

Methodology

Spanish Validation of the Short Mood and Feelings Questionnaire (SMFQ) in Children Aged 8-12

José P. Espada¹, María T. González^{1,2}, Iván Fernández-Martínez¹, Mireia Orgilés¹, and Alexandra Morales¹

¹ Miguel Hernández University.

² Valencian International University.

ARTICLE INFO

Received: January 31, 2022

Accepted: May 04, 2022

Keywords:

Short Mood and Feelings
Questionnaire
Depression
Psychometric properties
Assessment
Children

ABSTRACT

Background: The Short Mood and Feelings Questionnaire (SMFQ) is one of the few well-established available measures designed to assess childhood depression. The objective of the present study was to assess the psychometric properties and explore the differential item functioning (DIF) of the SMFQ in a community sample of Spanish children. Gender and age differences in SMFQ scores were also analyzed. **Method:** The sample included 824 students aged 8-12 years ($M_{age} = 9.64$, $SD = 1.2$) recruited from four public and private schools in urban areas in the southeast of Spain. **Results:** Findings from factor analysis and Polytomous Rasch analysis supported a unidimensional interpretation of the SMFQ, thereby replicating findings across cultures and languages. The SMFQ had good reliability, and test-retest analysis indicated fair to good temporal stability. Evidence of construct validity was provided by a path diagram of the SMFQ and SDQ subscales. No age or gender differences in the SMFQ scores were observed. However, two items (3 and 11) exhibited gender-related DIF. **Conclusions:** The use of SMFQ sum-scores as a continuous measure of the severity of depressive symptoms can be supported. The measure shows promise as brief, reliable, valid instrument for the assessment of depressive symptoms in Spanish children.

Validación Española del Short Mood and Feelings Questionnaire (SMFQ) en Niños de 8 a 12 Años

RESUMEN

Antecedentes: El Cuestionario Breve de Estado de Ánimo y Sentimientos (SMFQ, por sus siglas en inglés) es una de las pocas medidas bien establecidas para evaluar la depresión infantil. Este estudio evalúa las propiedades psicométricas y explora el funcionamiento diferencial de los ítems (DIF) del SMFQ en una muestra comunitaria de niños españoles. **Método:** La muestra incluyó 824 escolares de 8 a 12 años ($Medad = 9.64$, $DT = 1.2$) de cuatro colegios públicos y privados del sureste de España. **Resultados:** Los hallazgos del análisis factorial y el análisis de Rasch politómico respaldaron una interpretación unidimensional del SMFQ. El SMFQ tuvo una buena consistencia interna y el análisis test-retest indicó una estabilidad temporal de regular a buena. Se examinaron las evidencias sobre la validez de constructo mediante un diagrama de ruta de las subescalas SMFQ y SDQ, sin diferencias en el SMFQ por edad y sexo. Sin embargo, dos ítems presentaron DIF relacionada con el género. **Conclusiones:** Se respalda el uso de las puntuaciones sumatorias del SMFQ como una medida continua de la gravedad de los síntomas depresivos. La medida se muestra prometedora como instrumento breve, fiable y válido para la evaluación de los síntomas depresivos en niños españoles.

Palabras clave:

Short Mood and Feelings
Questionnaire
Depresión
Propiedades psicométricas
Evaluación
Niños

Mental health problems during childhood are very common (Sucupira et al., 2017). Depression is considered by the World Health Organization (WHO, 2017) as the leading cause of ill health and disability worldwide. Symptomatology of depression often occurs for the first time in childhood or adolescence (Hankin, 2015). In Spain, Canals-Sans et al. (2018) found prevalence rates of 11.6% of depressive symptoms in school children from 11 to 12 years old, with no significant differences between girls (11.8%) and boys (11.3%), and no gender or age differences in depressive symptoms. As for differences in depression by gender and age, the literature points to differences in the prevalence of depression starting to become more apparent in adolescence at about 13 years of age, being more common in girls than in boys. However, it appears that the results are contradictory and inconclusive among studies based on questionnaire scores measuring depression in children and adolescents (Angold et al., 2002). Evidence suggests that children who have internalizing problems that refer to difficulties of a personal nature, such as depression, anxiety, withdrawal and somatic complaints (Santa Cruz et al., 2021), remain at risk of developing psychological disorders in later life (Clark et al., 2007). Furthermore, the early onset of depression may increase the probability of manifesting depressive episodes in adolescence and adulthood (Avenevoli et al., 2015), so accurate identification and early detection are important to ensure appropriate services and for prevention of later disorders (Canals-Sans et al., 2018).

The development and finding evidence of validity is a continuous and ongoing process (Turner et al., 2014). The use of instruments that have shown evidence of validity can help to rapidly screen and identify depression among children. Also, self-reported questionnaires are efficient and less expensive (Stevanovic, 2012). Several well-validated self-report instruments have been developed to assess depression among children (e.g., Beck, 1961; Kovacs, 1992). However, many measures do not offer open access, which restricts their use (Jeffreys et al., 2016). The Mood and Feelings Questionnaire (MFQ; Angold et al., 1987) does not present this limitation. It was developed for the early identification of depressive symptomatology in children and adolescents. Also, this instrument is recommended by the National Institute for Health and Clinical Excellence (NICE) as an assessment tool for the age range 8-18 years to screen and monitor depression treatment response (Lawton et al., 2016), and is one of the most useful tests for screening childhood depression (Bernaras et al., 2019; Lawton et al., 2016).

A short version of the MFQ (SMFQ) was developed by Angold et al. (1995) as an alternative to the MFQ, as it requires half the time to complete and allows the quick evaluation of depressive symptomatology in youth. Empirical evidence suggests that depression can be understood as a unidimensional phenomenon (Aggen et al., 2005). In this regard, the strong unidimensional structure and high internal consistency of SMFQ have been documented in previous studies, which have widely supported the SMFQ as an instrument based on a one-factor model (e.g., Angold et al., 2002; Costello et al., 1991; Kuo et al., 2005; Lundervold et al., 2013; Messer et al., 1995; Sharp et al., 2006; Turner et al., 2014). The SMFQ items assess a broad set of symptoms that characterize childhood depression according to *The Diagnostic and Statistical Manual of Mental Disorders 5th edition* (DSM-5; American Psychiatric Association [APA], 2013),

such as depressed mood and anhedonia, tiredness, restlessness, concentration difficulties, and several aspects of negative self-evaluation. According to McKenzie et al. (2011), it is an excellent instrument to identify clinical levels of depression in children and adolescents at risk for the development of major depression in adulthood.

The significance of cross-cultural finding evidence of validity and adaptation of the content for psychometric instruments has been pointed out by previous researchers (Adewuya et al., 2006). In recent years, the original English version of the self-report version of the SMFQ has been validated and translated in Serbian (Stevanovic, 2012), Norwegian (Olsen, 2015), Bangladeshi (Deeba et al., 2015), Brazilian-Portuguese (Pinto, 2014), Chinese (Yu et al., 2017), New Zealander (Thabrew et al., 2018), and Thai (Lerthattasilp et al., 2020) populations. However, despite the prevalence of depression among Spanish children and the potential value of the SMFQ, it is not yet available for the Spanish-speaking population.

Therefore, the current study extends the literature by examining for the first time the properties of the scores and factorial structure of the SMFQ in a community of Spanish-speaking children aged 8-12 years. Specifically, this study examined (a) the salience of a general factor underlying the items from the SMFQ using confirmatory factor analysis (CFA) and the unidimensional Rating Scale model (Andrich, 1978), (b) the internal consistency and test-retest reliability, and (c) the convergent and divergent validity of the SMFQ, (d) analyze possible differences in the mean total SMFQ score by children's gender and age, and (e) examine differential item functioning (DIF) of the SMFQ by gender. Based on the original study of the SMFQ (Angold et al., 1995), it is expected to find a unifactorial structure, evidence of reliability (coefficient above .75) and evidence of validity. Specifically, it was expected moderate correlations between the SMFQ and measures of internalizing symptoms (emotional problems and peer problems scales), and between the SMFQ and externalizing problems (conduct problems and hyperactivity). In keeping with prior findings with Spanish children (e.g., Canals-Sans et al., 2018), no differences in the mean total score by children's gender and age will be found. To the best of our knowledge, gender-DIF for the SMFQ has not been previously explored, so no hypotheses are previously formulated.

