



## The psychometric properties of the parenting scale for Spanish mothers with children aged between 2 and 7 years

Gemma López-Fernández, RNRM<sup>a,b,\*</sup>, Juana Gómez-Benito, PhD Psych<sup>a,c</sup>, Maite Barrios, PhD Psych<sup>a,c</sup>

<sup>a</sup> Department of Social Psychology and Quantitative Psychology, Faculty of Psychology, Universitat de Barcelona, Barcelona, Spain

<sup>b</sup> Department of Gynecology and Obstetrics, Corporació Sanitària Parc Taulí, Sabadell, Barcelona, Spain

<sup>c</sup> Group on Measurement Invariance and Analysis of Change (GEIMAC), Institute of Neurosciences, Universitat de Barcelona, Barcelona, Spain

### ARTICLE INFO

#### Article history:

Received 23 July 2021

Revised 22 October 2021

Accepted 5 November 2021

#### Keywords:

Mothers

Nurses

Parenting

Psychometrics

Statistical factor analysis

Validation study

### ABSTRACT

**Purpose:** This study aimed to evaluate the psychometric properties of the Spanish version of the Parenting Scale in a large sample of Spanish mothers.

**Design and methods:** A two-stage cross-sectional study of the adaptation and cultural validation of the Parenting Scale in a Spanish-speaking environment. In Stage I, the Parenting Scale was translated and back-translated and its semantic, linguistic and contextual equivalence was assessed. In Stage II, the Spanish-language version was validated after its application to 662 Spanish mothers with healthy children aged between 2 and 7 years. Several factor structure models of the Parenting Scale were compared by confirmatory factor analysis. Convergent validity, internal consistency and test-retest reliability were also examined.

**Results:** The model of Irvine et al. (1999) presented the best fit to our data. This model demonstrated adequate reliability (internal consistency and stability). The total score and each factor of the Parenting Scale correlated positively with perceived stress in mothers, difficulties in mother-child bonding and child hyperactivity, and negatively with child prosocial behavior.

**Conclusions:** The Spanish version of the Parenting Scale is a valid and reliable measure that can be used by healthcare professionals and scientists to assess dysfunctional parenting in Spanish mothers of children aged 2 to 7 years.

**Practice implications:** This study will allow the use of the Parenting Scale in epidemiological and cross-cultural studies in a variety of applied contexts. Additionally, health professionals who work with families in Spain will have access to a valid and reliable instrument for the assessment of mothers' parenting styles.

© 2021 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

### Background

Parenting style is conceptualized as a constellation of attitudes or a pattern of parental authority that are conveyed to the child (Baumrind, 1971). In 1971, Baumrind introduced three parenting styles as patterns of parental authority: authoritarian, authoritative, and permissive. These styles differ in terms of their levels of demandingness (i.e., the extent to which parents exert control, power, and supervision over their children, and set limits) and responsiveness (i.e., the extent to which parents show their children emotional warmth and acceptance, give support, and reason with them) (Merlin et al., 2013). Thus, parents with an authoritarian style are characterized by a low level of responsiveness and a high level of demandingness; those with an

authoritative style are characterized by a high level of both responsiveness and demandingness; and those with a permissive style are characterized by a low level of both these traits (Baumrind, 1991).

Studies exploring the influence of different parenting styles on the development of externalizing behaviors have reported that the authoritative style has positive effects on adaptation, promoting resilience, self-esteem and psychological adjustment in children, while authoritarian and permissive styles put children at risk of suffering externalizing behavioral problems (Rhoades & O'Leary, 2007; Ruiz-Hernández et al., 2019). The early detection of dysfunctional parenting styles and the introduction of measures to correct them are of great importance for children's mental health. Pediatric nurses, as health professionals in constant contact with families, are the professionals of reference for detecting dysfunctional parenting styles in early stages. To be able to do so, however, they need valid and reliable measures for assessing parenting styles.

The Parenting Scale (PS; Arnold et al., 1993) is widely used for this purpose (Pritchett et al., 2011). It consists of a 30-item self-report

\* Corresponding author at: Department of gynecology and obstetrics, Corporació Sanitària Parc Taulí, Parc del Taulí 1, 08208 Sabadell, Barcelona, Spain.  
E-mail address: [glopez@tauli.cat](mailto:glopez@tauli.cat) (G. López-Fernández).

scale and was developed to measure dysfunctional discipline styles associated with child externalizing behavior (Arnold et al., 1993). Parents report how they might respond to different discipline situations by choosing between an effective or ineffective parental response on a 7-point Likert scale. After reversing some items higher scores reflect a greater degree of dysfunctional parenting (Salari et al., 2012).

The first study published on the PS revealed a three-factor structure: laxness, overreactivity, and verbosity. Laxness reflects permissiveness and inconsistent discipline, overreactivity emotional and harsh discipline, and verbosity lengthy verbal responses to misbehavior. The instrument was validated observing parenting behavior and child misbehavior, distinguishing between mothers of children in clinical and non-clinical populations (Arnold et al., 1993). Since the publication of the PS, several studies have examined its dimensionality using exploratory and confirmatory approaches (Arney et al., 2008; Collett et al., 2001; Del Vecchio, Jerusalem, & Terjesen, 2017; Harvey et al., 2001; Irvine et al., 1999; Karazsia et al., 2008; Kliem et al., 2019; Prinzie et al., 2007; Reitman et al., 2001; Rhoades & O'Leary, 2007; Salari et al., 2012; Steele et al., 2005). However, only Arney et al. (2008) identified a factor structure that was similar to the original one reported by Arnold et al. (1993), and even then they reported low internal consistency and test-retest reliability for verbosity. The other studies did not support the original three-factor structure and proposed alternative factor structures of one (Del Vecchio, Jerusalem, & Terjesen, 2017), two (i.e., laxness and overreactivity; Collett et al., 2001; Harvey et al., 2001; Irvine et al., 1999; Kliem et al., 2019; Prinzie et al., 2007; Reitman et al., 2001) or three factors (i.e., laxness, overreactivity, and hostility; Rhoades & O'Leary, 2007).

Several studies focusing on Spanish populations have used the PS (Grau Sevilla, 2007; Miranda et al., 2009), and have highlighted the need for an instrument of this type in the Spanish context. At present, however, there are no studies reporting on the translation of the PS from English into Spanish or on the psychometric properties of the Spanish version. The lack of information on the translation and adaptation of the PS in the studies using Spanish populations calls into question the validity of their findings (Hambleton et al., 2005). The International Test Commission Guidelines for Translating and Adapting Test states that it cannot be assumed that a translated version has the same psychometric qualities as the standardized version in its primary language (International Test Commission, 2017). Consequently, the use of instruments that are merely translated, without sufficient verification of the results, does not guarantee reliable results in intercultural research, since equivalence between the translated and original versions of the instrument cannot be taken for granted. Thus, the purpose of this study was two-fold: first, to translate and adapt the PS into Spanish and, second, to compare the different factor structure models of the PS in the Spanish population and to test its psychometric properties (reliability and convergent validity) based on the factor structure that best fits the data.

