commonly scales used to measure anxiety and depression. We conducted a study to examine if adolescent males and females from Chile differed on how they endorsed the items of the Youth Self-Report (YSR) anxious/depressed problem scale. We used data from a cross-sectional sample consisting of 925 participants (Mean age = 14, SD=1.3, 49% females) of low to lower-middle socioeconomic status. A two-parameter logistic (2PL) IRT DIF model was fit. Results revealed differential item endorsement (DIF) by gender for six of the 13 items with adolescent females being more likely to endorse a depression item while males were found more likely to endorse anxiety items. Findings suggest that items found in commonly utilized measures of anxiety and depression symptoms may not equally capture true levels of these behavioral problems among adolescent males and females. Given the high levels of mental disorders in Chile and surrounding countries, further attention should be focused on increasing the number of empirical studies examining potential gender differences in the assessment of mental health problems among Latin American populations to better aid our understanding of the phenomenology and determinants of these problems in the region.

Keywords

anxiety; depression; adolescent females; international

A concerning aspect of depression and anxiety disorders is the greater prevalence of these disorders among adult and adolescent females (Halbreich & Kahn, 2007). Across a number of different contexts, rates of anxiety disorders have been found to be greater for adolescent females than for males (Campbell & Rapee, 1994; Gullone & King, 1993). Similarly, greater rates of depression or depressive disorders have been found among adolescent females than males (Garcia, Skay, Sieving, Naughton, & Bearinger, 2008; Robles-Pina, Defrance, & Cox, 2008; Wade, Cairney, & Pevalin, 2002). Despite these differential rates of mental health problems, understanding which items adolescent females are more likely to endorse compared to males on commonly used measures of anxiety and depression constructs remains limited, particularly among Latin American populations. We posit that better understanding which items are more likely to be supported by adolescent females, if any, may serve to better inform interventions to specifically target the concept suggested by the items.

We suspect that, as Breslau and colleagues have outlined elsewhere, there might be two pathways to explain differences in item endorsement that result from characteristics of the individual (Breslau, Javaras, Blacker, Murphy, & Normand, 2008). One pathway suggests that adolescent females may simply be more likely than adolescent males to endorse all items on common measures of depression and anxiety because adolescent females are more willing to disclose personal information. Thus, similar rates of gender differences for all of the items would lead to a higher score but would not reflect a true higher prevalence of an underlying mental health problem. Another possible pathway might be that higher rates of mental health problems might be observed among adolescent females due to the different ways by which males and females interpret the items which are designed to measure these constructs. As an example, a common item measuring depression is whether one feels hopeless. How adolescent females interpret hopelessness might differ from the interpretations that adolescent males use to answer that item. In such a case, some survey items would be endorsed differently between adolescent females and males, but there would be no differences in response to other items that were interpreted consistently by both genders. In this case, differences in endorsing the items would be consistent with differences between the two genders in the relationship between true levels of depression and anxiety and individual items, but not with differences between the two genders in level of depression and anxiety.

To understand the source of possible gender differences in item endorsement it is necessary to examine whether responses to individual anxiety and depression items are linked systematically to personal characteristics such as gender, age, or ethnicity. Statistically, finding differences in item endorsement is typically referred to as differential item functioning (DIF) (Holland, Wainer, & Educational Testing, 1993). DIF is an approach that allows for this type of analysis while controlling for the trait level being measured.

Early studies in the area of adolescent mental health using the Beck Depression Inventory found that adolescent males and females did not differ on their overall depression score, rather, studies found that adolescent males and females differed in their support of individual depression items. Adolescent males were found to report more symptoms of irritability, work inhibition, social withdrawal, and sleep disturbance (Baron & Joly, 1988). Similarly, adolescent males were also found to report greater levels of anhedonia, depressed morning mood, and morning fatigue (Bennett, Ambrosini, Kudes, Metz, & Rabinovich, 2005). Adolescent females, on the other hand, expressed depression by reporting symptoms of body image distortion, loss of appetite, weight loss, mood and lack of satisfaction (Baron & Campbell, 1993; Baron & Joly, 1988).