Methods

Participants

The sample consisted of 824 children (52.3% girls) aged between 8 and 12 years ($M_{age} = 9.64$, $SD = 1.2$). The age distribution was as follows: 22.6% ($n = 186$) were 8 years old; 23.8% ($n = 196$) were 9; 25.8% ($n = 213$) were 10; 22.7% ($n = 187$) were 11; and 5.1% ($n = 42$) were 12 years old. All children were residents and students in Spanish schools. Most children had been born in Spain (98.7%) and the rest had been born in other Western and Southeast European countries, and in South American or North Asian countries, but all of them were Spanish-speaking. More than half of the children had one sibling (60.9%), while the rest had two (18%), three (3.9%), four (1.1%), five (0.5%), six (0.4%), or no siblings (15.2%). The children were recruited from four public and private schools in urban areas in the southeast of

Spain. Concerning school grade, 24.9% of the children attended third grade of primary education, 24.3% fourth grade, 25.8% fifth grade, and 25% sixth grade.

Instruments

Demographic information was collected from the participants, such as age, gender, school, grade, number of siblings, and place of origin.

The Short Mood and Feeling Questionnaire (SMFQ; Angold et al., 1995) was developed to measure depression symptoms in youth aged 8-16 years old. It consists of 13 items with a common response format, rated on a 3-point Likert scale: 0 = not true; 1 = sometimes; 2 = always. Because the sum of the raw scores is used, the sum-score of the SMFQ ranges between 0 and 26. The higher the score, the greater intensity of existing depressive symptoms. This instrument has previously shown good internal consistency ($\alpha = .88$) and test-retest reliability ($r_{xx} = .84$) (Angold et al., 1995).

The European Spanish version of the scale was translated from English into European Spanish following the back-translation method and international recommendations for test adaptation (e.g., Hambleton, 2005; Hernández et al., 2020). Permission was first requested and obtained from the developers of the original version of the questionnaire to carry out the adaptation. Once permission was obtained, the items of the English version were translated into European Spanish by a bilingual psychologist and, afterwards, another bilingual psychologist translated it back into English. Finally, the bilingual translators compared the back-translated version with the original scale and resolved minor differences through discussion.

The Strengths and Difficulties Questionnaire (SDQ; Goodman, 1997) consists of 25-items assessing emotional and behavioral difficulties and positive behaviors in children aged 4-17 years. It comprises five subscales: Emotional Symptoms, Conduct Problems, Hyperactivity/Inattention, Peer Relationship Problems, and Prosocial Behavior, rated on a 3-point Likert scale ranging from 0 (*not true*) to 2 (*certainly true*), with a mixture of positively and negatively phrased items. The total subscale is the sum of the first four subscales. The SDQ internal consistency and its five-factor structure have been supported with Spanish children (Español-Martín et al., 2021). In this study, we also calculated the two composites or subscales of the SDQ, called Internalizing Problems (sum of the Emotional and Peer Problems subscales) and Externalizing Problems (sum of Hyperactivity and Conduct Problems subscales). Overall, the higher the score, the greater the symptomatology, except for the Prosocial Behavior subscale, where a higher score indicates more positive behavior. The Spanish version of the SDQ and instructions of how scoring it were obtained from the official SDQ website (www.sdq.org). For the five dimensions, ordinal alpha was .75 for Emotional Problems, .65 for Conduct Problems, .72 for Hyperactivity, .59 for Peer Problems, and .70 for Prosocial Behavior. Ordinal α in the current sample was .83 for the total difficulties score.

Procedure

The current study was authorized by the ethics committee of the Miguel Hernández University. Information was provided to

the school principals and parents. School approval and written parental consent were obtained for all participating children after the procedures had been fully explained and before their inclusion in the study. Children aged between 8 and 12 years were recruited from public and private primary schools located in urban areas in the southeast of Spain. Two public educational centers and two private educational centers agreed to participate in the study. These primary schools were selected for convenience, based on their availability and potential to represent the socioeconomic structure of Spanish population. The anonymous questionnaires were administered to participants in their classrooms. A researcher remained in each classroom to read the instructions and items aloud and resolve any questions. Participation was voluntary and anonymous. No incentives were provided.

Eight weeks after the first evaluation, a randomly selected subsample of children ($n = 421$) was requested to complete the same questionnaire to calculate test-retest reliability of the SMFQ.

Data Analysis

Sample descriptive analyses, as well as means and standard deviation of the items of the Spanish version of the SMFQ were run with SPSS v26. Following a decalogue for factor analysis of test items (Ferrando et al., 2022), measures of sampling adequacy for each item (MSA; Kaiser et al., 1974) and anti-image correlations (AIC; Mulaik, 2010) were also estimated. The MSA is an index between 0 and 1, values below .50 would be considered unacceptable, and would lead to the elimination of the item (Kaiser et al., 1974). Anti-image correlations refer to the reflection of the pairwise correlation remaining after partialing out the effects of other variables. Although there is not a standard rigorous cut-off point for AICs, a value of .30 has been proposed as a reasonable initial criterion by Ferrando et al. (2022), being values close to 0 more desirable. The Kolmogorov-Smirnov test indicated a violation of the assumption of normality ($p < .05$). Confirmatory factor analysis (CFA) and ordinal alpha were performed in the framework of R 3.5.2 with R Studio 1.1.453 (R Core Team, 2018). The one-factor structure of the SMFQ was tested using diagonally weighted least squares (DWLS), a variant of the weighted least squares estimator that utilizes polychoric correlations. Because the structural model includes variables with an ordinal scale, the DWLS estimator, rather than the usual maximum likelihood estimator, is deemed more appropriate (Forero et al., 2009). An acceptable fit of the model was determined by values greater than .95 for the comparative fit index (CFI), the Tucker-Lewis index (TLI) and the Goodness of fit index (GFI) and less than .06 for the root mean square error of approximation (RMSEA) and the standardized root mean square residual (SRMR) (Hu et al., 1999). We examined the unidimensionality of the SMFQ using the Rating Scale model (Andrich, 1978). Polytomous Rasch analysis was conducted using Jamovi 2.3 on all 13 SMFQ items together. When considering the average of estimated item thresholds, the Rating Scale model provides a difficulty index (δ) that indicates the item location in a single latent dimension. We considered the Information-weighted mean square statistic (Infit) and the Outlier-sensitive means square statistic (Outfit), which values tend to fall in the 0.6-1.40 interval (Bond et al., 2013). Ordinal alpha was calculated because it is considered a more accurate

estimator of reliability for unidimensional measures with dichotomic and ordinal items (Gadermann et al., 2012; Viladrich et al., 2017). Evidence of construct validity was provided by a path diagram of the SMFQ and SDQ subscales. The SDQ was selected because it assesses internalizing problems, following the strategy of the original paper of the SMFQ (Angold et al., 1995). Temporal stability of the SMFQ was examined using intraclass correlation (ICC) test-retest coefficients. Criteria values greater than .75 were considered excellent, from .40 to .75 fair to good, and values below .40 were poor (Fleiss, 1986). Due to the lack of normality in the distribution of the data, the Mann-Whitney U test were used to explore differences in the SMFQ total mean score by children's gender (group 1: girls and group 2: boys) and age (group 1: 8-9 years old and group 2: 10-12 years old), respectively. The equivalence of the SMFQ items by gender was tested using the *lordif* software package in R Studio (Choi et al., 2011). It uses ordinal logistic regression to evaluate DIF, a measurement invariance method that, in this case, informs whether females completing the Spanish version of the SMFQ endorse item response options at the same rate as males at different score levels. There are two different classes of DIF when there is a differential probability of answering depending on the gender: (a) the uniform DIF (indicates that there is a consistently different performance between groups across all score groups), and (b) the non-uniform DIF (refers to inconsistent probability of responding between groups). Models were tested at $\alpha = .01$ level.