## Stage I: translation and cultural adaptation

### Purpose

The aim of Stage I was to translate and culturally adapt the PS, for use with mothers of children aged between 2 and 7 years, to the Spanish-speaking context. The adaptation used a standardized method in order to guarantee semantic, linguistic and contextual equivalence with the original instrument.

### Method

#### Design

Cross-sectional study of cultural adaptation of instruments. The STROBE checklist (supplemental material) for cross-sectional studies was followed to enhance methodological rigour.

### Participants

At this stage of the study, 15 participants were recruited, the minimum recommended sample size for pilot tests in studies of cross-cultural validation of scales (Sousa & Rojjanasrirat, 2011). The participants, Spanish women of legal age with one or more children aged between 2 and 10, were recruited in November 2017 from the personal contacts of the researchers using convenience sampling. Information on the nature of the study was sent by email and all participants accepted the conditions of the study before they were recruited.

### Translation process

The PS was translated following the Brislin model (Brislin, 1970) for cross-cultural adaptation of research instruments. First, the original version in English was translated into Spanish independently by two of the authors of the present study (MBC and GLF), who are native speakers of Spanish, fluent in English and were previously unfamiliar with the PS. Second, the two Spanish versions of the PS were compared and reconciled into one forward translation. Third, another member of the team (JGB) reviewed the translation of the scale, resolved some queries and suggested some improvements, which were in turn evaluated and accepted by MBC and GLF. Fourth, a native English speaker fluent in Spanish translated this form of the PS back into English. Finally, the original instrument and the back-translated English version were compared. To guarantee that the conceptual meaning was captured, Professor O'Leary, author of the original instrument, was contacted. She approved all the items bar one, item 14, which was modified following her indications.

### Pilot testing

The revised forward translation was tested over two rounds between November 2017 and January 2018.

In the first round, nine women were asked to rate their level of comprehension of the instructions in the PS, an item response example, and each of the items using a five-point Likert scale, where zero indicated that the item was not understood at all and five indicated that the item was understood perfectly. Scores below five were considered inadequate and participants who rated an item lower than 5 points were requested to indicate why they thought that the item was problematic and how they would improve it. The test was performed online and was self-administered; there was no time limit and participants could verify their answers before responding.

The second round was held after incorporating the changes suggested by the participants in the first round. The nine original participants and six new participants rated their level of comprehension of the instructions, an item response example, and each item on the PS on a five-point Likert scale, using the same rating scale as before. Again, the test was online, self-administered, and with no time limit. Fig. 1 shows the methodological process followed.

### Ethical considerations

All procedures involving humans were performed in accordance with the World Medical Association Declaration of Helsinki Ethical Principles for Medical Research with Human Subjects and with national and institutional ethical guidelines. The study was approved by the Bioethics Commission of the University of Barcelona (IRB00003099). All electronic questionnaires were collected anonymously. The data were securely stored and used only for research purposes.

### Results

In the first round, the instructions and most of the items were scored 5 points by all nine participants. Only the example and items 23 and 27 were rated below five, each one by three participants. The participants suggested that the item response example would benefit from some further instructions concerning the seven-point scale and anchors; therefore, we added specific instructions in the example. In items 23 and 27, the participants had some difficulties with the anchors. Three

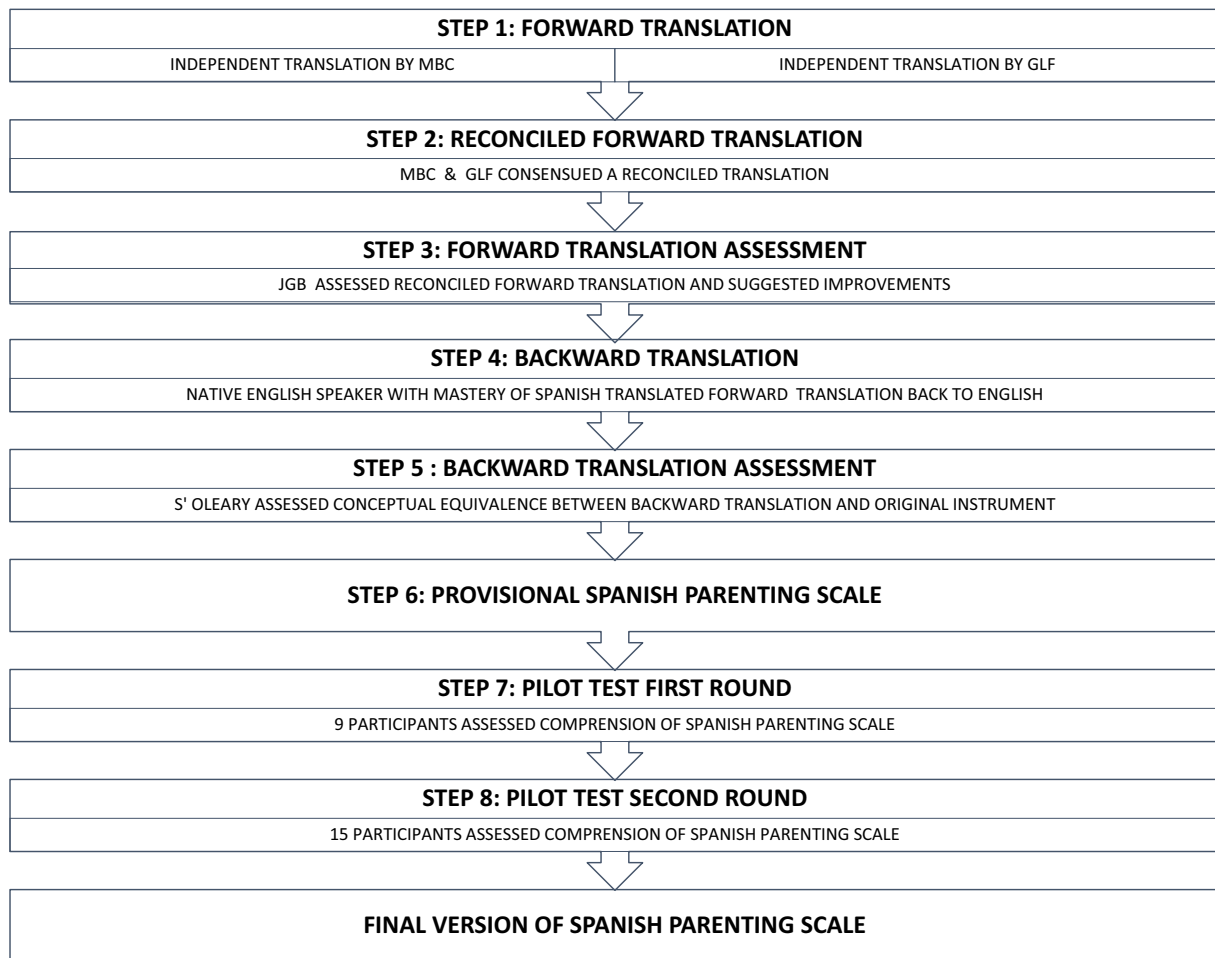


Fig. 1. Methodological process of Stage I.

participants found the second anchor in item 23 to be unclear, while three participants had the same problem with the first anchor in item 27. These were modified by adding some extra words to make them clearer, and the modifications were assessed and approved by Professor O'Leary. The final Spanish version is shown in Table 1.

