

Adaptation and psychometric properties of SCORE-15 to measure family functioning in Peruvian population

Sánchez-Villena, Andy Rick^a ; Castañeda-Valdivia, Edén^a ; Farfán, Eduardo^a 

Methodological Article

Abstract	Resumen	Tabla de Contenido
<p>SCORE-15 is a self-report instrument with 15 items which purpose is to measure family functioning. It was created in the United Kingdom and has been adapted in different European countries. Nevertheless, there are no studies about its psychometric properties in Latin America. Therefore, the aim of this research is to examine its psychometric properties in Peruvian population. For this, the original version of SCORE-15 was translated into Spanish using the double translation method. Subsequently, an exploratory and confirmatory factor analysis was carried out. The results show that SCORE-15 has two dimensions with adequate reliability, which is invariant according to sex and age. Likewise, the divergent and convergent validity was verified. In conclusion, the SCORE-15, in its Peruvian version, is a brief tool with good psychometric properties for evaluation of family functioning and can be used by researchers or psychotherapist in their clinical practice</p> <p>Keywords: Family assessment, systemic assessment, SCORE-15, validity.</p>	<p>Adaptación y propiedades psicométricas del SCORE-15 para medir el funcionamiento familiar en población peruana. El SCORE-15 es un instrumento de autorreporte con 15 ítems cuyo propósito es medir el funcionamiento familiar. Fue creado en el Reino Unido y ha sido adaptado en diferentes países europeos. Sin embargo, no existen estudios acerca de sus propiedades psicométricas en Latinoamérica. Por ello, la presente investigación tiene como objetivo principal examinar sus propiedades psicométricas en población peruana. Para ello, la versión original del SCORE-15 fue traducida al español con el método de la doble traducción. Posteriormente, se realizó un análisis factorial exploratorio y confirmatorio. Los resultados mostraron que el SCORE-15 tiene dos dimensiones con adecuada confiabilidad, la cual es invariante según sexo y edad. Asimismo, se comprobó la validez divergente y convergente. En conclusión, el SCORE-15, en su versión peruana es una herramienta breve y con buenas propiedades psicométricas para la evaluación del funcionamiento familiar, que puede ser usado por investigadores o psicoterapeutas en su práctica clínica.</p> <p>Palabras clave: Evaluación familiar, evaluación sistémica, SCORE-15, validez.</p>	<p>Introduction 99 Method 100 Participants 100 Instruments and measures 103 Procedure 103 Statistical analysis 103 Results 103 Discussion 105 Acknowledgements 107 References 107</p>

Recibido el 18 de junio de 2021; Aceptado el 2 de febrero de 2022

Editaron este artículo: Fernanda Molina, Paula Abate, Tamara Kobiec y Paula Aguirre

Systems, specifically families, exist to preserve their existence as such. Functional families are those capable of creating an environment that facilitates the personal development of their members. For this reason, they do not usually suffer from severe psychological disorders (González et al., 2012).

The family is an open system, which degree of communication and adaptability to changes, inherent in the history of the family life cycle, has repercussions on family functioning (Arnold & Osorio, 1998). Therefore, functioning or dysfunction delimit the health or disease of the

members. For this reason, for Olson et al. (1979), family functioning could be defined as a property in which family members have strong affective ties with each other without neglecting their autonomy (cohesion) and, at the same time, being able to self-correct when facing difficulties (adaptability). Furthermore, family functioning has been related to eating disorders in adolescents (Cerniglia et al., 2017), juvenile anxiety (Schleider et al., 2015), personality disorders (Cancrini & De Gregorio, 2018), substance use and aggressive behavior (Finan et al., 2015), psychosis (Cancrini & De Gregorio, 2018; Linares, 2018), and depression

^a Universidad Privada del Norte, Facultad de Ciencias de la Salud, Carrera Profesional de Psicología, Cajamarca, Perú.

*Enviar correspondencia a: Sánchez-Villena, A. R. E-mail: andysavi92@gmail.com

(Shao et al., 2020), among other factors involved in psychological well-being.

