Spanish Adaptation and Psychometric Properties of the Parent Version of the Short Mood and Feelings Questionnaire (SMFQ-P) in a Non-Clinical Sample of Young School-Aged Children

Iván Fernández-Martínez1, Alexandra Morales1, Francisco Xavier Méndez2, José Pedro Espada1 and Mireia Orgilés1

1 Universidad Miguel Hernández de Elche (Spain)
2 Universidad de Murcia (Spain)

Abstract. The parent version of the Short Mood and Feelings Questionnaire (SMFQ-P) is a brief 13-item tool for the screening of depression in children from the age of 6 years. Despite the wide use of the SMFQ-P, its psychometric properties and factor structure remain understudied, with few data available for young school-aged children. The objective of this study was to examine for the first time the factorial structure and psychometric properties of the SMFQ-P in a non-clinical sample of Spanish-speaking children aged 6–8 years. Participants were 181 children whose parents completed the Spanish-adapted version of the SMFQ-P along with the parent version of other measures of anxiety and general difficulties and positive attributes. The SMFQ-P demonstrated adequate internal consistency (α = .83) and test-retest reliability over an eight-week period (ICC = .80), and good convergent and divergent validity. Factor analysis confirmed the original 13-item model, thus supporting the unidimensionality of the measure in the Spanish sample. Overall, this study provides initial empirical evidence for the utility of the SMFQ-P with Spanish-speaking children from early school ages, and extends the international support of the measure.

Keywords: assessment, depression, parent report, Short Mood and Feelings Questionnaire, Spanish validation

Along with anxiety and behavior disorders, mood disorders are one of the most common mental health problems in children, with an estimated prevalence ranging from 0.6% to 3% for major depressive disorder (Merikangas et al., 2009; Polanczyk et al., 2015). Depressive symptoms can appear early in life, from preschool ages, remain stable or increase over time and have heterotypic continuity with other emotional problems such as anxiety (Bufford et al., 2012; Lewis et al., 2020). Furthermore, early manifestation of depressive symptoms in children strongly predicts subsequent disorders, such as major depression disorder, anxiety, and attention deficit hyperactivity disorder during school age and early adolescence (Luby et al., 2014). Thus, the need and importance of early identification of depressive symptoms has been underlined (Luby, 2010; Wesselhoeft et al., 2013). In this regard, there is a need to continue to develop and adapt tests to diagnose depression from an early age (Bernaras et al., 2019).

The value of a multi-informant assessment for an accurate assessment of children’s internalizing problems, including parent-reports, has been supported (Izquierdo-Sotorrío et al., 2016; Kerr et al., 2007). Although a low-to-moderate cross-informant correspondence, as well as parents’ tendency to under-report child depression have been found, the use of multiple informants allows a more comprehensive assessment, as the manifestation of children’s mental health problem may vary according to the context. This also allows the adaptation of treatment to the children’s needs, providing parents a valuable perspective of their child’s emotional problems (De Los Reyes et al., 2015; Kim et al., 2016). It has also been noted that self-reported scales

generally appear to be more helpful in adolescence, with more limited use in younger children (Rey et al., 2015). Moreover, following Dougherty et al.'s (2008) review, parents as a source of information are especially useful in the case of pre-adolescent and younger children.

Rating scales are cost-effective and widely employed for the screening of depression (Dougherty et al., 2008; Rey et al., 2015), and several measures for the screening of childhood depression with sound psychometric properties offering a parent-rating form can be found (Bernaras et al., 2019; Costello & Angold, 1988). However, it has been argued that few instruments are available free of charge, which is critical as it may limit their accessibility and use (Jeffreys et al., 2016). One of the scales available for measuring childhood depression through children and parents is the Short Mood and Feelings Questionnaire (SMFQ), a 13-item scale designed to provide a quick assessment and screening of core DSM-based depressive symptoms in children and adolescents aged 6–17 years (Angold et al., 1995).