In the second round, the instructions and items were scored 5 points by all the participants except for one. This participant suggested changing the rating scale to a yes or no answer format. However, our objective was not to create a new scale; further, given that the other participants had not expressed dissatisfaction with the scale format and the suggested change would lead to a loss of item variability and thereby weaken the scale, we considered the translation and adaptation of the PS to be complete.

## Stage II: psychometric validation

### Purpose

The aim of Stage II was to compare the different factor structure models of the PS and to test the psychometric properties of the structure that best fitted our data.

### Method

#### Design

Cross-sectional study of cultural adaptation of instruments. The STROBE checklist (supplemental material) for cross-sectional studies was applied to enhance the methodological rigour.

### Participants

The participants were recruited by convenience sampling (i.e., the snowball approach) through school family associations. The minimum sample size ( $n = 300$ ) was determined considering the numerical ratio between the cases and the number of items (10:1; Kline, 2016). The inclusion criteria were: Spanish women of legal age with one or more children between 2 and 7 years of age, having access to the internet, and being able to read and write in Spanish.

A total of 673 mothers participated, but only 662 provided valid and complete answers to all questionnaires and were therefore included in the data analyses. Eight were excluded for not meeting the study criteria and three for leaving some items unanswered. The mothers were not financially compensated, but they received a personal report on their results. Table 2 presents the demographic information for the Stage II participants.

### Measures

The PS was administered alongside other instruments measuring psychological stress, child behavior, and maternal bonding.

The *Perceived Stress Scale* (PSS) (Cohen et al., 1983) in its Spanish version (Remor, 2006) was used to assess the stress perceived by the participants. The Spanish version of PSS is 10-item scale that measures the degree to which situations are valued as stressors. Items are rated on a 5-point response scale (0 = never, 1 = almost never, 2 = once in a while, 3 = often and 4 = very often). The total score is obtained by reversing the score of items 4, 5, 7 and 8. A higher score indicates a higher level of perceived stress. The Spanish version (Remor, 2006)

**Table 1**  
Spanish version of the Parenting Scale.

ESCALA DE CRIANZA

Instrucciones:

En algún momento, todos los niños se portan mal o hacen cosas que pueden ser peligrosas, o que a los padres no les gustan. Algunos ejemplos son: Pegar a otras personas, lloriquear, tirar comida, olvidar los deberes del colegio, no recoger los juguetes, mentir, tener rabietas, negarse a ir a la cama, querer comer galletas antes de cenar, cruzar la calle sin mirar, replicar o volver tarde a casa. Los padres tienen muchas maneras diferentes o estilos de tratar este tipo de problemas. A continuación se presentan una serie de ítems que describen diferentes estilos de crianza. En cada ítem, marque el número que mejor describa su estilo de crianza con su hijo durante los dos últimos meses.

Ítem de ejemplo:

A la hora de comer...

Dejo que mi hijo decida qué cantidad quiere comer / Soy yo quien decide qué cantidad debe comer mi hijo

Las puntuaciones de los extremos (1 y 7) indican que usted se siente totalmente identificada con el extremo elegido, en caso de sentirse parcialmente identificada deberá elegir una puntuación intermedia, en función de si se siente más identificada con un extremo u otro.

1- Cuando mi hijo se porta mal...

Actúo de inmediato. / Actúo más tarde.

2- Antes de actuar ante un problema...

Le doy varios avisos o advertencias a mi hijo. / Le doy únicamente un aviso o advertencia.

3- Cuando estoy alterado o estresado...

Soy exigente con mi hijo y le estoy encima. / No soy más exigente de lo habitual.

4- Cuando le digo a mi hijo que no haga algo ...

Soy breve. / Me extiendo hablando.

5- Cuando mi hijo me fastidia...

Puedo ignorar que me está fastidiando. / No puedo ignorar que me está fastidiando.

6- Cuando mi hijo se porta mal...

Habitualmente entro en una larga discusión con mi hijo. / No entro a discutir.

7- Amenazo con hacer cosas que ...

Estoy seguro puedo cumplir. / Sé que en realidad no haré.

8- Soy el tipo de madre que...

Pongo límites en lo que le permito hacer mi hijo. / Dejo que mi hijo haga lo que quiera.

9- Cuando mi hijo se porta mal ...

Le doy un sermón. / Trato el tema de forma breve y directa.

10- Cuando mi hijo se porta mal...

Le levanto la voz o le grito. / Le hablo de forma calmada.

11- Si el decir "no", no tiene un efecto inmediato...

Tomo otro tipo de medida. / Sigo hablándole y trato de convencerle.

12- Cuando quiero que mi hijo deje de hacer algo...

Le digo firmemente que pare. / Le intento convencer o le ruego que pare.

13- Cuando mi hijo está fuera de mi vista ...

Generalmente no sé lo que estará haciendo/ Siempre tengo una idea bastante clara de lo que estará haciendo.

14- Después de que haya habido un problema con mi hijo/a...

A menudo me siento resentida durante un tiempo./Las cosas vuelven a la normalidad rápidamente.

15- Cuando no estamos en casa...

Trato a mi hijo igual que lo hago en casa. / Dejo que se salga con la suya mucho más.

16- Cuando mi hijo/a hace algo que no me gusta...

Hago algo al respecto cada vez que ocurre. / A menudo lo dejo pasar.

17- Cuando hay un problema con mi hijo...

Las cosas me superan y hago cosas que no quiero hacer / Las cosas no se me van de las manos.

18- Cuando mi hijo se porta mal le agarro, le doy un cachete, un azote o le pego...

Nunca o casi nunca. / La mayoría de las veces.

19- Cuando mi hijo no hace lo que le pido...

A menudo lo dejo pasar o lo termino haciendo yo. / Tomo otro tipo de medida.

20- Cuando doy una advertencia o hago una amenaza justa...

A menudo no la cumplo. / Siempre cumplo con lo que he dicho.

21- Si decir "no" no funciona ...

Tomo algún otro tipo de medida. / Le ofrezco algo que le guste para que se porte bien.

23- Cuando mi hijo se porta mal...

Exijo que me diga porqué lo ha hecho. / Le digo "que pare de portarse mal" o tomo otro tipo de medida.

24- Si mi hijo se porta mal y luego se arrepiente...