Even though most mental health scales are constructed to contain items that do not exhibit DIF, there are still some studies which have found DIF in mental health items. Existing research with adolescents has focused on examining the Conduct Disorder (CD) criteria contained in the Diagnostic Interview Schedule for Children for evidence of DIF. Gelhorn and colleagues found that most of the items used as evidence of CD criteria did not exhibit gender based differential item functioning, except for three: *destruction of property* which was endorsed less by males, *steals without confrontation* which was endorsed less by females, and *runs away* again endorsed less by female adolescents (Gelhorn, et al., 2009).

Applying DIF to the diagnostic criteria of Major Depressive Disorder, as operationalized by the DSM-IV (American Psychiatric Association, 2000), Small and colleagues found evidence of DIF when comparing adolescents classified into different types of comorbid depression groups (Small, et al., 2008). Adolescents with comorbid internalizing disorders were found to be less likely to endorse the item measuring psychomotor agitation/retardation than adolescents belonging to the depression or externalizing groups. Adolescents belonging

to the depression group were found to be less likely to endorse the item measuring cognitive impairments (Small, et al., 2008).

Despite the existence of findings pointing to differential item endorsement between males and females, potential gender differences in the endorsement of items of anxiety and depression scale remain under-explored, particularly in Latin America. In Chile, anxiety and depression are one of the most frequent disorders (WHO, 2004; Vicente et al., 2007) and the rate of lifetime depressive and anxiety disorders among the female population is twice to three times greater than that of males (Vicente et al., 2007; 2006). It is thus important to explore whether gender influences how individuals endorse items measuring depression and anxiety symptoms. In the present study, we examined if there were gender differences in the endorsement of anxiety and depression items from the anxious-depression scale found in the Youth Self Report (Achenbach & Rescorla, 2001) with a community based sample of adolescents from Santiago, Chile. Besides increasing the international normative and descriptive data base on adolescent anxiety and depressive symptoms, an examination of the contribution of gender to item endorsement can further our understanding of the phenomenology and determinants of anxiety and depression in a country with one of the highest rates of these disorders.

Method

Sample and Procedures

This study used data from the first part of a two-part assessment, scheduled two years apart, of adolescents participating in the Santiago Longitudinal Study (SLS), a study of community-dwelling adolescents in Santiago, Chile. This is a collaborative project between U.S. and Chilean institutions with funding from the U.S. National Institute on Drug Abuse (NIDA). The sample for the study consisted of 925 participants (Mean age = 14.2, age range = 11.9 – 17.8, SD=1.3, 49.4% female) of mid-to-low socioeconomic status. The average number of years of parental education was 9.9 years and average monthly family income was 315,000 pesos (500 Chilean pesos is approximately \$1.00 US dollar), respectively. About 68% of the youth lived with both their father and mother. Participants for this study were recruited from a sample of approximately 1,700 families that participated in a study of nutrition when the youth were in infancy and 1,200 when the youth were 10 years old (Lozoff, et al., 2003). Between 2008 and 2010, 1,025 youth were successfully recruited. In 2008–2009, youth completed Wave 1 assessments consisting of a 2-hr interviewer-administered questionnaire with comprehensive questions on demographics, mental health, peer, family, and school characteristics. Wave 2 assessments have begun and will be completed in 2011.

Each adolescent participant completed a two-hour interviewer-administered questionnaire with standardized measures that were pilot tested and validated with the population under investigation prior to conducting the current study. Some of the topics assessed ranged from the quality of the relationship with parents, adolescents' self perceptions, behavior, health status, and substance use. All of the interviews were conducted in Spanish in a private office at the Chilean site by Chilean psychologists trained in the administration of standardized instruments. Adolescent assent and parental consent were obtained by the interviewers prior to commencing the interviews. The study received Institutional Review Board approval from the institutional review boards of the corresponding universities.

Measures

Assessment of depression and anxiety—We used the ASEBA School-Age Form Youth Self Report (YSR) to measure levels of depression and anxiety and other disorders

(Achenbach & Rescorla, 2001). The stem question for the thirteen items that make up the depression-anxiety scale is “Below is a list of items that describe kids. For each item that describes you now or within the past 6 months, please tell me if the item is (3) *very true or often true*, (2) *somewhat or sometimes true*, or (1) *is not true* (0).