Results

Validity related to the internal structure

CFA was performed to determine whether the unidimensional structure fit the Spanish data. The R algorithm ran the model using a bounds-constrained quasi-Newton optimization method (also known as NLMINB) and ended normally after 20 iterations. The results suggested a good fit for the 13-item single-factor structure: $\chi^2(65) = 117.50$, $\chi^2/df = 1.80$, CFI = .978, TLI = .974, GFI = .993, RMSEA = .031, 95% CI [.030, .040], SRMR = .05. Figure 1 provides the factor loadings, all statistically significant, with standardized values exceeding .25.

Rasch analysis

Rasch analysis of dimensionality uses principal component analysis and correlations of item residuals. Entering all 13 items into the analysis showed one cluster. All items showed acceptable fit statistics when tested all together, with mean square Infit and Outfit statistics within the recommended range of 0.6–1.4 (Bond et al., 2013). The measure of difficulty (δ) of the items varied between $0.44 \leq \delta \leq 2.19$. The values of Infit of the items varied between 0.80 and 1.23 and, Outfit between 0.80 and 12. Taken together, the Rasch analysis provided further evidence of the unidimensionality for the SMFQ, and that items were well fitting.

A Wright Map (Figure 2) shows the distribution of persons (on the left) and items difficulty (on the right). On the left, most of the items are located in a position centered and that the items, in general, achieve an adequate distribution in the continuum, although the curve is slightly shifted up. If the SMFQ was a

measure of ability (instead of a screening measure of depressive symptoms), it might be necessary to add some items to cover the lower sectors of the continuum. The students tended to pick answer options at or under the middle of the ordinal scale. This means that students tended to do not agree or half agree with the SMFQ statements (e.g. "I felt miserable or unhappy" or "I didn't enjoy anything"). Most students tended to choose answer options 0 (= not true) and 1 (= sometimes). This makes sense considering that a non-clinical sample was used in the current study. In a clinical sample, the option 2 "Always" would be expected to be chosen more frequently. The data also indicates that the SMFQ has a medium to high level of difficulty. The easiest items were item 4 ("I felt very agitated") and item 7 ("It was hard to think or to concentrate"), which are slightly distanced from the rest. The most difficult was item 2 ("I didn't enjoy anything"). Questions of equivalent difficulty lie at the same point on the logit scale (e.g., questions 6, 8, 9 and 11 and questions 3 and 13).

Table 1.
Rating scale analysis of the SMFQ

	δ	S.E.	Infit	Outfit
Item 1	1.19	0.06	0.80	0.80
Item 2	2.79	0.10	1.06	1.06
Item 3	1.89	0.07	1.18	1.18
Item 4	0.44	0.05	1.23	1.22
Item 5	2.13	0.08	0.99	0.99
Item 6	2.23	0.08	0.99	0.99
Item 7	0.76	0.06	0.99	0.98
Item 8	2.32	0.08	1.07	1.07
Item 9	2.40	0.08	1.02	1.02
Item 10	1.68	0.07	0.98	0.98
Item 11	2.43	0.08	1.06	1.06
Item 12	1.45	0.06	1.07	1.07
Item 13	1.73	0.07	0.91	0.91

δ = Difficulty index; S.E. = Standard Error; Infit= Information-weighted mean square statistic; Outfit= Outlier-sensitive means square statistic.

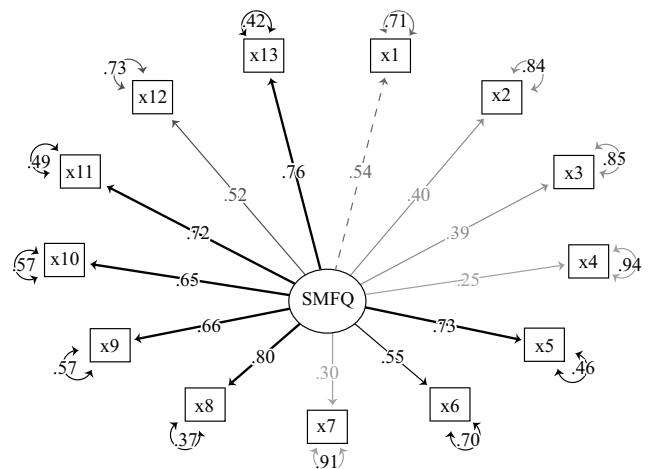


Figure 1.
Factor loading for the single-factor model.

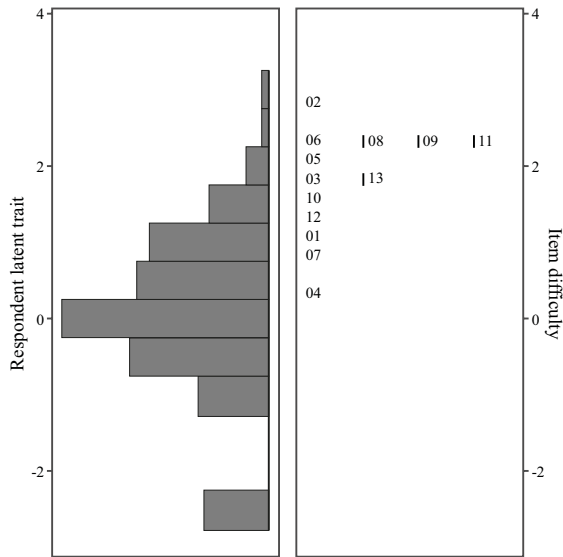


Figure 2. Wright Map of the unidimensional analysis of the SMFQ ($n = 13$ items). Individuals in the analysis are displayed on the left and items difficulty are shown on the right.

Psychometric Properties and Reliability

Table 2 shows the psychometric characteristics of the Spanish version of the SMFQ, including the mean score and standard deviation (SD) for each item, as well as the corrected item-total correlations (r_{it}^c), among others. The r_{it}^c coefficients were considered adequate. All were statistically significant, with values equal to or above .37, except for Item 2 (“I didn’t enjoy anything,” $r_{it}^c = .22$), which obtained the lowest values, followed by Item 3 (“I felt so tired that I just used to sit down without doing anything,” $r_{it}^c = .26$), Item 4 (“I felt very agitated,” $r_{it}^c = .18$), and Item 7 (“It was hard to think or concentrate,” $r_{it}^c = .23$). The internal consistency of the SMFQ was satisfactory, obtaining an ordinal α coefficient of .85.

Table 3 shows the MSA and AIC values for the Spanish version of the SMFQ. Kaiser’s measure of sampling adequacy (MSA), assessing the factorability of the observed variables, was appropriate, with MSA values between .78 and .90 (Kaiser et al., 1974). The AIC values were close to 0, and in all cases lower than 0.30, so redundant items were not found (Ferrando et al., 2022).

Table 2. Psychometric characteristics of the Spanish version of the SMFQ.