Me ocupo del problema como lo suelo hacer. / Lo dejo pasar por esta vez.

25- Cuando mi hijo se porta mal...

Casi nunca digo palabrotas. / Casi siempre digo palabrotas.

26- Cuando le digo a mi hijo que no puede hacer algo...

Le permito que lo haga de todas formas. / Me mantengo firme en lo que he dicho.

27- Cuando tengo que ocuparme de un problema con mi hijo.

Le digo que lamento tener que reñirle o castigarle. / No le digo que lo lamento.

28- Cuando mi hijo hace algo que no me gusta, le insulto o le digo cosas crueles.

Nunca o casi nunca. / La mayoría de las veces.

29- Si mi hijo replica o se queja cuando me encargo de un problema...

Ignoro sus quejas y me mantengo firme en lo que he dicho. / Le doy un sermón sobre no quejarse.

30- Si mi hijo se enfada cuando digo "no" ...

Me echo atrás y cedo ante mi hijo./ Me mantengo firme en lo que he dicho.

**Table 2**  
Sociodemographic characteristics of the sample.

	%	M	SD	min – max
Maternal age		38.13	4.23	24 to 51
Maternal educational level				
Primary education	3.1			
Secondary education	22.5			
University education	74.5			
Maternal employment status				
Employed	72.2			
Unemployed	13.1			
Businesswoman (Entrepreneur)	9.9			
Retired due to a disability	0.2			
Retired due to preference	4.5			
Family life				
Both parents and children live together	89.7			
The mother and her children live with a new partner	1.1			
The mother and children live alone	7.9			
Another situation	1.4			
No. of children in household		1.81	0.74	1 to 5
Gender of target child				
Male child	56.2			
Female child	43.7			
Age of target child		4.36	1.93	2 to 7

Note. Abbreviations: M, mean; SD, standard deviation, min; minimum value; max, maximum value.

reported adequate internal consistency; in this study we reported a Cronbach's alpha coefficient of 0.84.

The *Strengths and Difficulties Questionnaire (SDQ)* (Goodman, 1997) is a 25-item scale for evaluating emotional and behavioral disorders in children and adolescents. Four scales measure problem behaviors, under the headings: Emotional symptoms, Conduct problems, Hyperactivity and Problems with peers, while the fifth scale, the Prosocial behavior scale, refers to positive behaviors. Each scale consists of five items rated on a 3-point scale (0 = not true, 1 = somewhat true, and 2 = certainly true). Higher scores on the four scales that measure behavior problems indicate more difficulties, while higher scores on the prosocial behavior scale indicate more strengths. This scale has been translated into and validated in many languages, including Spanish. Spanish versions for children aged between two and four (Ezpeleta et al., 2013), and for children over four (Gómez-Beneyto et al., 2013) were used in this study. In our sample, the Cronbach's alpha coefficients for the five scales were: Emotional symptoms ( $\alpha = 0.66$  for children between two and four years;  $\alpha = 0.74$  for children over four years); Conduct problems ( $\alpha = 0.72$  for children between two and four years;  $\alpha = 0.63$  for children over four years); Hyperactivity ( $\alpha = 0.83$  for children between two and four years;  $\alpha = 0.84$  for children over four years); Peer problems ( $\alpha = 0.67$  for children between two and four years;  $\alpha = 0.73$  for children over four years); and Prosocial behavior ( $\alpha = 0.75$  for children between two and four years;  $\alpha = 0.71$  for children over four years).

The *Mother-to-Infant Bonding Scale (MIBS)* (Taylor et al., 2005) is an eight-item scale designed to assess the feelings of a mother toward her child. Items are scored on a 4-point Likert scale (0 = not at all, 3 = very much) according to the intensity of the maternal emotion toward the child, with higher scores indicating more bonding challenges. The Spanish version of the scale has been shown to have good internal consistency (Palacios-Hernández, Subirà Álvarez, & García-Esteve, 2015). In our sample, Cronbach's alpha was 0.74.

#### Data collection

Between February and July 2018, the study was disseminated among school family associations throughout Spain. The study was only available online and remained open until September 2018. Participants could contact the research team via email if they had any questions regarding the research.

A subsample of 354 participants agreed to participate in the retests and were re-administered the PS between four and six weeks after the initial assessment.

Informed consent was obtained from all participants. Previously they were informed of the nature of the research and the study's objectives, and it was made clear that participation was voluntary and that all data would always remain confidential.

#### Ethical considerations

Ethical considerations had been explained in Stage I.

#### Data analysis

Confirmatory factor analyses (CFA) were used to investigate the validity of the original three-factor model (Arnold et al., 1993), as well as the two-factor models proposed by Irvine et al. (1999), Collet et al. (2001), Reitman et al. (2001), Harvey et al. (2001), Prinzie et al. (2007), Arney et al. (2008), and Kliem et al. (2019) and the one-factor and three-factor models proposed by Del Vecchio, Jerusalem, & Terjesen, 2017 and Rhoades and O'Leary (2007) respectively. All analyses were performed with the IBM SPSS AMOS 25 software, using maximum likelihood estimation with robust standard errors.

To assess the goodness of fit of the proposed models, we applied the following criteria: (1) the root mean square error of approximation (RMSEA) with its 90% confidence interval; (2) the goodness-of-fit index (GFI); (3) the adjusted goodness-of-fit index (AGFI); (4) the comparative fit index (CFI) and (5) the non-normed fit index (NNFI). The first three criteria were used to assess absolute model fit, while the last two criteria were used to compare the fitted model to the "null" model (relative fit). Additionally, the Akaike information criterion (AIC) and  $\chi^2/df$  were calculated. The AIC, which is based on the maximum likelihood estimate and the number of parameters (independent variables) in the model, is one of the most widely used criteria for model selection. The lower the AIC values, the better the relative fit of the model to the data.

RMSEA values <0.05 represent a good fit, values between 0.05 and 0.08 an acceptable fit, values between 0.08 and 0.10 a marginal fit and values >0.10 an unacceptable fit. Regarding CFI, NNFI, GFI, and AGFI, values >0.90 suggest an adequate fit, while values >0.95 indicate a good fit.  $\chi^2/df$  ratios close to or less than two also indicate good fit, while those or three or below represent an acceptable fit (Schermelleh-Engel et al., 2003).

The reliability of the subscales in the model that best fitted our data was examined by calculating Cronbach's alpha coefficients for internal consistency and Pearson correlation coefficients between test-retest scores over a four to six-week interval. Finally, the association of the adapted PS with the other measures was also examined by calculating Pearson correlation coefficients.

#### Results

##### Factor structure

As shown in Table 3, the  $\chi^2$  value was significant and the ratio relating its degree of freedom ( $\chi^2/df$ ) was greater than three for all the models, indicating a questionable fit. RMSEA values were also relatively similar for most of the models (0.060 to 0.079), indicating a reasonable fit. The values of the other criteria used were less consistent across the different models: NNFI (0.562 to 0.903), CFI (0.634 to 0.922), GFI (0.824 to 0.961), AGFI (0.757 to 0.936) and AIC (143.493 to 1539.898).