Analytic Method

Analyses proceeded in several steps. All analyses were conducted using Mplus (Muthén & Muthén, 2009). First, an exploratory factor analysis was conducted to obtain an initial idea of the factor structure of the items from the depression-anxiety scale. Once we had an idea of the basic factor structure, we confirmed that structure using confirmatory factor analysis in order to decide upon the optimal number of factors for the model. Note that both of the factor analytic procedures treated item response data as categorical variables.

Once the basic factor structure had been established, individual items were examined for evidence of DIF. Item Response Theory (IRT) modeling was used to estimate DIF by gender on depression and anxiety items from the Youth Self Report. IRT models are utilized to examine the observed indicators of a scale measuring a latent construct. These types of models center on identifying whether participant responses to individual scale items are linked systematically to participants’ characteristics such as age, gender, race, or ethnicity (Osterlind & Everson, 2009). IRT modeling provides information about the probability that an adolescent with a particular level on the measured trait would endorse a particular item. To identify what depression or anxiety items had DIF we fit a polytomous two-parameter logistic (2PL) IRT model, which is equivalent to graded response model, with the assumption that the mean DIF is 0. Note that DIF parameters are not identified by any DIF analysis, and they need some scaling reference point for them to be identified. The zero-mean DIF assumption is a commonly utilized approach that scales the magnitudes of DIF to the mean DIF (Osterlind & Everson, 2009). There was no previous information in the literature about which items were free of differential item functioning, and there was no evidence that many items displayed DIF in the same direction. Therefore, we judged that the zero-mean DIF assumption would be a reasonable approach.

Once items showing evidence of DIF had been identified from the IRT analysis, we conducted follow-up analyses to locate the level of the responses (thresholds) that demonstrated differential item functioning. Specifically, we examined each item to determine whether the DIF occurred at the threshold between responses 0 (“not true”) and 1 (“somewhat true”) or between responses 1 (“somewhat true”) and 2 (“often”). Identification of the exact threshold at which

DIF occurred was carried out by fitting three Rasch Multigroup (MG) models. The first Rasch model allowed only the first threshold (responses 0 to 1) to vary, while the second Rasch model allowed the second threshold (responses 1 to 2) to vary. The third Rasch model let both thresholds vary. We used the following criteria to choose between models: BIC (Bayes Information Criterion), ABIC (adjusted Bayesian Information Criterion), and AIC (Akaike’s Information Criterion).

Results

Preliminary Results

The initial exploratory factor analyses resulted in two factors with eigenvalues greater than one. Based on these results each scale item was classified into the two factors using two criteria. The first criteria involved examining the factor loadings for each item. A cut-off point of 0.41 was used such that only items with factors loadings greater than 0.41 were classified into one of the two factors. The second criterion involved examining the content

of the items to determine what construct each item was measuring and thus assigning a label to each factor. Based on these two criteria, twelve of the thirteen items loaded in one of the two factors and one item loaded in the two factors. The way the items loaded in two factors suggested the existence of one latent trait which could be termed depression and another which could be characterized as anxiety.

To confirm the results of the exploratory factor analyses, two Confirmatory Factor Analyses were fitted and compared. The first model was a confirmatory factor analysis assuming one underlying structure (depression and anxiety, together). The second confirmatory factor analysis that was tested assumed two separate factors (one for depression and the other for anxiety). Table 1 presents the results for these analyses along with factor loadings and fit indices. The best fitting model for our data was the two-factor structure model, according to the fit indices (TLI, CFI, and RMSEA) and the Robust Chi Square Difference Testing with Mean and Variance Adjusted Test Statistics (41.635, $p < 0.001$) (Asparouhov & Muthén, 2006).

Matching criteria and focal group—The unobserved latent factors of depression and anxiety were used as matching criterion for the analyses to represent the trait levels of depression and anxiety. The focal and reference groups were determined using adolescent gender, where female adolescents were the focal group and male adolescents the reference group.