The use of the SMFQ offers sound advantages, such as being a brief, simple, and easy-to-complete screening instrument, specifically developed for children and adolescents, and available free of charge. The SMFQ allows the assessment of childhood depression including affective and cognitive items, but also others related to restlessness, lack of concentration or tiredness. Both the self-report and parent-report (SMFQ-P) have been shown to have a unifactorial structure and good psychometric properties in community samples, highlighting their ability to detect children at risk of depression and discriminate clinical from non-clinical cases (e.g., Rhew et al., 2010; Sharp et al., 2006; Thapar & McGuffin, 1998). Recently, this measure has been cross-culturally adapted to Norwegian (Lundervold et al., 2013), Serbian (Stevanovic, 2012), Bangladeshi (Deeba et al., 2015), New Zealand (Thabrew et al., 2018), Chinese (Cheng et al., 2009), and Brazilian-Portuguese (Sucupira et al., 2017) populations. However, it has not yet been validated for use in Spanish-speaking children (Bernaras et al., 2019).

On another hand, although the SMFQ-P has shown to be a useful scale to assess childhood depression, even for evaluation of treatment response in early school-aged children (Creswell et al., 2020; Tennant et al., 2017), few data are available about its psychometric properties in young children (i.e., 6–8 years), as studies have generally focused on older ages (e.g., Karevold et al., 2017; Sucupira et al., 2017). Therefore, it would be valuable to provide further data regarding the factor structure, reliability (i.e., including test-retest reliability) and validity of the SMFQ-P with young school-aged children.

Thus, the aim of the present study was to examine in a sample of primary school Spanish-speaking children aged 6 to 8: (a) The factor structure; (b) the internal consistency and test-retest reliability; and (c) the convergent and divergent validity of the SMFQ-P. Based on findings in previous studies (e.g., Angold et al., 1995; Karevold et al., 2017; Sucupira et al., 2017), an adequate fit of the Spanish data to the single-factor structure of the SMFQ-P was hypothesized, and reliability was expected to be good. Regarding convergent and divergent validity, positive significant correlations were expected between the SMFQ and measures of emotional psychopathology (i.e., depression and anxiety), and lower correlations with a non-emotional construct measuring prosocial behavior (e.g., Muris et al., 2003; Zhao et al., 2012).

Method

Participants

The sample included 181 children aged between 6 and 8 years ($M = 6.87$, $SD = .79$); 98 boys (54.1%) and 83 girls (45.9%), all of them Caucasian. The age distribution was as follows: 38.1% were 6 years old, 37% were 7 years old, and 24.9% were 8. Participants were recruited from ten primary schools from southeastern Spain, selected according to their availability to represent the socio-economic structure of Spanish population. Children were in Grades 1 (44.8%), 2 (35.4%), and 3 (19.8%) of primary education. Most children had been born in Spain (97.8%), whereas the remaining 2.2% had been born in other countries, but all the children were Spanish-speakers. Of all the children, 76.2% had between one and three siblings, the remaining 23.8% had none. A total of 181 parents (81.8% mothers and 18.2% fathers) completed questionnaires about their children. Regarding parental status, 86.2% were married, 12.2% divorced, and 1.6% single parents. Most of participants came from middle-class families. Parents’ levels of study were as follows: Primary education, 16.6%, secondary education, 30.9%; and higher education, 52.5%.

Measures

Short Mood and Feelings Questionnaire - Parent Version (SMFQ-P). The SMFQ-P (Angold et al., 1995) consists of a brief 13-item measure to assess depressive symptomatology in children and adolescents, mainly assessing affective and cognitive symptoms. It comprises descriptive phrases (e.g., s/he was very restless) on which parents rate how their children have been acting and feeling over the past two weeks on a 3-point Likert scale (0 = not true; 1 = sometimes; 2 = true). The total SMFQ-P score is obtained by summing all the responses (score range: 0–39), with higher scores indicating more severe symptoms of depression. The
SMFQ-P has shown good validity and reliability ($\alpha = .84$–.87) (Angold et al., 1995; Rhew et al., 2010).

In the current study, the original English version of the SMFQ was translated into European-Spanish, after receiving permission from the author, following the accepted back-translation guidelines proposed by Hambleton (2005). Thus, the 13 items of the scale were translated from English to Spanish by a bilingual psychologist. Then, another bilingual psychologist, who had not seen the original version, translated the items back to English. Finally, the back-translated version was compared with the original version, resolving minor differences by agreement among the Spanish-speaking translators.