The two models that provided the best fit for our data were the one proposed by Irvine et al. (1999) and the one proposed by Reitman et al. (2001). Although the  $\chi^2/df$  ratio was greater than three and the RMSEA, AIC, AGFI and GFI values were acceptable for both models, only the model of Irvine et al. (1999) had acceptable NNFI and CFI values ( $\chi^2/df = 3.37$ ; NNFI = 0.903; CFI = 0.922; GFI = 0.957; AGFI = 0.936; and AIC = 366.27). Consequently, we considered that the model that provided the best fit for our data was the model identified by

**Table 3**  
Goodness-of-fit indices for confirmatory factor analysis.

Model	Factors	$\chi^2$ (df)	$\chi^2/df$	NNFI	CFI	GFI	AGFI	AIC	RMSEA (90% CI)
Arnold et al. (1993)	3	1239.976 (294)**	4.22	0.700	0.729	0.864	0.838	1353.962	0.070 (0.066–0.074)
Irvine et al. (1999)	2	178.889 (53)**	3.37	0.903	0.922	0.957	0.936	228.889	0.060 (0.050–0.070)
Collet et al. (2001)	2	1347.631 (298)**	4.52	0.673	0.700	0.849	0.823	1453.631	0.073 (0.069–0.077)
Reitman et al. (2001)	2	139.847 (34)**	4.11	0.846	0.883	0.961	0.936	181.847	0.069 (0.057–0.081)
Harvey et al. (2001)	2	1280.196 (251)**	5.10	0.662	0.692	0.846	0.816	1378.196	0.079 (0.075–0.083)
Rhoades and O’Leary (2007)	3	245.991 (63)**	3.90	0.833	0.865	0.946	0.922	301.991	0.066 (0.058–0.075)
Prinzie et al. (2007)	2	813.983 (169)**	4.81	0.749	0.776	0.884	0.856	895.983	0.076 (0.071–0.081)
Arney et al. (2008)	2	1419.898 (346)**	4.10	0.678	0.705	0.860	0.835	1539.898	0.069 (0.065–0.072)
Del Vecchio et al. (2017)	1	475.272 (54)**	8.80	0.569	0.647	0.878	0.824	532.272	0.109 (0.100–0.118)
Kliem et al. (2019)	2	119.493 (19)**	6.20	0.846	0.895	0.956	0.916	143.493	0.089 (0.074–0.105)

Note. Abbreviations:  $\chi^2$ , Satorra-Bentler  $\chi^2$ ; df, degrees of freedom;  $\chi^2/df$ , Satorra-Bentler  $\chi^2$ /degrees of freedom ratio; NNFI, Non-Normed Fit Index; CFI, Comparative Fit Index; GFI, Goodness of fit index; AGFI, Adjusted goodness of fit index; AIC, Akaike information criterion; RMSEA, Root mean square error of approximation; CI, Confidence interval; \*\*  $p < .001$ .

Irvine et al. (1999). Fig. 2 displays the path diagram of the model of Irvine et al. (1999).

laxness (0.752), overreactivity (0.759), and total score (0.764) also showed adequate temporal stability.

*Internal consistency and test-retest reliability*

Based on the two-factor model of Irvine et al. (1999), Cronbach’s alpha coefficient and test-retest reliability were calculated. Laxness, overreactivity, and total score showed acceptable internal consistency (0.726, 0.726, and 0.738, respectively). Regarding test-retest reliability,

*Relationship with other measures*

Table 4 presents the Pearson correlations between the PS and the other measures of psychological stress, child behavior and maternal bonding. Correlations are given for laxness and overreactivity, as well as for the total score.

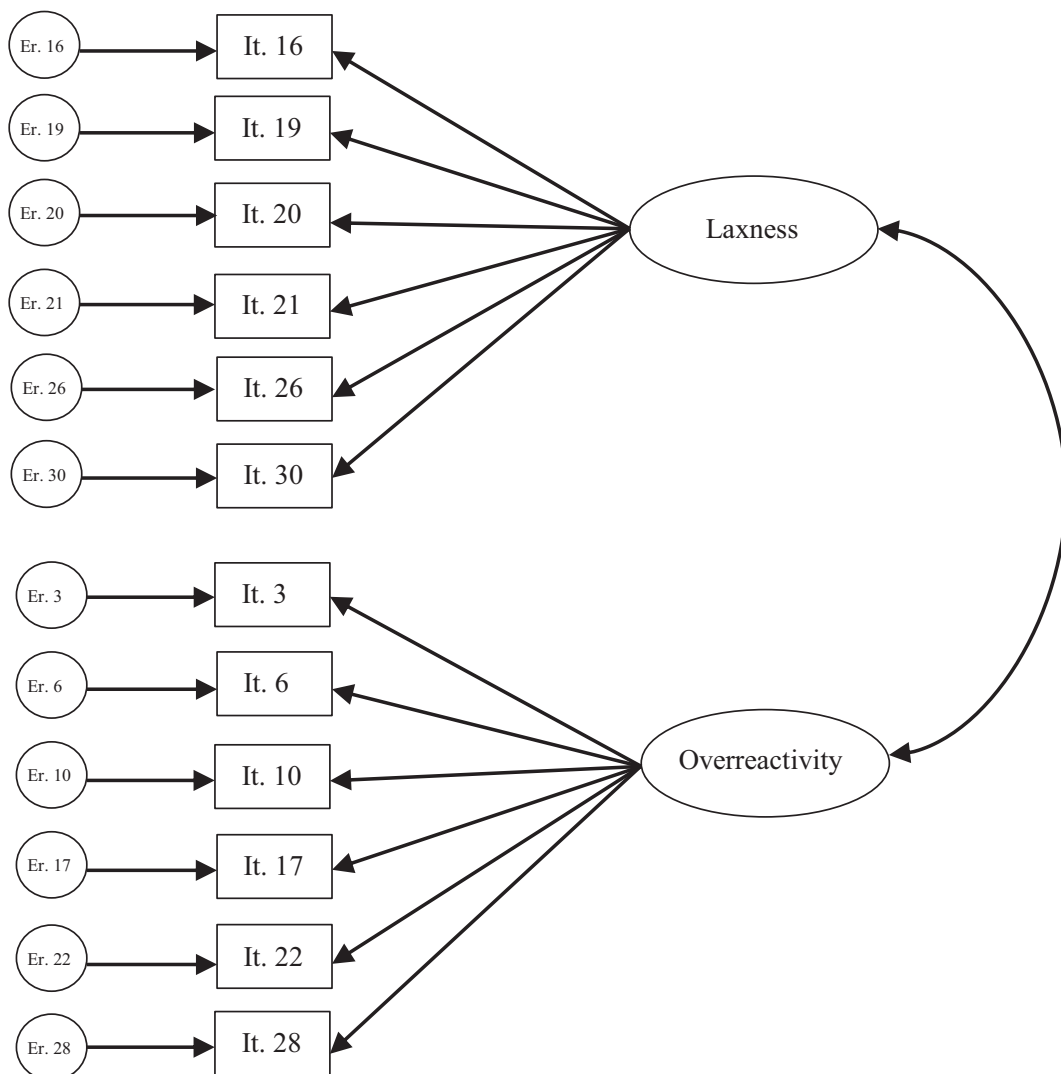


Fig. 2. Path diagram of the model of Irvine et al. (1999).