Differential Item Functioning Results

Using the two-factor anxious and depression structure, a polytomous two-parameter IRT model was fit to examine for Differential Item Functioning (DIF) with the assumption that the mean DIF is 0. The results indicated that six items of the 13 items had DIF. Figure 2 depicts the results of the analysis. Specifically, there was evidence for differential item functioning in two items in the depression factor: *I cry a lot* and *I feel that I have to be perfect*. In addition, evidence of DIF was found in four items comprising the anxiety factor: *I'm afraid that I might think or do something bad*, *I am too fearful or anxious*, *I feel too guilty*, and *I am self-conscious or easily embarrassed*.

The analysis showed that for the depression factor adolescent females with the same level of depression as adolescent males had a greater probability of endorsing the item *I cry a lot*. For the other depression item, adolescent males with the same level of depression as females had a greater probability of endorsing the item *I feel that I have to be perfect*. In other words, female adolescents were more likely to report that they cry a lot while male adolescents were more likely to report that they have to be perfect.

For the anxiety factor adolescent males were more likely to endorse the item *I feel too guilty*, *I am afraid I might think or do something bad*, *I am too fearful or anxious*, and *I am self-conscious or easily embarrassed*. Thus, for the same level of anxiety, boys in our sample were more likely to report feeling too guilty, being afraid of thinking or doing something bad, feeling too anxious, and being easily embarrassed.

For the six items that showed evidence of DIF we calculated the magnitude of DIF and classified it using ETS criteria (Dorans & Holland, 1993). Based on the DIF point estimate and its confidence interval, only one item showed DIF that was classified as a large magnitude difference. This was the item *I cry a lot* while the other five items were classified as having negligible DIF.

Results of DIF by Threshold Analysis

Table 2 shows the results for the analyses of the items that had DIF by threshold. Each item had two thresholds: threshold A, going from response category 0 to response category 1 and threshold B, going from response category 1 to response category 2. The results showed that the item *I cry a lot* had significant and large DIF magnitude in both thresholds using ETS criteria. In other words, female adolescents compared to males with the same level of depression were not only more likely to endorse the item *I cry a lot* but also they were more likely to endorse the transition in responses from 0 to 1 and from 1 to 2. Even though there were other significant DIF by threshold results, they all grouped into ETS category that indicates negligible DIF.

Discussion

This study makes several contributions to the literature. There is a longstanding conversation in the literature on child and adolescent mental health on whether anxiety and depression are more appropriately conceptualized as two co-occurring disorders (Bares, Andrade, Delva, Grogan-Kaylor, in-press), or as one single disorder (Chavira, Stein, & Bailey, 2004; Muris, 2006; Oneil, Podell, Benjamin, & Kendall, 2010). The suggestion of this research is that with this adolescent population anxiety and depression are best thought of as related but separate constructs. If anxiety and depression are correlated but distinct, they may have different etiologies and further research is needed to ascertain if there are unique differentiable predictive factors for anxiety and depression. Similarly, correlated but distinct constructs might have different sequelae, and further research on the distinct consequences of anxiety and depression is worth undertaking.

With regard to our core analysis of DIF, our analyses revealed that 6 of the 13 items of the Depression and Anxiety scale of the Youth Self Report showed differential item endorsement by gender but according to ETS criteria only one item evidenced a large magnitude of DIF. Specifically, for the same level of depression, adolescent females were more likely than adolescent males in our sample to endorse the item, *I cry a lot*. This finding is consistent with previous research showing that non-depressed women report crying more frequently and more intensely than non-depressed men (Lombardo, Cretser, & Roesch, 2001) and it may be possible that this item is capturing the greater tendency that females have to cry than to actually be measuring an aspect of depression. Although crying is seen as a feature of depression there is little evidence for this linkage. In fact, Rottenberg and colleagues have indicated that there exist no differences between the likelihood of crying between depressed and non-depressed adults (Rottenberg, Gross, Wilhelm, Najmi, & Gotlib, 2002). Although not considered very large, the other gender differences in item endorsement suggest that the items showed differential item functioning might be capturing differences in how Chilean youth express and manifest depression and anxiety. Further cross-cultural work in other Latin American countries is needed to examine the generalizability of this finding.