Strengths and Difficulties Questionnaire-Parent version (SDQ-P). The SDQ-P (Goodman, 2001) is a widely used, 25-item scale assessing general difficulties (i.e., emotional and behavioral) and positive behaviors used, 25-item scale assessing general difficulties (SDQ-P). The SDQ-P (Goodman, 2001) is a widely speaking translators.

Minor differences by agreement among the Spanish- speaking translators. Those who had not seen the original version, translated the items back to English. Finally, the back-translated version was compared with the original version, resolving minor differences by agreement among the Spanish-speaking translators.

The present study was approved by the ethics boards of all institutions. Twelve schools from urban areas were invited to the study and 10 agreed to participate. School principals’ approval was obtained, and schools distributed a letter to parents with information about the research. Information was sent to 1,400 parents of children aged 6 to 18, of which 181 (12.9%) participated. Similar to other works addressing children’s emotional problems (e.g., Orgilés, Fernández-Martínez, et al., 2019; Orgilés et al., 2020), in this study parents were asked to complete an online assessment. Parents who were interested in participating accessed an online form and voluntarily completed the questionnaires, which took around 15 minutes. Prior to the completion of questionnaires, parents were informed about the purposes and procedure of the research, that data would be treated confidentially, and that they could withdraw at any time, as participation was voluntary. All of them accepted to participate and gave informed consent. Parents were asked to complete the online form again 8 weeks later (test-retest period). No incentives were given for their collaboration.

### Statistical Analyses

We run data analysis after checking that there were no missing data. The aim of the current study was to test the original unifactorial structure proposed by Angold et al. (1995) to validate the SMFQ-P for Spanish children. To meet this objective, we carried out confirmatory factor analysis (CFA) using the Lavaan package in R Studio (Rossel, 2012). The lack of normality of the sample distribution in the SMFQ-P scores was confirmed through the Kolmogorov-Smirnov test (Steinskog et al., 2007). Values of skewness and kurtosis less than 1.0 were considered slight nonnormality, the values between 1.0 and about 2.3 were considered as moderate nonnormality, and the values beyond 2.3 were considered as severe nonnormality (Lei & Lomax, 2005). Factor analysis was based on the polychoric correlation matrix as it is recommended for ordinal polytomous items that are asymmetric or with an excess of kurtosis. The diagonally weighted least squares (DWLS) Robust was used due to its robustness with ordinal data (Forero et al., 2009). The model fit indices examined to determine the appropriateness of the factor solution were: $\chi^2$/degrees of freedom ($df$), comparative fit index (CFI), Tucker-Lewis index (TLI), and root mean square error of approximation (RMSEA). An adequate fit of the model of the SMFQ-P for the Spanish data was set as follows: $\chi^2$/df $\leq 3$, CFI and TLI $\geq .90$, and RMSEA $\leq .08$ (Hu & Bentler, 1999). Because two models were run (one with all items, and another removing item 1), $\Delta\chi^2$ and $\Delta df$ were also reported.

Furthermore, we computed descriptive statistics (means, standard deviations, corrected item-total correlation, and $t$ if the item is removed) and internal consistency ($\alpha$). Temporal stability of the measure was explored in a subsample of 42 children ($M = 6.88, SD = .80, 40.5$% girls) (23.20% of the sample) using the intraclass correlation coefficient (ICC) for test-retest. We compared scores obtained at baseline and a posttest after two months. ICC values above 0.60 were considered good (Anastasi, 1998; Baumgartner & Chung,
Attrition analyses were conducted to test possible differences in sociodemographic variables and in the SMFQ-P, SCAS-P and SDQ-P subscales between participants involved in the test-retest and those who were not. Convergent and divergent validity was explored through conducting bivariate correlations with the SCAS-P total score and SDQ-P Emotional symptoms and Prosocial behavior scores. Differences in SMFQ-P scores by children's gender and age were explored using t-test and one-way ANOVA. All descriptive analyses were conducted using SPSS v25.