Item	<i>M</i>	<i>SD</i>	Skew	Kurtosis	r_{it}^c	$\alpha-i$
1. I felt miserable or unhappy/ Me he sentido triste	0.51	0.57	0.58	-0.65	.37	.84
2. I didn’t enjoy anything/ No he disfrutado nada	0.14	0.38	2.64	6.52	.22	.85
3. I felt so tired that I just used to sit down without doing anything/ Me he sentido tan cansando/a que no tenía ganas de hacer nada	0.31	0.56	1.68	1.79	.26	.85
4. I felt very agitated/ He estado muy inquieto/a	0.80	0.75	0.34	-1.19	.18	.85
5. I felt worthless/ He sentido que ya no hacía nada bien	0.25	0.53	2	3.06	.50	.83
6. I cried a lot/ He llorado mucho	0.23	0.48	1.96	3.04	.37	.84
7. It was hard to think or to concentrate/ Me ha costado pensar o concentrarme	0.67	0.65	0.45	-0.73	.23	.85
8. I hated myself/ Me he odiado a mí mismo/a	0.22	0.52	2.36	4.55	.53	.82
9. I was a bad person/ He sentido que era mala persona	0.20	0.47	2.30	4.58	.42	.83
10. I felt lonely/ Me he sentido solo/a	0.36	0.60	1.44	0.98	.47	.83
11. I thought that nobody loved me/ He pensado que nadie me quería	0.20	0.48	2.44	5.18	.45	.83
12. I thought that I would never be as good as other children or adolescents/ He pensado que nunca podría ser tan Bueno/a como otros/as niños/as	0.43	0.64	1.22	0.30	.37	.84
13. I did everything wrong/ He sentido que todo lo hacía mal	0.35	0.59	1.52	1.20	.57	.82

Note. SMFQ: Self-report version of the Short Mood and Feelings Questionnaire; *M*: mean; *SD*: standard deviation; MSA = Measure of Sampling Adequacy; r_{it}^c : corrected item-total correlation; $\alpha-i$: Ordinal α if the item is removed.

Table 3. Anti-image correlation matrix for the SMFQ.

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	.86 ^a	-.04	.009	.005	-.10	-.19	-.08	-.04	-.06	-.11	-.001	.01	-.05
2	-.04	.82 ^a	.01	.02	-.12	-.13	-.05	-.01	.02	-.04	.01	.002	-.04
3	.009	.01	.86 ^a	.02	-.009	-.04	-.12	-.06	-.06	-.06	.02	-.02	-.09
4	.005	.02	.02	.80 ^a	-.05	-.06	-.07	.04	-.05	-.01	.009	-.03	-.10
5	-.10	-.12	-.009	-.05	.87 ^a	-.009	.05	-.16	-.04	-.11	-.17	-.06	-.14
6	-.19	-.13	-.04	-.06	-.009	.85 ^a	.04	-.06	.003	-.07	-.07	-.03	-.06
7	-.08	-.05	-.12	-.07	.05	.04	.78 ^a	-.03	.003	-.02	.001	-.03	-.11
8	-.04	-.01	-.06	.04	-.16	-.06	-.03	.87 ^a	-.21	-.03	-.14	-.10	-.17
9	-.06	.02	-.06	-.05	-.04	.003	.003	-.21	.88 ^a	-.04	-.12	.005	-.12
10	-.11	-.04	-.06	-.01	-.11	-.07	-.02	-.03	-.04	.89 ^a	-.18	-.08	-.09
11	-.001	.01	.02	.009	-.17	-.07	.001	-.14	-.12	-.18	.87 ^a	-.07	-.04
12	.01	.002	-.02	-.03	-.06	-.03	-.03	-.10	.005	-.08	-.07	.90 ^a	-.15
13	-.05	-.04	-.09	-.10	-.14	-.06	-.11	-.17	-.12	-.09	-.04	-.15	.88 ^a

^a Measures of Sampling Adequacy (MSA).

Temporal stability of the SMFQ was explored in a randomly selected subsample of 421 children (51.09% of the whole sample) who completed the questionnaire again eight weeks later. The test-retest subsample and those who were not involved in the second evaluation were equivalent in terms of gender [$\chi^2(1) = 0.21$; $p = .64$], SDQ total score [$t(792) = 0.69$; $p = .48$], and emotional symptoms subscale of the SDQ [$t(792) = .04$; $p = .96$]. Children who responded the retest were slightly older than children who did not answer the second assessment [$t(792) = -2.61$; $p = .009$; $d = 0.19$]. ICC test-retest coefficient was .67 for the SMFQ [95% CI (.57, .74), $F = 42.92$, $p < .001$], which is considered fair to good test-retest reliability, according to Fleiss (1986).

Evidence of validity related to the relation with other variables

Figure 3 shows a path diagram to provide evidences of validity of the SMFQ along with SDQ subscales. The significant and direct correlations obtained between the SMFQ and related measures of the SDQ provided evidence of convergent validity. Specifically, good evidence of validity was obtained between the SMFQ and the SDQ Internalizing Problems subscales. The correlation found between the SMFQ and the Emotional Problems subscale was the highest, providing excellent evidence of validity, according to the criteria proposed in the European model for the evaluation of the quality of the tests (CET-R; Hernández et al., 2016). Moderate correlations were found between SMFQ and Externalizing Problems subscales, including Conduct Problems and Hyperactivity.

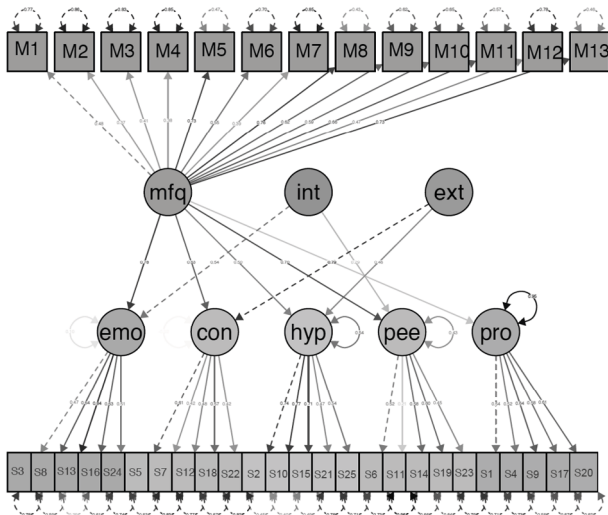


Figure 3. Path diagram including SMFQ and SDQ items. Emotional problems (emo): S3, S8, S13, S16 and S24. Conduct problems (con): S5, S7, S12, S18 and S22. Hyperactivity (hyp): S2, S10, S15, S21 and S25. Peer problems (pee): S6, S11, S14, S19 and S23. Prosocial behaviour (pro): S1, S4, S9, S17 and S20.

Differences in the mean total SMFQ score by gender and age

No statistically significant difference in the mean total SMFQ score by gender ($p = .09$) was observed. When comparing the groups of children aged 8-9 and children aged 10-12, no statistically significant differences were found ($p = .04$). The

correlation between the SMFQ total core and age was not statistically significant.

Differential item functioning (DIF)

The dataset on the 13 items of the SMFQ was analyzed for DIF related to gender. The reference and focal groups were defined as boys ($n = 393$) and girls ($n = 431$). All items had the same response-style with three response categories (0 = not true; 1 = sometimes; 2 = always). This tool is constructed in a way that higher scores mean higher level of depression. The *lordif* package terminated flagging and two items were found displaying gender related DIF: item 3 (“I felt so tired that I just used to sit down without doing anything”) and item 11 (“I thought that nobody loved me”).

Figure 4 displays the individual-level DIF impact (from *lordif*) comparing females and males. It considers differences in the SMFQ between using scores that account DIF and those that not account for DIF. On the left, the box plot shows the interquartile range (box in grey color). It represents the middle 50% of the differences, ranges roughly from -0.01 to +0.01 with a median of approximately 0. On the right, the graph represents the initial theta on the horizontal axis and the difference in theta (not accounting for DIF and after accounting for DIF) in the vertical axis. On the left of this graph, the positive values suggest that in some observations, accounting for DIF led to slightly higher scores for those with lower levels of depression, notable this appears to happen more for male individuals. On the right of this graph, the negative values show that for those with higher level of depression, accounting for DIF led to slightly higher scores, especially for females. Therefore, differences were inconsistent between genders when considering DIF items.