Limitations

Although the study has several strengths, the findings should be viewed in the context to the following three limitations. First, the study relied on a convenience sample of adolescents of lower socioeconomic status. It is not clear if the findings can be generalized to the larger population of Chilean adolescents. Second, although T-scores can be available to indicate the possibility of a youth being a 'case' of depression-anxiety, the instrument is not one intended to classify youth as meeting diagnostic criteria for depression and/or anxiety. Therefore, the extent to which there is DIF in relation to a diagnosis of mental disorder remains to be investigated. In a similar vein, it is not clear the effect on a diagnostic measure should items with DIF be eliminated, if DIF is found. Third, we did not examine if DIF was

present as a function of other characteristics such as age, ethnicity, and family socioeconomic status, among other possible identities, or intersections of these. Further understanding the behavior of depression and anxiety items by these and other identities could further shed light into the construction of these scales.

Notwithstanding these limitations, to our knowledge this is the first study of DIF with a Latin American sample. As such, these findings contribute to the research that seeks to better measure depression and anxiety among youth in diverse cultural settings. These findings may have significant implications with regard to practice in child mental health. For instance, practice guidelines already suggest different treatment approaches for anxiety and depression. However, as we continue to learn more about the separate etiologies and sequelae of anxiety and depression among adolescents, though among adults as well, we may begin to be able to develop more effective, tailored, treatments for these disorders. Furthermore, practitioners may also benefit from an awareness of the differential functioning of these items. For example, practitioners should be aware that adolescent males may be less likely to say that they cry a lot while adolescent females may be less likely to say that they feel too guilty. As an implication, adolescent males who indicate that they cry a lot may actually be experiencing higher levels of depression than standard instruments might indicate. Thus, at times, awareness of DIF, may provide a grounding for awareness of hitherto unnoticed problems.

Acknowledgments

We are extremely grateful to the adolescents and their families for taking the time to participate in this study. This research was funded by NIDA Grant # R01 DA021181 and the Vivan A and James L. Curtis School of Social Work Research and Training Center at the University of Michigan.

References

- Achenbach, TM.; Rescorla, L. Manual for the ASEBA School-Age Forms and Profiles. Burlington: University of Vermont, Research Center for Children, Youth, and Families; 2001.
- Alarcón RD. Mental health and mental health care in Latin America. *World Psychiatry*. 2003; 2(1):54. [PubMed: 16946892]
- American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4. Washington, DC: APA; 2000. TR
- Andrade L, Caraveo-Anduaga J, Berglund P, Bijl R, De Graaf R, Vollebergh W. The epidemiology of major depressive episodes: results from the International Consortium of Psychiatric Epidemiology (ICPE) Surveys. *International journal of methods in psychiatric research*. 2003; 12(1):3–21. [PubMed: 12830306]
- Asparouhov, T.; Muthén, BO. Robust Chi-Square Difference Testing with Mean and Variance Adjusted Test Statistics. *Mplus Web Notes*: No. 10. 2006. Retrieved from <http://www.statmodel.com/download/webnotes/webnote10.pdf>
- Bares, Andrade; Delva, Grogan-Kaylor. Examining the factor structure of anxiety and depression symptom items among adolescents in Santiago, Chile. *Journal of the Society for Social Work and Research*. (in-press).
- Baron P, Campbell T. Gender differences in the expression of depressive symptoms in middle adolescents: An extension of earlier findings. *Adolescence*. 1993; 28(112):903–911. [PubMed: 8266843]
- Baron P, Joly E. Sex differences in the expression of depression in adolescents. *Sex Roles*. 1988; 18(1):1–7.
- Bennett D, Ambrosini P, Kudes D, Metz C, Rabinovich H. Gender differences in adolescent depression: Do symptoms differ for boys and girls? *Journal of Affective Disorders*. 2005; 89(1–3): 35–44. [PubMed: 16219362]