**Results**

**Descriptive Statistics**

Table 1 reports descriptive statistics (means, standard deviations, skewness, kurtosis, corrected item-total correlation, and α if the item is removed) of the SMFQ-P. SMFQ-P mean scores were 4.59 (SD = 4.27). No differences in SMFQ-P scores by children's gender and age were found. Most items presented an excess of kurtosis or a non-normality, ranging from moderate to severe (Lei & Lomax, 2005). The corrected item-total correlations were considered adequate, all statistically significant, with values equal to or above .41, except for Item 1 ("S/he felt miserable or unhappy"), which was lower (.13). Cronbach alpha if the item removed was computed and indicated that removing Item 1 would slightly increase the internal consistency of the scale, suggesting that the reliability of this item is not adequate for the Spanish sample.

<table>
<thead>
<tr>
<th>Item</th>
<th>M</th>
<th>SD</th>
<th>Kurtosis</th>
<th>Skewness</th>
<th>r_{it}^{c}</th>
<th>α-1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. S/he felt miserable or unhappy</td>
<td>0.76</td>
<td>0.79</td>
<td>-1.65</td>
<td>-22</td>
<td>.13</td>
<td>.85</td>
</tr>
<tr>
<td>2. S/he didn't enjoy anything at all</td>
<td>0.29</td>
<td>0.48</td>
<td>0.82</td>
<td>1.37</td>
<td>.55</td>
<td>.81</td>
</tr>
<tr>
<td>3. S/he felt so tired that s/he just sat around and did nothing</td>
<td>0.17</td>
<td>0.47</td>
<td>7.16</td>
<td>2.80</td>
<td>.41</td>
<td>.82</td>
</tr>
<tr>
<td>4. S/he was very restless</td>
<td>0.55</td>
<td>0.69</td>
<td>-0.45</td>
<td>0.88</td>
<td>.41</td>
<td>.82</td>
</tr>
<tr>
<td>5. S/he felt s/he was no good anymore</td>
<td>0.29</td>
<td>0.54</td>
<td>2.17</td>
<td>1.75</td>
<td>.65</td>
<td>.80</td>
</tr>
<tr>
<td>6. S/he cried a lot</td>
<td>0.50</td>
<td>0.68</td>
<td>-0.21</td>
<td>1</td>
<td>.49</td>
<td>.81</td>
</tr>
<tr>
<td>7. S/he found it hard to think properly or concentrate</td>
<td>0.50</td>
<td>0.65</td>
<td>-0.21</td>
<td>0.94</td>
<td>.50</td>
<td>.81</td>
</tr>
<tr>
<td>8. S/he hated him/herself</td>
<td>0.07</td>
<td>0.25</td>
<td>13.97</td>
<td>3.68</td>
<td>.46</td>
<td>.82</td>
</tr>
<tr>
<td>9. S/he felt s/he was a bad person</td>
<td>0.13</td>
<td>0.35</td>
<td>5.76</td>
<td>2.55</td>
<td>.48</td>
<td>.81</td>
</tr>
<tr>
<td>10. S/he felt lonely</td>
<td>0.25</td>
<td>0.49</td>
<td>2.46</td>
<td>1.80</td>
<td>.59</td>
<td>.81</td>
</tr>
<tr>
<td>11. S/he thought nobody really loved him/her</td>
<td>0.26</td>
<td>0.51</td>
<td>2.57</td>
<td>1.83</td>
<td>.47</td>
<td>.81</td>
</tr>
<tr>
<td>12. S/he thought s/he could never be as good as other kids</td>
<td>0.42</td>
<td>0.67</td>
<td>0.44</td>
<td>1.33</td>
<td>.63</td>
<td>.81</td>
</tr>
<tr>
<td>13. S/he felt s/he did everything wrong</td>
<td>0.40</td>
<td>0.61</td>
<td>0.57</td>
<td>1.28</td>
<td>.65</td>
<td>.80</td>
</tr>
<tr>
<td>Total</td>
<td>4.59</td>
<td>4.27</td>
<td>.83</td>
<td>.80</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Note.** M = Mean; SD = Standard Deviation; r_{it}^{c} = corrected item-total correlation; α-1 = Cronbach alpha if the item is removed; α = Cronbach alpha; ICC = Intraclass correlation coefficient for test-retest.