**Table 4**  
Pearson correlation coefficients between the Parenting Scale and other measures.

	Laxness	Overreactivity	Total Score
PSS	0.188**	0.311**	0.325**
MIBS	0.311**	0.438**	0.389**
<b>SDQ 2–4 years</b>			
Conduct symptoms	0.093	0.342**	0.284**
Hyperactivity symptoms	0.121*	0.221**	0.221**
Emotional symptoms	0.005	0.164**	0.112*
Peer symptoms	0.038	−0.064	−0.019
Prosocial behavior	−0.196**	−0.143**	−0.215**
<b>SDQ 5–7 years</b>			
Conduct symptoms	0.074	0.286*	0.237**
Hyperactivity symptoms	−0.002	0.191*	0.127**
Emotional symptoms	0.089	0.092	0.115*
Peer symptoms	0.078	0.068	0.093
Prosocial behavior	−0.128**	−0.238**	−0.236**

Note. Abbreviations: PSS, Perceived Stress Scale; MIBS, Mother-to-Infant Bonding Scale; SDQ, Strengths and Difficulties Questionnaire. \*\*  $p < .01$ , \*  $p < .05$ .

All correlations of the PS were in the expected direction: positive with child externalizing and internalizing behavior, parental stress, and difficulties in mother-child bonding, and negative with child prosocial behavior.

Overreactivity and total score showed a moderate correlation with perceived stress in mothers and mother-infant bonding difficulties. Correlations with externalizing problems (i.e., conduct and hyperactivity symptoms) were mainly small (ranging from 0.127 to 0.342) and most correlations with child internalizing problems (i.e., emotional and peer symptoms) were non-significant or very small (ranging from 0.107 to 0.164). Additionally, overreactivity, laxness and total score showed a small and negative correlation with child prosocial behavior (ranging from −0.128 to −0.238).

For laxness, the correlations were small with perceived stress in mothers, mother-infant bonding difficulties, and child prosocial behavior. Correlations with child externalizing and internalizing behavior were non-significant, except for a small correlation with hyperactivity symptoms in children aged between two and four years.

## Discussion

In this paper, we report the process of translating and culturally adapting into Spanish one of the most commonly used parenting scales. The scale was translated using a method that is widely applied in the adaptation of assessment instruments in psychological research with families and children (Brislin, 1970). With the growing interest in intercultural studies, it is important to have reliable and valid measures that can be used in different languages and countries (Hambleton et al., 2005). The use of a standardized method in the translation and adaptation process is crucial to ensure the quality of the translation and equivalence between the adapted scale and the original instrument (International Test Comision, 2017). Studies that do not use a well-translated instrument adapted to the culture jeopardize the quality of their findings.

There were some difficulties when translating some of the items. The PS is full of English idioms and sometimes it was difficult to find equivalent colloquial expressions in Spanish. Eventually, however, we achieved an acceptable translation of all the items; when the conceptual equivalence of the original instrument and the back translation was assessed, only minor modifications were needed.

The need for additional instructions in the item response example of the Spanish version may have been due to the participants' unfamiliarity with the response format. The results of the second round of the pilot test showed that the level of understanding of the item response example had improved after the addition of an explanation of the seven-point scale and anchors. Two items (23 and 27) needed to have one of their anchors modified, since the participants found them ambiguous. These anchors were improved by adding information to make them more

specific, without affecting their semantic meaning. All the steps followed in Stage I ensured semantic, linguistic and contextual equivalence between the original instrument and the final Spanish version, guaranteeing appropriate content validity of the adapted version.

The next step in the adaptation of the PS was to test it in a large sample and provide evidence of its psychometric properties. Thus, in Stage II, the factor structure of the scale was assessed in a large sample of Spanish mothers, bearing in mind the various factor models proposed by different groups of researchers.

In agreement with previous studies (Collett et al., 2001; Del Vecchio, Jerusalmi, & Terjesen, 2017; Harvey et al., 2001; Irvine et al., 1999; Karazsia et al., 2008; Kliem et al., 2019; Prinzie et al., 2007; Reitman et al., 2001; Rhoades & O'Leary, 2007; Salari et al., 2012), our data did not confirm the original three-factor structure proposed by Arnold et al. (1993). A subsequent CFA showed that the model reported by Irvine et al. (1999) provided the best fit in our sample. This model consists of two factors, overreactivity and laxness, which reproduce the original overreactivity and laxness scales of Arnold et al. (1993) and conforms to the constructs of authoritarian and permissive parenting styles described by Baumrind (1991). However this model removes the verbosity scale described in the original study, on the grounds that this construct rarely appears in the literature on deficient parental interaction in parents of adolescents. Although in our study the target population does not have children in this age group, the model of Irvine et al. (1999) presented better fit indices to our population than the model of Arnold et al. (1993). This may be due to differences in the composition of the samples, since our study was carried out on a community sample and the study of Arnold et al. (1993) included clinical and non-clinical populations. This hypothesis is consistent with the findings of Rhoades and O'Leary (2007) and Salari et al. (2012) who tested (in community samples) the fit indices for the models identified in earlier studies and reported that the model of Irvine et al. (1999) produced the best fit when focusing only on mothers. By contrast, in mixed samples of parents, Rhoades and O'Leary (2007) and Karazsia et al. (2008) found that the best fit is produced in the model proposed by Reitman et al. (2001). In our study, the two-factor model proposed by Reitman et al. (2001) showed an adequate fit for most of the criteria studied; however, the NNFI and CFI values did not. This lack of fit of the model proposed by Reitman et al. (2001) in our study could have been due to differences in the composition of the samples. The model of Reitman et al. (2001) was originally tested in a mixed sample of mothers and fathers, while our sample was composed exclusively of mothers. This could also explain the better fit of our data with the model of Irvine et al. (1999), which was originally tested in a sample composed mostly of women.