- Breslau J, Javaras KN, Blacker D, Murphy JM, Normand SLT. Differential item functioning between ethnic groups in the epidemiological assessment of depression. *Journal of nervous and mental disease*. 2008; 196(4):297–306. [PubMed: 18414124]
- Campbell M, Rapee R. The nature of feared outcome representations in children. *Journal of Abnormal Child Psychology*. 1994; 22(1):99–111. [PubMed: 8163777]
- Chavira D, Stein M, Bailey K. Comorbidity of generalized social anxiety disorder and depression in a pediatric primary care sample. *Journal of Affective Disorders*. 2004; 80(2–3):163–171. [PubMed: 15207929]
- Costello EJ, Mustillo S, Erkanli A, Keeler G, Angold A. Prevalence and development of psychiatric disorders in childhood and adolescence. *Archives of general psychiatry*. 2003; 60(8):837–844. [PubMed: 12912767]
- Dorans, NJ.; Holland, PW. DIF detection and description: Mantel-Haenszel and standardization. In: Holland, PW.; Wainer, H., editors. *Differential Item Functioning*. Hillsdale, NJ: Lawrence Erlbaum; 1993.
- Garcia C, Skay C, Sieving R, Naughton S, Bearinger L. Family and racial factors associated with suicide and emotional distress among Latino students. *The Journal of School Health*. 2008; 78(9): 487–495. [PubMed: 18786041]
- Gelhorn H, Hartman C, Sakai J, Mikulich-Gilbertson S, Stallings M, Young S, et al. An item response theory analysis of DSM-IV conduct disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*. 2009; 48(1):42–50. [PubMed: 19034046]
- Gullone E, King N. The fears of youth in the 1990s: Contemporary normative data. *The Journal of Genetic Psychology*. 1993; 154(2):137–153. [PubMed: 8366328]
- Halbreich U, Kahn L. Atypical depression, somatic depression and anxious depression in women: Are they gender-preferred phenotypes? *Journal of Affective Disorders*. 2007; 102(1):245–258. [PubMed: 17092565]
- Hankin BL. Cognitive vulnerability-stress model of depression during adolescence: Investigating depressive symptom specificity in a multi-wave prospective study. *Journal of Abnormal Child Psychology*. 2008; 36(7):999–1014. [PubMed: 18437551]
- Holland, PW.; Wainer, H. *Educational Testing, S. Differential item functioning*. Hillsdale: Lawrence Erlbaum Associates; 1993.
- Lombardo WK, Cretser GA, Roesch SC. For crying out loud--The differences persist into the '90s. *Sex Roles*. 2001; 45(7–8):529–547.
- Lozoff B, De Andraca I, Castillo M, Smith JB, Walter T, Pino P. Behavioral and developmental effects of preventing iron-deficiency anemia in healthy full-term infants. *Pediatrics*. 2003; 112(4): 846(849). [PubMed: 14523176]
- Muris P. The pathogenesis of childhood anxiety disorders: Considerations from a developmental psychopathology perspective. *International Journal of Behavioral Development*. 2006; 30(1):5–11.
- Muthén, BO.; Muthén, L. *Mplus Version 5.21*. 2009.
- Oneil KA, Podell JL, Benjamin CL, Kendall PC. Comorbid depressive disorders in anxiety-disordered youth: Demographic, clinical, and family characteristics. *Child psychiatry and human development*. 2010; 41(3):330–341. [PubMed: 20066489]
- Osterlind, SJ.; Everson, HT. *Differential item functioning*. Los Angeles: Sage; 2009.
- Perry, G.; Olarreaga, M. Trade liberalization, inequality, and poverty reduction in Latin America. In: Bourguignon, F.; Pleskovic, B., editors. *Annual World Bank Conference on Development Economics-Regional 2007: Beyond Transition*. Washington: World Bank Inst; 2007. p. 103-151.
- Robles-Pina RA, Defrance E, Cox DL. Self-concept, early childhood depression and school retention as predictors of adolescent depression in urban Hispanic adolescents. *School psychology international*. 2008; 29(4):426–441.
- Rottenberg J, Gross JJ, Wilhelm FH, Najmi S, Gotlib IH. Crying threshold and intensity in major depressive disorder. *Journal of Abnormal Psychology*. 2002; 111(2):302–312. [PubMed: 12003451]
- Small D, Simons A, Yovanoff P, Silva S, Lewis C, Murakami J. Depressed adolescents and comorbid psychiatric disorders: Are there differences in the presentation of depression? *Journal of Abnormal Child Psychology*. 2008; 36(7):1015–1028. [PubMed: 18509755]