**Confirmatory Factor Analysis**

The original one-factor model of the SMFQ-P provided a good fit to the data in the Spanish sample. The model fit indices examined suggested an appropriate fit for the single-factor solution, $\chi^2 = 117.12, df = 65, \chi^2/df = 1.80$, CFI = .96, TLI = .96, and RMSEA = .06, 95% CI [.04, .08]. Table 2 provides the factor loadings, all statistically significant with standardized values exceeding .30, except for Items 12 and 13, which had slightly lower values. It should be noted that Item 1 ("S/he felt miserable or unhappy") showed the highest factor load value (.77). However, based on above described results suggesting the reliability improvement by removing Item 1, we ran a CFA to test the one-factor model, but removing this item. The new specification of the model achieved slightly better fit indices than the 13-item one-factor model: $\chi^2 = 94.03, df = 54, \chi^2/df = 1.74$, CFI = .97, TLI = .96, and RMSEA = .06, 95% CI [.04, .08], $\Delta \chi^2 = 23.09, \Delta df = 11$. The factor loadings were all statistically significant with standardized values exceeding .30, except for Items 5 (.30), 12 (.23), and 13 (.25), which obtained lower values than in the 13-item model (Table 2).

**Reliability and Validity**

As shown in Table 1, internal consistency of the SMFQ-P was good, with a Cronbach’s alpha coefficient of .83 for the total score, indicating a considerable homogeneity of the items measuring the same construct, depressive symptoms. Test-retest analysis was calculated in a sub-sample (n = 42). Results indicated an adequate reliability.
of the measure over time. The attrition analyses revealed that the sample of participants included in the test-retest was equivalent to those not included ($n = 139$) in sociodemographic variables, as well as in the SMFQ-P, SCAS-P and SDQ-P subscales ($p > .05$).

Convergent and divergent validity was explored by conducting bivariate correlations (Table 3). Convergent validity was supported by moderate correlations between the total score of the SMFQ-P and the SDQ Emotional symptoms subscale ($r = .60$), and between the total score of the SMFQ-P and the SCAS-P, which reached a lower magnitude ($r = .48$). The SDQ Prosocial behavior subscale had a lower and negative correlation with the SMFQ-P total score ($r = -.35$), thus supporting the divergent validity of the SMFQ-P.

### Table 2. Factor Loadings for the SMFQ-P

<table>
<thead>
<tr>
<th>Item</th>
<th>13-item</th>
<th>12-item</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. S/he felt miserable or unhappy</td>
<td>.77</td>
<td></td>
</tr>
<tr>
<td>2. S/he didn’t enjoy anything at all</td>
<td>.46</td>
<td>.50</td>
</tr>
<tr>
<td>3. S/he felt so tired that s/he just sat around and did nothing</td>
<td>.55</td>
<td>.56</td>
</tr>
<tr>
<td>4. S/he was very restless</td>
<td>.69</td>
<td>.68</td>
</tr>
<tr>
<td>5. S/he felt s/he was no good anymore</td>
<td>.31</td>
<td>.30</td>
</tr>
<tr>
<td>6. S/he cried a lot</td>
<td>.60</td>
<td>.63</td>
</tr>
<tr>
<td>7. S/he found it hard to think properly or concentrate</td>
<td>.57</td>
<td>.57</td>
</tr>
<tr>
<td>8. S/he hated him/herself</td>
<td>.46</td>
<td>.46</td>
</tr>
<tr>
<td>9. S/he felt s/he was a bad person</td>
<td>.35</td>
<td>.33</td>
</tr>
<tr>
<td>10. S/he felt lonely</td>
<td>.41</td>
<td>.43</td>
</tr>
<tr>
<td>11. S/he thought nobody really loved him/her</td>
<td>.54</td>
<td>.54</td>
</tr>
<tr>
<td>12. S/he thought s/he could never be as good as other kids</td>
<td>.24</td>
<td>.23</td>
</tr>
<tr>
<td>13. S/he felt s/he did everything wrong</td>
<td>.27</td>
<td>.25</td>
</tr>
</tbody>
</table>