Figure 5 displays the trait distribution of all SMFQ items by gender. Theta distribution for gender broadly overlapped, although, females tended to present higher levels of depression than males.

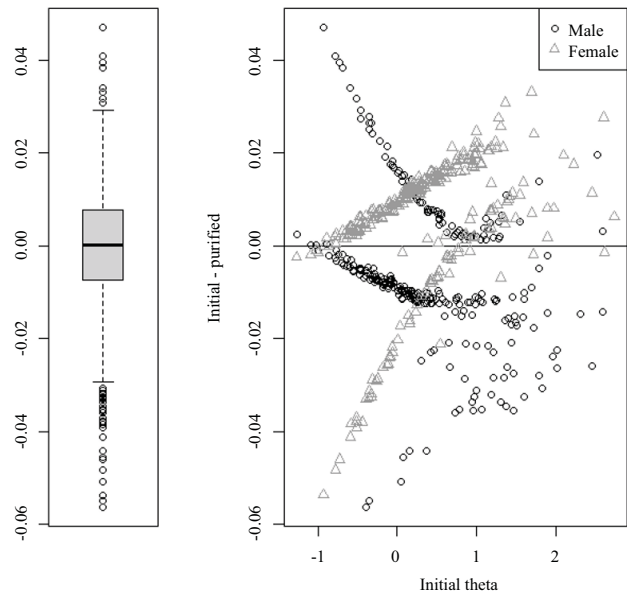


Figure 4. Individual-level DIF impact (from *lordif*) comparing females and males

Figure 6 shows the test characteristic curves (TCCs) for females and males individuals. The left plot is based on item parameter estimates for the 13 items. The differences in the estimated total scores for females and males are minimal. The right plot is only about the DIF items (items 3 and 11), and the differences are at a lower and higher level of depression (theta) for females and males in this population.

The flagged DIF items, 3 and 11, are displayed in Figures 7 and 8, respectively. Each figure includes four diagnostic plots: item characteristic curves (ICCs) (upper-left graph), the absolute difference between the ICCs for female and male groups (upper-right), the item response functions for both groups (lower-left) and the absolute difference between ICCs (upper-right graph) (lower-right). For both DIF items, individual who selects a higher response, the level of theta or depression is higher.

For item 3, the LR χ^2 was significant when comparing Model 1 and Model 2 ($p = .008$) and 1-df test for comparing Model 2 and 3 ($p = .01$); however, for 2-df test for comparing Model 2 and 3, it was not significant ($p = .19$). In this case, the slope of the male group was slightly higher than that for the female group across almost the entire continuum, except at higher values of theta or depression, suggesting a uniform DIF.

For item 11, the LR χ^2 was significant when comparing Model 1 and Model 2 ($p = .0005$) and 1-df test for comparing Model 2 and 3 ($p = .002$); however, for 2-df test for comparing Model 2 and 3, it was not significant ($p = .68$). In this case, the slope of female group was higher than that for male group indicating a uniform DIF.

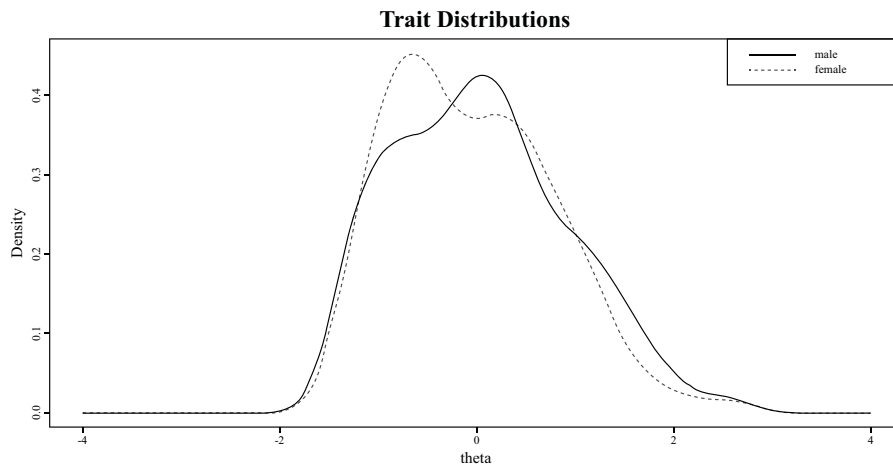


Figure 5. Trait distributions of the SMFQ items for females and males.

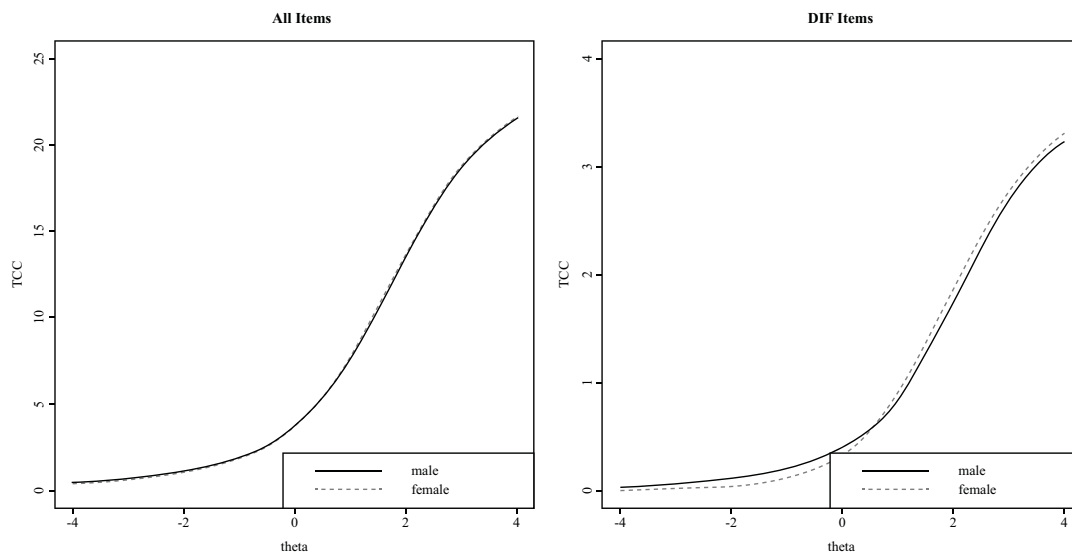


Figure 6. Differential item functioning impact with all SMFQ items included (left) and with only items with DIF included (right) comparing females and males.

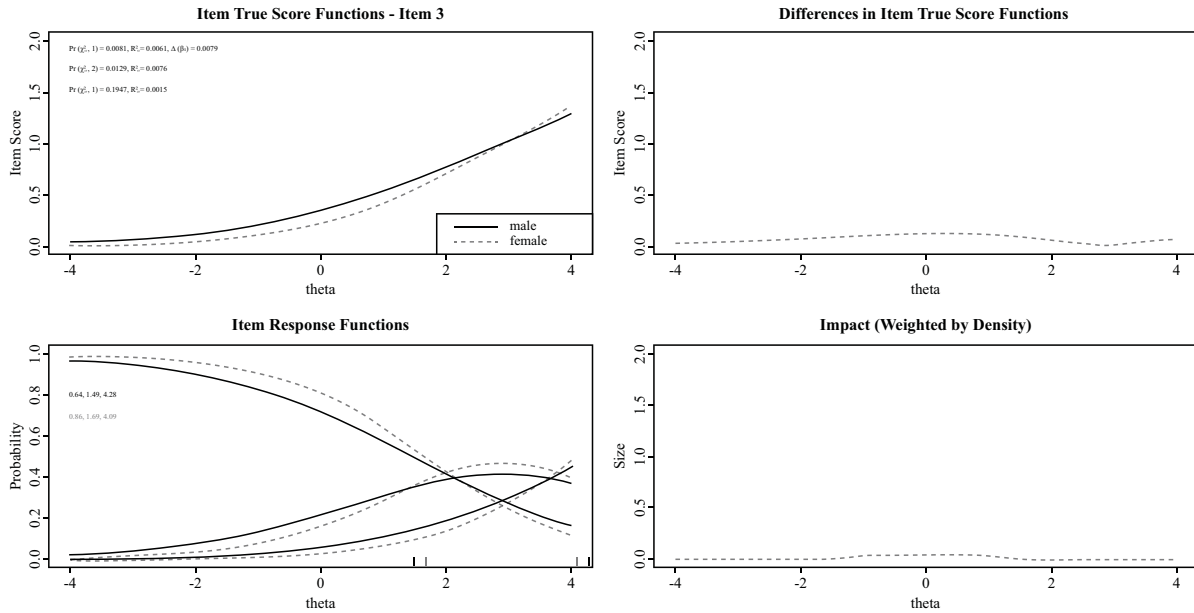


Figure 7. Visualizing the non-uniform DIF for the item 3 in terms of children’s gender.