- Titelman D, Caldentey EP, Pineda R. The bigness of smallness: The financial crisis, its contagion mechanisms and its effects in Latin America. [Article]. *Cepal Review*. 2009; (98):7–33.
- Vicente B, Kohn R, Rioseco P, Saldivia S, Navarrete G, Torres P. Regional differences in psychiatric disorders in Chile. *Social Psychiatry and Psychiatric Epidemiology*. 2006; 41(12):935–942. [PubMed: 17036264]
- Vicente B, Kohn R, Saldivia S, Rioseco P. Burden of psychiatric diseases in Chile. *Revista médica de Chile*. 2007; 135(12):1591–1599.
- Wade TJ, Cairney J, Pevalin DJ. Emergence of gender differences in depression during adolescence: National panel results from three countries. *Journal of the American Academy of Child & Adolescent Psychiatry*. 2002; 41(2):190–198. [PubMed: 11837409]
- WHO. The global burden of disease: 2004 update. World Health Organization; 2004.

Table 1

Factor Loadings for Exploratory and Confirmatory Factor Analysis of YSR Depression and Anxiety Items (n=925)

Item	Exploratory Factor Analysis*		Confirmatory Factor Analysis*	
	Factor 1	Factor 2	Factor 1	Factor 2
Item 14. I cry a lot.	0.368	0.326	0.614	0.648
Item 29. I am afraid of certain places, or situations.	-0.123	0.609	0.415	0.443
Item 30. I am afraid of going to school	0.407	0.197	0.538	0.558
Item 31. I am afraid I might think or do something bad.	0.183	0.342	0.462	0.489
Item 32. I feel that I have to be perfect.	0.282	0.177	0.408	0.428
Item 33. I feel that no one loves me.	0.854	0.000	0.776	0.820
Item 35. I feel worthless or inferior.	0.766	0.009	0.721	0.752
Item 45. I am nervous or tense.	0.247	0.354	0.528	0.561
Item 50. I am too fearful or anxious.	0.001	0.704	0.596	0.637
Item 52. I feel too guilty.	0.333	0.373	0.622	0.129
Item 71. I am self-conscious or easily embarrassed.	0.116	0.355	0.413	0.436
Item 91. I think about killing myself.	0.770	-0.094	0.629	0.663
Item 112. I worry a lot	0.223	0.289	0.452	0.475
<hr/>				
	CFI=0.988		CFI=0.933	
	TLI=0.982		TLI=0.952	
	RMSEA = 0.030		RMSEA=0.049	
			CFI=0.960	
			TLI=0.971	
			RMSEA=0.039	

*The Robust Chi Square Difference Testing with Mean and Variance Adjusted Test Statistics was 41.635 ($p < 0.001$)

Table 2

YSR Depression and Anxiety Items with evidence of Differential Item Functioning: Analysis by Threshold

Item	2PL IRT Model			
	Estimate	CI	DIF ETS criteria	
			Point estimate	CI
Item 14, threshold a. I cry a lot.	1.368	0.823; 1.913	C	C
Item 14, threshold b. I cry a lot.	1.428	0.545; 2.311	C	B
Item 31, threshold a. I am afraid I might think or do something bad. *	-0.360	-0.854; 0.134	A	A
Item 50, threshold a. I am too fearful or anxious.	0.676	0.263; 1.089	C	A
Item 50, threshold b. I am too fearful or anxious.	0.725	-0.033; 1.483	C	A
Item 52, threshold a. I feel too guilty. *	-0.625	-1.250; 0.0003	C	A
Item 32, threshold a. I feel that I have to be perfect. *	-0.654	-1.119; -0.189	C	A
Item 71, threshold a. I am self-conscious or easily embarrassed. *	-0.426	-0.779; -0.073	A	A

ETS Criteria: A DIF is negligible; B is moderate DIF and C is large DIF; ~ $p \leq 0.1$, * $P \leq 0.05$, ** $p \leq 0.01$, *** $p \leq 0.001$.

* For threshold b there was no evidence of DIF