### Table 3. Pearson correlations among the SMFQ-P and the SCAS-P, the emotional symptoms and the prosocial behavior subscales of SDQ-P

<table>
<thead>
<tr>
<th></th>
<th>SMFQ-P</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCAS-P</td>
<td></td>
</tr>
<tr>
<td>Total score</td>
<td>.48**</td>
</tr>
<tr>
<td>SDQ-P</td>
<td></td>
</tr>
<tr>
<td>Emotional symptoms</td>
<td>.60**</td>
</tr>
<tr>
<td>Prosocial behavior</td>
<td>-.35**</td>
</tr>
</tbody>
</table>

** Correlation is significant at the .01 level (2-tailed).

(ICC = .80) of the measure over time. The attrition analyses revealed that the sample of participants included in the test-retest was equivalent to those not included ($n = 139$) in sociodemographic variables, as well as in the SMFQ-P, SCAS-P and SDQ-P subscales ($p > .05$).

Convergent and divergent validity was explored by conducting bivariate correlations (Table 3). Convergent validity was supported by moderate correlations between the total score of the SMFQ-P and the SDQ Emotional symptoms subscale ($r = .60$), and between the total score of the SMFQ-P and the SCAS-P, which reached a lower magnitude ($r = .48$). The SDQ Prosocial behavior subscale had a lower and negative correlation with the SMFQ-P total score ($r = -.35$), thus supporting the divergent validity of the SMFQ-P.

### Discussion

The aim of the present study was to validate the SMFQ-P for Spanish-speaking children aged 6–8 years. The internal consistency of the scale was adequate ($\alpha = .83$) and similar to those reported by previous studies conducted in child and adolescent samples of other countries such as the United States ($\alpha = .84–.87$; Angold et al., 1995; Rhew et al., 2010), Norway ($\alpha = .83$; Karevold et al., 2017), Brazil ($\alpha = .87$; Sucupira et al., 2017), and Australia ($\alpha = .89$; Tennant et al., 2017). Additionally, test-retest results (ICC = .80) supported the scale as a consistent measure over time and are in line with previous findings (ICC = .79; Sucupira et al., 2017). Thus, these results showed that the Spanish version of the SMFQ-P is a reliable instrument to measure depressive symptoms. Furthermore, to our knowledge, test-retest analyses have scarcely been carried out for the SMFQ-P, so our results extend the literature supporting the stability of the scale over an eight-week period. Regarding Item 1 (“S/he felt miserable or unhappy”), the explanation of the low item-total correlations obtained (.13) and the results indicating a slight improvement of the internal consistency ($\alpha = .85$) if the item was removed remain unclear. However, the factor loading obtained for Item 1 (.77) was the highest in this study, indicating the considerable association between the item and the latent factor measured. In this regard, Angold et al. (1995) suggested the appropriateness of considering the item’s loading in the factor under study through factor analysis, rather than through item-total correlations, as this method allows excluding some of the error variance involved in the total score.

To examine the convergent and divergent validity, we computed correlations between the SMFQ-P, two measures of emotional problems including depression and anxiety (i.e., the SDQ Emotional symptoms subscale and the SCAS-P total score), and another variable measuring a non-emotional construct (i.e., the SDQ Prosocial behavior subscale). The correlation between the SMFQ-P and the SDQ Emotional symptoms was the strongest, whereas a lower but moderate positive correlation was also found between the SMFQ-P and the...
Besides, prior relationship with the underlying dimension of the scale (e.g., Carrillo et al., 2012; Zhao et al., 2012), and attending to the fact that the SCAS-P is not measuring depressive symptoms, in contrast to the SDQ Emotional symptoms subscale (Goodman, 2001). Furthermore, consistent with findings of other authors (Muris et al., 2003), a lower and negative correlation was found between our measure of depression and the SDQ Pro-social behavior subscale, supporting divergent validity. Low but significant correlations between the SDQ Pro-social behavior and another measure of emotional-related psychopathology, such as anxiety (i.e., SCAS), has also been previously considered evidence of divergent validity (e.g., Di Riso et al., 2013). These findings are also in line with previous research in the Spanish-speaking population that found positive significant correlations between measures of depression, anxiety and SDQ Emotional symptoms, as well as lower (positive or negative) correlations between the SDQ Pro-social behavior subscale and emotion-related measures (e.g., Carrillo et al., 2012; Orgilés, Rodríguez-Menchón et al., 2019; Park et al., 2016; Piqueras et al., 2017).