The present study also provides evidence for the scale's internal consistency, test-retest reliability and convergent validity with measures of child behavior problems. The values for internal consistency and test-retest reliability obtained by Irvine et al. (1999) were slightly higher than ours. Those authors also observed moderate correlations between child externalizing behavior and overreactivity and total score, as well as between child internalizing behavior and overreactivity. Furthermore, they reported small correlations between child externalizing behavior and laxness, as well as between child internalizing behavior and laxness and total score. Although these correlations were in the same direction as those found in our data, their values were slightly higher than ours, perhaps due to differences in the ages of the children and in the instruments used to assess child behavior. In the study of Irvine et al. (1999), the children's mean age was 12.2 years, while our study included children aged between two and seven years. Moreover, Irvine et al. (1999) used the Child Behavior Checklist (Achenbach, 1991) whereas we applied the SDQ. Interestingly, Salari et al. (2012), who also used the SDQ to measure child internalizing behavior (i.e., emotional symptoms), found small correlations with total score. Those researchers also found small correlations between total score, overreactivity, laxness and measures of parental stress, which was congruent with our findings.

## Implications for practice

This study will allow the use of the PS in epidemiological and cross-cultural studies in a variety of applied contexts, such as clinical and educational settings. Additionally, for the first time in Spain, pediatric nurses and other health professionals who work with families will have access to a valid and reliable instrument that allows an assessment of mothers' parenting styles and of the quality of interventions and programs implemented aimed at correcting disruptive parenting styles. The importance of maternal parenting style in children's behavioral development makes the early detection of disruptive parenting styles absolutely necessary in order to provide support to families. This scale may be helpful in identifying the maternal skills that most need to be targeted and addressed by individual mothers. In addition, having knowledge of mothers' parenting styles will be useful for pediatric nurses whose main responsibility is to provide direct care for children and to create alliances with parents in order to promote their children's health (Yoo & Cho, 2020).

## Limitations

This study had a number of limitations. First, the sample was recruited by convenience sampling and was not representative of mothers in the general population. Mothers with a lower education level were poorly represented in this study, maybe because the participants were self-recruited and study participation required online access and use of electronic devices. Second, fathers were not included in the study; it should not be assumed that the same psychometric properties would be observed with fathers as respondents. Third, all measures (i.e., parenting styles, attachment, stress and child behavior) were assessed through self-report responses and were not observed directly. Fourth, although the PS correlated moderately with the other measures, effect sizes were small.

Despite these limitations, the study has several noteworthy strengths. The first is the large sample size. Second, the study confirms and validates the factor structure proposed by Irvine et al. (1999), in an age group different from the one for which it was designed and validated. As a result, the study offers health professionals not only a valid and reliable instrument to evaluate disruptive parenting styles in mothers with children between two and seven years old, but also a more agile (and shorter) instrument than the one originally designed by Arnold et al. (1993). Third, the Spanish validation of the PS will allow its use in multiple countries, facilitating cross-cultural studies. According to the 2021 annual report of the Cervantes Institute, Spanish is spoken in more than 20 countries and there are more than 591 million Spanish speakers in the world (7.5% of the world's population), making Spanish the second most spoken native language in the world after Mandarin (Instituto Cervantes, 2021). This, coupled with the fact that the PS is an easy-to-apply and reliable instrument for evaluating dysfunctional parenting styles, makes this cultural adaptation necessary.

Additional studies that apply random sampling, including parents of both sexes and children from both clinical and non-clinical populations and of different age ranges, are now recommended.

## Conclusion

The Spanish adaptation of the PS described in the present study shows reliable and valid scores for assessing dysfunctional parenting styles in mothers of toddlers and elementary school-aged children. Its adaptation to the Spanish-speaking environment can be considered to be successful.

## Compliance with ethical standards

### Funding

This work was supported by the Agency for the Management of University and Research Grants of the Government of Catalonia [grant

2017SGR1681]. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

### Conflict of interest

The authors declare that they have no conflict of interest.

### Author contributions

All authors have agreed on the final version and meet at least one of the following criteria (recommended by the ICMJE): 1- substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data; 2- drafting the article or revising it critically for important intellectual content.

### Data availability statement

The data that support the findings of this study are available from the corresponding author on reasonable request.

## References

- Achenbach, T. (1991). *Manual for the child behavior checklist/4–18 and 1991 profile*. Burlington, VT, USA: University of Vermont, Department of Psychiatry.
- Arney, F., Rogers, H., Baghurst, P., Sawyer, M., & Prior, M. (2008). The reliability and validity of the parenting scale for Australian mothers of preschool-aged children. *Australian Journal of Psychology*, 60(1), 44–52. <https://doi.org/10.1080/00049530701458076>.
- Arnold, D. S., O'Leary, S. G., Wolff, L. S., & Acker, M. M. (1993). The parenting scale: A measure of dysfunctional parenting in discipline situations. *Psychological Assessment*, 5(2), 137–144. <https://doi.org/10.1037/1040-3590.5.2.137>.
- Baumrind, D. (1971). Current patterns of parental authority. *Developmental Psychology*, 4 (1, Pt.2), 1–103. <https://doi.org/10.1037/h0030372>.
- Baumrind, D. (1991). The influence of parenting style on adolescent competence and substance use. *The Journal of Early Adolescence*, 11(1), 56–95. <https://doi.org/10.1177/02724316911111004>.
- Brislin, R. W. (1970). Back-translation for cross-cultural research. *Journal of Cross-Cultural Psychology*, 1, 185–216.
- Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior*, 24(4), 385–396. <https://doi.org/10.2307/2136404>.
- Collett, B. R., Gimpel, G. A., Greenson, J. N., & Gunderson, T. L. (2001). Assessment of discipline styles among parents of preschool through school-age children. *Journal of Psychopathology and Behavioral Assessment*, 23(3), 163–170.
- Del Vecchio, T., Jerusalem, D., & Terjesen, M. D. (2017). Psychometric characteristics of the parenting scale in a Vietnamese sample. *International Journal of Psychology*, 52(6), 482–490. <https://doi.org/10.1002/ijop.12242>.
- Ezpeleta, L., Granero, R., La, O., De, N., Penelo, E., & Domènech, J. M. (2013). Psychometric properties of the strengths and difficulties questionnaire 3–4 in 3-year-old preschoolers. *Comprehensive Psychiatry*, 54(3), 282–291. <https://doi.org/10.1016/j.comppsy.2012.07.009>.
- Gómez-Beneyto, M., Nolasco, A., Moncho, J., Pereyra-Zamora, P., Tamayo-Fonseca, N., Munarriz, M., ... Girón, M. (2013). Psychometric behaviour of the strengths and difficulties questionnaire (SDQ) in the Spanish national health survey 2006. *BMC Psychiatry*, 13, 95. <https://doi.org/10.1186/1471-244X-13-95>.
- Goodman, R. (1997). The strengths and difficulties questionnaire: A research note. *Journal of Child Psychology and Psychiatry*, 38(5), 581–586. Retrieved from <https://doi.org/10.1111/j.1469-7610.1997.tb01545.x>.
- Grau Sevilla, M. D. (2007). *Análisis del contexto familiar en niños con TDAH*. Universitat de València. <http://www.tesisenred.net/handle/10803/10230>.
- Hambleton, R. K., Spielberger, C. D., & Merenda, P. F. (2005). Issues, designs, and technical guidelines for adapting tests into multiple languages and cultures. *Adapting educational and psychological tests for cross-cultural assessment* (pp. 15–50). New York: Psychology Press. <https://doi.org/10.4324/9781410611758-6>.
- Harvey, E., Danforth, J. S., Ulaszek, W. R., & Eberhardt, T. L. (2001). Validity of the parenting scale for parents of children with attention-deficit/hyperactivity disorder. *Behaviour Research and Therapy*, 39(6), 731–743. [https://doi.org/10.1016/S0005-7967\(00\)00052-8](https://doi.org/10.1016/S0005-7967(00)00052-8).
- International Test Commission (2017). *The ITC Guidelines for Translating and Adapting Tests* (Second). Retrieved from [www.InTestCom.org](http://www.InTestCom.org).
- Irvine, A. B., Biglan, A., Smolkowski, K., & Ary, D. V. (1999). The value of the parenting scale for measuring the discipline practices of parents of middle school children. *Behaviour Research and Therapy*, 37(2), 127–142. [https://doi.org/10.1016/S0005-7967\(98\)00114-4](https://doi.org/10.1016/S0005-7967(98)00114-4).
- Karazsia, B. T., Van Dulmen, M. H. M., & Wildman, B. G. (2008). Confirmatory factor analysis of Arnold et al.'s Parenting Scale across race, age, and sex. *Journal of Child and Family Studies*, 17(4), 500–516. <https://doi.org/10.1007/s10826-007-9170-1>.
- Kliem, S., Lohmann, A., Mößle, T., Foran, H. M., Hahlweg, K., Zenger, M., & Brähler, E. (2019). Development and validation of a parenting scale short form (PS-8) in a