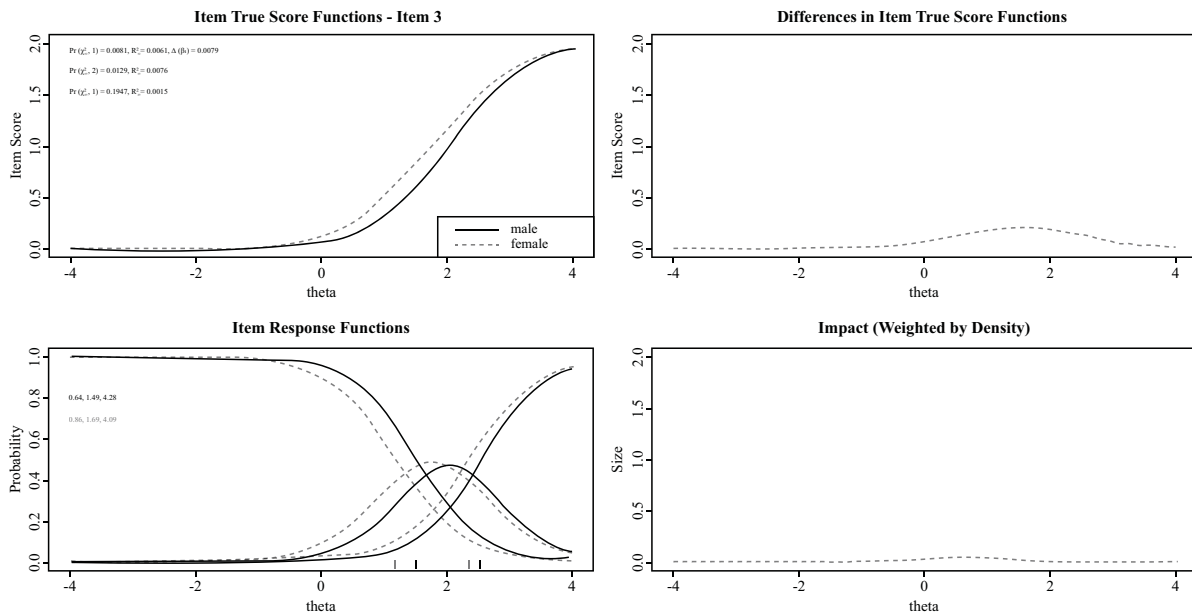


Figure 8. Visualizing the non-uniform DIF for the item 11 in terms of children’s gender.

Discussion

The SMFQ (Angold et al., 1995) is a brief, easy-to-use, self-report questionnaire that captures specific information about depressive symptoms, and can be used as an objective screening tool. The main purpose of the present study was to examine the psychometric properties and factor structure of the Spanish version of the SMFQ. The reliability and validity were supported in this

study. A high internal coefficient ($\alpha = .85$) was found. This data is in keeping with the coefficients reported by previous finding evidence of validity studies, with reliability coefficients (α) ranging from .72 to .87 (Deeba et al., 2015; Kuo et al., 2005; Olsen, 2015; Stevanovic, 2012; Sucupira et al., 2017), which have supported the internal consistency of the SMFQ across several countries and cultures. Furthermore, fair to good stability of the SMFQ was found after eight-week interval for the SMFQ (ICC = .67).

The CFA goodness-of-fit index supported the unidimensional structure of the SMFQ in the Spanish sample, with item loadings of .25 to .80. Thus, this study indicates acceptable fit to the data, implying that the items indeed measure one underlying construct, depressive symptoms (based on DSM criteria). The Rasch analysis also supported the unidimensionality for the SMFQ, and that items were well fitting. Our findings are in line with other studies analyzing the internal structure of the SMFQ (Banh et al., 2012; Deeba et al., 2015; Messer et al., 1995; Sharp et al., 2006; Stevanovic, 2012), and with the parent-reported version of MFQ in a Spanish sample (Fernández-Martínez et al., 2020), which found a single-factor structure and strong loadings for the items.

It should be noted that Item 4 (“I felt very agitated”) of the SMFQ negatively affected the internal consistency of the scale. However, its removal or that of any other item would not significantly improve the scale properties. Similarly, other studies found excellent internal consistency for all the items, except for Item 4 (Lundervold et al., 2013; Sucupira et al., 2017). Sharp et al. (2006) reported that Items 3, 4, and 7 had no discriminatory power, especially for high scores, but they contributed towards screening for patients reporting low scores on the SMFQ. The variables of restlessness and tiredness, which are assessed by these items, may be related to changes in the sleep-wake pattern during adolescence, resulting from physiological and psychological factors (Gradisar et al., 2011). In this study, the lowest scores reported were for Items 2, 3, 4, and 7, which assessed enjoyment, restlessness, tiredness, and concentration problems, respectively, all of which are symptoms of depression (Lundervold et al., 2013). Low scores may suggest that, despite their distress about their physical appearance, the children were mentally healthy (Sucupira et al., 2017).

The validity of the SMFQ was tested by comparing it with the SDQ, a measure that assesses emotional and behavioral difficulties and positive behaviors. The Spanish version of the SDQ has been cross-culturally adapted and validated as an instrument that measures a variety of aspects of mental health. Convergent validity of the SMFQ was examined through its correlations with the SDQ Emotional Problems and Internalizing Symptoms subscales, and with the total SDQ measures. The positive moderate correlations showed acceptable convergent validity. Moreover, the higher correlation with the SDQ subscales (i.e., anxiety and depression) provides further evidence of the convergent validity. In contrast, the correlations between the SMFQ and the SDQ Prosocial Behavior subscale were low and negative, supporting the divergent validity of the SMFQ. These findings are consistent with previous research reporting similar correlations between the SMFQ and other measures of anxiety and depression (Deeba et al., 2015; Turner et al., 2014), and with the parent-reported version of the MFQ with Spanish sample, which also showed low and negative correlations with the SDQ Prosocial subscale (Fernández-Martínez et al., 2020).

In line with what was hypothesized based on prior findings with Spanish children (Canals-Sans et al., 2018), we did not find statistically significant differences in the mean total SMFQ score by gender and age. This suggests that in this sample the level of depressive symptomatology is similar between girls and boys, and among all participants aged 8 to 12. Nonetheless, authors

such as Angold et al. (2002) have pointed out that findings from studies with children and adolescents on age and gender effects based on depression questionnaires scores tend to be inconsistent, with mixed and contradictory results. Thus, this study provides additional data with a sample of Spanish-speaking school-aged children, but more research is needed to better understand this issue.

Moreover, analyzing the measurement invariance of the SMFQ, it was found that two items (3 and 11) showed gender-related DIF. It means that answers to these items were dependent on children's gender. DIF revealed uniform DIF, that is to say, there is a consistently different performance between groups across all score groups. Both items are not fair based on children's gender; however, the observed DIF was almost insignificant. Based on these results, the SMFQ presents validity related to measurement invariance across gender and it may be used in both genders. Despite of this, the interpretation of the scores must consider our results of DIF across gender.