The results of the CFA supported the original one-factor model in Spanish children, showing that the Spanish version of the SMFQ-P measures the underlying factor of general depression. Our findings agree with previous research supporting the unidimensional structure of the scale (Angold et al., 1995; Karevold et al., 2017). Moreover, our results suggest that a 12-item model provides a slightly more reliable measure and a better fit than the 13-item model, by deleting Item 1. Although we do not have a clear explanation for this finding, it could be due to methodological reasons, for example because of the small sample size, so further research with larger and more representative samples should be addressed. It could also be due to the Spanish translation of the Item 1 “El/a/Ella/Él se sintió triste o infeliz” and a further revision of the translation may be advisable. However, as noted above, it is also worth mentioning that the factorial load of this item was the highest, even higher than that reported in the original study (.71) (Angold et al., 1995), which shows its strong relationship with the underlying dimension of the scale. Besides, prior findings support the importance and accuracy of this item, showing its high sensitivity and predictive value, and that it is also highly reported by youth (McKenzie et al., 2011). Furthermore, maintaining this item can be important as it seems clinically relevant, measuring one of the core symptoms of depression such as depressed or low mood (i.e., sadness or unhappiness) (Rey et al., 2015). Therefore, although good results are obtained in the CFA by removing Item 1 from the model, it should be noted that in the current study this item was dropped only on the basis of statistical criteria, while this method may increase reliability but reduce validity. Consequently, this decision should be based on other criteria and not only on statistical criteria, so that in this study the original 13-item model is supported and maintained. Moreover, the mean total score for depression obtained in our sample was 4.59 (SD = 4.27), which is within the range found in other SMFQ-P studies with non-diagnosed children (e.g., overall mean scores range 3.1 – 8.17; Rhew et al., 2010; Tennant et al., 2017). Also, no differences in gender and age were found. This is consistent with the literature reporting that the frequency of childhood depression is typically similar for boys and girls, and that adolescence is where higher depression rates are observed, especially for girls (Zahn-Waxler et al., 2008).

This study has some limitations that should be considered when interpreting the results. First, the sample of the study was small and was drawn from a specific region of Spain, which limits the generalization of the results. In addition, parents with higher education were over-represented in the study sample, and this may have influenced the results. The findings of this study should be addressed in larger samples and in different regions of Spain in future studies, and the psychometric properties of the SMFQ-P need to be confirmed with older children. However, it should be noted that the sample used is even slightly larger than that used in the original study (Angold et al., 1995). Second, this study did not analyze the validity of the measure using a diagnostic standard, which should be a focus of attention in future studies to determine the diagnostic accuracy of the measure. Third, the study relied on the reports of one of the parents of the children, so it would be interesting for future research to evaluate the responses of both parents to explore whether there are differences according to the informant.

In conclusion, despite the above limitations, this study provides for the first time initial empirical support for the use of the original 13-item version of the SMFQ-P with Spanish-speaking children, also extending the international support for the measure. The internal consistency, test-retest reliability, validity, and factor structure of the SMFQ-P were found to be adequate, suggesting its suitability for use as a brief screening tool to identify depressive symptoms in Spanish children. Finally, this study add support to the utility of this parent-report measure of depression for young children aged 6 to 8 years, thus addressing this gap in the
literature. The SMFQ-P is a valuable tool, as allows a quick assessment of core depressive symptoms (anhedonia and depressed mood), but also others, such as feelings of being unloved, self-hatred or worthlessness, which have been found to have a substantial predictive value in youths (McKenzie et al., 2011).

References