- representative population sample. *Journal of Child and Family Studies*, 28(1), 30–41. <https://doi.org/10.1007/s10826-018-1257-3>.
- Kline, R. B. (2016). *Principles and practice of structural equation modeling*. New York, NY: Guilford Press.
- Merlin, C., Okerson, J. R., & Hess, P. (2013). How parenting style influences children: A review of controlling, guiding, and permitting parenting styles on children's behavior, risk-taking, mental health, and academic achievement. *The William & Mary Educational Review*, 2(1) Retrieved from [http://publish.wm.edu/wmer/vol2/iss1/14](http://publish.wm.edu/wmer%0Ahttp://publish.wm.edu/wmer/vol2/iss1/14).
- Miranda, A., Grau, D., Rosel, J., & Meliá, A. (2009). Understanding discipline in families of children with attention-deficit/hyperactivity disorder: A structural equation model. *The Spanish Journal of Psychology*, 12(2), 496–505 Retrieved from [http://journals.cambridge.org/abstract\\_S1138741600001876](http://journals.cambridge.org/abstract_S1138741600001876).
- Palacios-Hernández, B., Subirà Álvarez, S., & García-Esteve, L. (2015). Comparative study of the factors involved in the mother-infant bonding in women with and without postpartum depression. Universitat Autònoma de Barcelona.
- Prinz, P., Onghena, P., & Hellinckx, W. (2007). Reexamining the parenting scale: Reliability, factor structure, and concurrent validity of a scale for assessing the discipline practices of mothers and fathers of elementary-school-aged children. *European Journal of Psychological Assessment*, 23(1), 24–31. <https://doi.org/10.1027/1015-5759.23.1.24>.
- Pritchett, R., Kemp, J., Wilson, P., Minnis, H., Bryce, G., & Gillberg, C. (2011). Quick, simple measures of family relationships for use in clinical practice and research. A systematic review. *Family Practice*, 28(2), 172–187. <https://doi.org/10.1093/fampra/cmq080>.
- Reitman, D., Currier, R. O., Hupp, S. D. A., Rhode, P. C., Murphy, M. A., & O'Callaghan, P. M. (2001). Psychometric characteristics of the parenting scale in a head start population. *Journal of Clinical Child and Adolescent Psychology*, 30(4), 514–524. [https://doi.org/10.1207/S15374424JCCP3004\\_08](https://doi.org/10.1207/S15374424JCCP3004_08).
- Remor, E. (2006). Psychometric properties of a European Spanish version of the perceived stress scale (PSS). *The Spanish Journal of Psychology*, 9(1), 86–93. <https://doi.org/10.1017/S1138741600006004>.
- Rhoades, K. A., & O'Leary, S. G. (2007). Factor structure and validity of the parenting scale. *Journal of Clinical Child and Adolescent Psychology*, 36(2), 137–146. <https://doi.org/10.1080/15374410701274157>.
- Ruiz-Hernández, J. A., Moral-Zafra, E., Llor-Esteban, B., & Jiménez-Barbero, J. A. (2019). Influence of parental styles and other psychosocial variables on the development of externalizing behaviors in adolescents: A systematic review. *European Journal of Psychology Applied to Legal Context*, 11(1), 9–21. <https://doi.org/10.5093/ejpalc2018a11>.
- Salari, R., Terreros, C., & Sarkadi, A. (2012). Parenting scale: Which version should we use? *Journal of Psychopathology and Behavioral Assessment*, 34(2), 268–281. <https://doi.org/10.1007/s10862-012-9281-x>.
- Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, 8(2), 23–74. <https://doi.org/10.1002/0470010940>.
- Sousa, V. D., & Rojjanasrirat, W. (2011). Translation, adaptation and validation of instruments or scales for use in cross-cultural health care research: A clear and user-friendly guideline. *Journal of Evaluation in Clinical Practice*, 17(2), 268–274. <https://doi.org/10.1111/j.1365-2753.2010.01434.x>.
- Steele, R. G., Nesbitt-Daly, J. S., Daniel, R. C., & Forehand, R. (2005). Factor structure of the parenting scale in a low-income African American sample. *Journal of Child and Family Studies*, 14(4), 535–549. <https://doi.org/10.1007/s10826-005-7187-x>.
- Taylor, A., Atkins, R., Kumar, R., Adams, D., & Glover, V. (2005). A new mother-to-infant bonding scale: Links with early maternal mood. *Archives of Women's Mental Health*, 8(1), 45–51. <https://doi.org/10.1007/s00737-005-0074-z>.
- Yoo, S. Y., & Cho, H. (2020). Exploring the influences of nurses' partnership with parents, attitude to families' importance in nursing care, and professional self-efficacy on quality of pediatric nursing care: A path model. *International Journal of Environmental Research and Public Health*, 17(15), 1–13. <https://doi.org/10.3390/ijerph17155452>.
- Instituto Cervantes. (2021). El Español: Una lengua viva. Informe 2017. Retrieved from [https://cvc.cervantes.es/lengua/espanol\\_lengua\\_viva/pdf/espanol\\_lengua\\_viva\\_2021.pdf](https://cvc.cervantes.es/lengua/espanol_lengua_viva/pdf/espanol_lengua_viva_2021.pdf)