This research presents some strengths but also some limitations to be considered. The main limitation concerns the data, collected from a community sample, and thus, it remains unclear whether the results could be generalized to clinical samples. However, the use of community-based or population-based samples enables one to generalize results to a greater degree (Olsen, 2015). Second, the sample comes from a localized geographical area of Spain, the southeast region, so future studies should be carried out with samples from other regions to facilitate greater generalization of the results. Third, in this study, no other specific measures were used to examine criterion validity, so it would be interesting to examine the validity by including other specific measures for depression and anxiety in future research. Finally, the lower discrimination coefficient of item 2 may be due to the item construction; it is a double negation that may make it difficult to understand and answer. Future revisions of the SMFQ should consider its wording to improve it.

Despite the above limitations, this research extends the literature by providing additional evidence of the psychometric properties and factor structure of the SMFQ in Spanish-speaking children. The SMFQ exhibited a unidimensional internal structure, thereby replicating prior findings across culture and language. Therefore, the use of SMFQ as a continuous measure of severity of depressive symptoms is also supported in the Spanish population. This is important for the continued improvement in the early detection of children at risk of mental health problems. In this regard, to provide proper preventive care or implement appropriate interventions, valid, reliable, and thoroughly vetted tools are needed to map and identify at-risk or affected children and adolescents. Thus, for the first time, initial evidence is provided of the usefulness of a version of the SMFQ for the assessment of depression in Spanish-speaking children, further supporting the suitability of the use of this tool by clinicians and researchers for the early detection of childhood depression.

Funding

This study did not receive specific funding from public sector agencies, the commercial sector, or non-profit bodies.

References

- Adeyuya, A. O., Ola, B. A., Dada, A. O., & Fasoto, O. O. (2006). Validation of the Edinburgh Postnatal Depression Scale as a screening tool for depression in late pregnancy among Nigerian women. *Journal of Psychosomatic Obstetrics and Gynecology*, 27(4), 267-272. <https://doi.org/10.1080/01674820600915478>
- Aggen, S. H., Neale, M. C., & Kendler, K. S. (2005). DSM criteria for major depression: Evaluating symptom patterns using latent-trait item response models. *Psychological Medicine*, 35, 475-487. <https://doi.org/10.1017/S0033291704003563>
- American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders, fifth edition (DSM-5)*. American Psychiatric Publishing. <https://doi.org/10.4135/9781483365817.n433>
- Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika*, 43(4), 561-573. <https://doi.org/10.1007/BF02293814>
- Angold, A., Costello, E. J., Messer, S. C., Pickles, A., Winder, F., & Silver, D. (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.
- Angold, A., Costello, E. J., Pickles, E., & Winder, F. (1987). *The development of a questionnaire for use in epidemiological studies of depression in children and adolescents*. Unpublished manuscript.
- Angold, A., Erkanli, A., Silberg, J., Eaves, L., & Costello, E. J. (2002). Depression scale scores in 8-17-year-olds: Effects of age and gender. *Journal of Child Psychology and Psychiatry*, 43(8), 1052-1063. <https://doi.org/10.1111/1469-7610.00232>
- Avenevoli, S., Swendsen, J., He, J. P., Burstein, M., & Merikangas, K. R. (2015). Major depression in the national comorbidity survey-adolescent supplement: Prevalence, correlates, and treatment. *Journal of the American Academy of Child & Adolescent Psychiatry*, 54(1), 37-44. <https://doi.org/10.1016/j.jaac.2014.10.010>
- Banh, M. K., Crane, P. K., Rhew, I., Gudmundsen, G., Stoep, A. V., Lyon, A., & McCauley, E. (2012). Measurement equivalence across racial/ethnic groups of the Mood and Feelings Questionnaire for childhood depression. *Journal of Abnormal Child Psychology*, 40, 353-367. <https://doi.org/10.1007/s10802-011-9569-4>
- Beck, A. T. (1961). *Beck Depression Inventory* [Measurement Instrument]. Center for Cognitive Therapy.
- Bernaras, E., Jaureguizar, J., & Garaigordobil, M. (2019). Child and adolescent depression: A review of theories, evaluation instruments, prevention programs, and treatments. *Frontiers in Psychology*, 10, Article 543. <https://doi.org/10.3389/fpsyg.2019.00543>
- Bond, T. G., & Fox, C. M. (2013). *Applying the Rasch model: Fundamental measurement in the human sciences*. Psychology Press.
- Canals-Sans, J., Hernández-Martínez, C., Sáez-Carles, M., & Arijá-Val, V. (2018). Prevalence of DSM-5 depressive disorders and comorbidity in Spanish early adolescents: Has there been an increase in the last 20 years? *Psychiatry Research*, 268, 328-334. <https://doi.org/10.1016/j.psychres.2018.07.023>
- Choi, S. W., Gibbons, L. E., & Crane, P. K. (2011). Lordif: An R package for detecting differential item functioning using iterative hybrid ordinal logistic regression/item response theory and Monte Carlo simulations. *Journal of Statistical Software*, 39(8), 1-30. <https://doi.org/10.18637/jss.v039.i08>
- Clark, C., Rodgers, B., Caldwell, T., Power, C., & Stansfeld, S. (2007). Childhood and adulthood psychological ill health as predictors of midlife affective and anxiety disorders: The 1958 British birth cohort. *Archives of General Psychiatry*, 64, 668-678. <https://doi.org/10.1001/archpsyc.64.6.668>
- Costello, E. J., Benjamin, R., Angold, A., & Silver, D. (1991). Mood variability in adolescents: A study of depressed, nondepressed and comorbid patients. *Journal of Affective Disorders*, 23(4), 199-212. [https://doi.org/10.1016/0165-0327\(91\)90101-w](https://doi.org/10.1016/0165-0327(91)90101-w)
- Deeba, F., Rapee, R. M., & Prvan, T. (2015). Psychometric properties of two measures of childhood internalizing problems in a Bangladeshi sample. *British Journal of Clinical Psychology*, 54, 214-232. <https://doi.org/10.1111/bjc.12071>
- Español-Martín, G., Pagerols, M., Prat, R., Rivas, C., Sixto, L., Valero, S., Artigas, M. S., Ribasés, M., Ramos-Quiroga, J. A., Casas, M., & Bosch, R. (2021). Strengths and Difficulties Questionnaire: Psychometric Properties and Normative Data for Spanish 5 to 17 Years Old. *Assessment*, 28(5), 1445-1458. <https://doi.org/10.1177/1073191120918929>
- Fernández-Martínez, I., Morales, A., Espada, J. P., & Orgilés, M. (2020). Psychometric properties and factorial structure of the Spanish version of the parent-report Mood and Feelings Questionnaire (MFQ-P). *Clinical Child Psychology and Psychiatry*, 25(2), 533-544. <https://doi.org/10.1177/1359104519897939>
- Ferrando, P. J., Lorenzo-Seva, U., Hernández-Dorado, A., & Muñoz, J. (2022). Decalogue for the factor analysis of test items. *Psicothema*, 34(1), 7-17. <https://doi.org/10.7334/psicothema2021.456>
- Fleiss, J. L. (1986). *The design and analysis of clinical experiments*. Wiley.
- Forero, C. G., Maydeu-Olivares, A., & Gallardo-Pujol, D. (2009). Factor analysis with ordinal indicators: A Monte Carlo study comparing DWLS and ULS estimation. *Structural Equation Modeling*, 16(4), 625-641. <https://doi.org/10.1080/10705510903203573>
- Gademann, A. M., Guhn, M., & Zumbo, B. D. (2012). Estimating ordinal reliability for Likert-type and ordinal item response data: A conceptual, empirical, and practical guide. *Practical Assessment, Research & Evaluation*, 17(3), 1-13.
- Goodman, R. (1997). The Strengths and Difficulties Questionnaire: A research note. *Journal of Child Psychology & Psychiatry & Allied Disciplines*, 38, 581-586. <https://doi.org/10.1037/00540-000>
- Gradisar, M., Gardner, G., & Dohnt, H. (2011). Recent worldwide sleep patterns and problems during adolescence: A review and meta-analysis of age, region, and sleep. *Sleep Medicine*, 12(2), 10-118. <https://doi.org/10.1016/j.sleep.2010.11.008>
- Hambleton, R. K. (2005). Issues, designs, and technical guidelines for adapting tests into multiple languages and cultures. In R. K. Hambleton, P. F. Merenda, & C. D. Spielberger (Eds.), *Adapting educational and psychological tests for cross-cultural assessment* (pp. 3-38). Lawrence Erlbaum Associates. <https://doi.org/10.4324/9781410611758-6>
- Hankin, B. L. (2015). Depression from childhood through adolescence: Risk mechanisms across multiple systems and levels of analysis. *Current Opinion in Psychology*, 4, 13-20. <https://doi.org/10.1016/j.copsyc.2015.01.003>
- Hernández, A., Hidalgo, M. D., Hambleton, R. K., & Gómez-Benito, J. (2020). International Test Commission guidelines for test adaptation: A criterion checklist. *Psicothema*, 32(3), 390-398. <https://doi.org/10.7334/psicothema2019.306>
- Hernández, A., Ponsoda, V., Muñoz, J., Prieto, G., & Elosua, P. (2016). Revisión del modelo para evaluar la calidad de los tests utilizados en España [Review of the model to evaluate the quality of the tests used in Spain]. *Papeles del Psicólogo*, 37(3), 192-197.
- 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), 1-55. <https://doi.org/10.1080/10705519909540118>
- Jeffreys, M., Rozenman, M., Gonzalez, A., Warnick, E. M., Dauser, C., Scahill, L., Woolston, J., & Weersing, V. R. (2016). Factor structure of the parent-

- report Mood and Feelings Questionnaire (MFQ) in an outpatient mental health sample. *Journal of Abnormal Child Psychology*, 44, 1111-1120. <https://doi.org/10.1007/s10802-015-0108-6>
- Kaiser, H. F., & Rice, J. (1974). Little jiffy, mark IV. Educational and Psychological Measurement, 34(1), 111-117. <https://doi.org/10.1177/001316447403400115>
- Kovacs, M. (1992). *Children's Depression Inventory Manual*. Multi-Health Systems. <https://doi.org/10.1037/t00788-000>
- Kuo, E. S., Stoep, A. V., & Stewart, D. G. (2005). Using the Short Mood and Feelings Questionnaire to detect depression in detained adolescents. *Assessment*, 12(4), 374-383. <https://doi.org/10.1177/1073191105279984>
- Lawton, A., & Moghraby, O. S. (2016). Depression in children and young people: Identification and management in primary, community and secondary care (NICE guideline CG28). *Archives of Disease in Childhood: Education and Practice*, 101, 206-209. <https://doi.org/10.1136/archdischild-2015-308680>
- Lerthattasilp, T., Tapanadechopone, P., & Butrdeewong, P. (2020). Validity and reliability of the Thai version of the Short Mood and Feelings Questionnaire. *East Asian Archives of Psychiatry*, 30(2), 48-51. <https://doi.org/10.12809/eaap1875>
- Lundervold, A. J., Breivik, K., Posserud, M. B., Stommark, K. M., & Hysing, M. (2013). Symptoms of depression as reported by Norwegian adolescents on the Short Mood and Feelings Questionnaire. *Frontiers in Psychology*, 4, Article 613. <https://doi.org/10.3389/fpsyg.2013.00613>
- McKenzie, D. P., Toumbourou, J. W., Forbes, A. B., Mackinnon, A. J., McMorris, B. J., Catalano, R. F., & Patton, G. C. (2011). Predicting future depression in adolescents using the Short Mood and Feelings Questionnaire: A two-nation study. *Journal of Affective Disorders*, 134, 151-159. <https://doi.org/10.1016/j.jad.2011.05.022>
- Messer, S. C., Angold, A., Costello, E. J., Loeber, R., Van Kammen, W., & Stouthamer-Loeber, M. (1995). Development of a short questionnaire for use in epidemiological studies of depression in children and adolescents. Factor composition and structure across development. *International Journal of Methods in Psychiatric Research*, 5, 251-262.
- Mulaik, S. A. (2010). *Foundations of factor analysis*. Chapman & Hall.
- Olsen, K. S. (2015). *Symptoms of depression in adolescents: Examining the psychometric properties of a Norwegian translation of the Short Mood and Feelings Questionnaire* [Master's thesis]. University of Oslo.
- Pinto, E. R. (2014). *Short Mood and Feelings Questionnaire: tradução para língua portuguesa, adaptação cultural e validação*. [Master's thesis]. University of Sao Paulo.
- R Core Team (2018). RStudio: integrated development for R (Version 1.1.453). Computer software.
- Santa Cruz, C., Espinoza, V., & Hohlberg, E. (2021). Problemas socioemocionales en niños con discapacidad auditiva, discapacidad visual y desarrollo típico [Socioemotional problems in children with hearing impairment, visual impairment and typical development]. *Revista Latinoamericana de Educación Inclusiva*, 15(1), 95-116. <https://doi.org/10.4067/S0718-73782021000100095>
- Sharp, C., Goodyer, I. M., & Croudace, T. J. (2006). The Short Mood and Feelings Questionnaire (SMFQ): A unidimensional item response theory and categorical data factor analysis of self-report ratings from a community sample of 7 through 11-year-old children. *Journal of Abnormal Child Psychology*, 34(3), 365-377. <https://doi.org/10.1007/s10802-006-9027-x>
- Stevanovic, D. (2012). Childhood depression and anxiety disorders in Serbia: A psychometric study of four screening questionnaires. *Epidemiology and psychiatric sciences*, 21(1), 111-116. <https://doi.org/10.1017/S20457960110000655>
- Sucupira, E., Sabino, M., Luiz, E., Moreira, G., Azevedo, M. J., & Masako, L. (2017). Short Mood and Feelings Questionnaire for screening children and adolescents for plastic surgery: Cross-cultural validation study. *Sao Paulo Medical Journal*, 135(6), 518-528. <https://doi.org/10.1590/1516-3180.2017.0036030517>
- Thabrew, H., Stasiak, K., Bavin, L. M., Frampton, C., & Merry, S. (2018). Validation of the mood and feelings questionnaire (MFQ) and short mood and feelings questionnaire (SMFQ) in New Zealand help-seeking adolescents. *International Journal of Methods in Psychiatric Research*, 27(3), 1-9. <https://doi.org/10.1002/mpr.1610>
- Turner, N., Joinson, C., Peters, T. J., Wiles, N., & Lewis, G. (2014). Validity of the Short Mood and Feelings Questionnaire in late adolescence. *Psychological Assessment*, 26(3), 752-762. <https://doi.org/10.1037/a0036572>
- Viladrich, C., Angulo-Brunet, A., & Doval, E. (2017). A journey around alpha and omega to estimate internal consistency reliability. *Anales de Psicología*, 33(3), 755-782. <https://doi.org/10.6018/analesps.33.3.268401>
- World Health Organization (2017). *Depression and other common mental disorders: Global health estimates*. World Health Organization.
- Yu, M., Xu, W., Xie, Q., Zhu, Y., Chasson, G. S., & Wang, J. (2017). Automatic thoughts as a predictor of internalizing and externalizing problems in Chinese adolescents: A test of the cognitive content-specificity hypothesis with age effects. *Scandinavian Journal of Psychology*, 58(5), 351-358. <https://doi.org/10.1111/sjop.12373>