Given the importance of family functioning, different instruments have been created for its measurement, such as the McMaster Family Assessment Device (Epstein et al., 1983), the Family functioning in adolescence questionnaire (Roelofse & Middleton, 1985), the Family Cohesion and Adaptability Assessment Scale (Olson, 1986), the Family APGAR (Smilkstein, 1978), the Family Relationship Assessment Scale (Kim et al., 2021) or SCORE-40, SCORE-15, SCORE-28, and SCORE-29 (Cahill et al., 2010; Stratton et al., 2010, 2014). The SCORE-15 has advantage over other instruments due to its practicality, since it is a short instrument (Vilaça et al., 2015) that can be used in both adults and adolescents (Zetterqvist et al., 2020). Furthermore, it is based on a current theoretical model, which is why it is the main instrument to measure family functioning by the European Family Therapy Association and the UK Association for Family Therapy (Carr & Stratton, 2017).

Specifically, SCORE-15 (Stratton et al., 2010) is a three-dimensional instrument (communication, adaptability, and difficulties) created in the United Kingdom that has been translated into different languages, and its psychometric properties have been studied in different countries. For example, its scores have been validated in Ireland (Fay et al., 2013; Hamilton et al., 2015; O'Hanrahan et al., 2017), Poland (Józefik et al., 2016), Thailand (Limsuwan & Prachason, 2020), Italy (Paolini & Schepisi, 2020), Korea (Shine et al., 2020), Portugal (Vilaça et al., 2015, 2018), Sweden (Zetterqvist et al., 2020), and Spain (Rivas & Pereira, 2016). However, regarding its internal structure, some studies have found that two indices of goodness of fit did not have satisfactory values when modeling the three dimensions of SCORE-15 (Carr & Stratton, 2017; O'Hanrahan et al., 2017). This would suggest the possibility that the structure is not replicable in some populations (Carr & Stratton, 2017).

In addition, it has been reported that correlations between factors are significant, especially between communication and difficulties dimensions, overcoming .80, even .90 (Hamilton et al., 2015; O'Hanrahan et al., 2017; Zetterqvist et al., 2020), which could indicate that both would

form a single factor (Reise, 2012). This would explain the little adjustment of the data and the possibility of a two-dimensional structure of SCORE-15 (see Table 1).

For this reason, the aim of this research is to examine the psychometric properties of SCORE-15 in the Peruvian general population, which is justified by the following reasons: First, because there is no history of validation or adaptation of this instrument in the Latin American context. Perhaps the closest antecedent, from the linguistic point of view, is the Spanish translation of SCORE-15. However, we consider important to have a Peruvian adaptation, because the items in the Spanish version contain words that do not achieve the semantic equivalence to the intended population. Therefore, provide a version of SCORE-15 adapted to the Spanish spoken in Peru, can maximize the suitability of the test to Peruvian population (International Test Commission, 2017). Second, because SCORE-15 is an instrument that has shown to be useful in psychotherapy, especially from systemic approach. Third, because by translating and validating SCORE-15, an accurate and theoretically supported tool could be available for mental health professionals, such as family psychotherapists, to adequately measure and evaluate family functioning, so it would have a practical purpose.

Method

Participants

This study considered a total of 1637 Peruvian participants (53.1% women), mostly residents of Cajamarca (75.4%), Lima (11.9%), and La Libertad (7.0%). The first sample consisted of 507 participants (55.2% women), who aged from 12 to 69 years ($M = 22.49$, $SD = 9.36$); they were part of the study for Exploratory Factor Analysis (EFA). The second sample consisted of 1130 people (51.9% women), aging from 12 to 72 years old ($M = 21.93$, $SD = 9.04$); they were part of the study of Confirmatory Factor Analysis (CFA). It should be noted that to collect the data, a non-probabilistic by snowball sampling was used through Google forms (<https://www.docs.google.com/forms>), which was shared by social networks in June 2020.

Table 1
Summary of previous studies

Authors	Country	<i>n</i>	Method	Ndim	CFI	RMSEA	SRMR	WRMR	α_{Total}	α_{F1}	α_{F2}	α_{F3}	Φ
Fay et al. (2013)	Ireland	403	CFA; alpha	3	.92	.07	.05	-	.83	.76	.58	.71	.58
Hamilton et al. (2015)	Ireland	701 ^a	CFA; alpha	3	.98	.06	.04	-	.90	.83	.78	.85	.97
Józefik et al. (2016)	Poland	-	Translation	-	-	-	-	-	-	-	-	-	-
Limsuwan and Prachason (2020)	Thailand	135	Criterion validity (non-clinical vs clinical); test-retest	-	-	-	-	-	.94	.84	.86	.91	-
O'Hanrahan et al. (2016)	Ireland	199	CFA; alpha	3	.95	.10	-	.92	.92	.88	.77	.83	.94
Paolini and Schepisi (2020)	Italy	203	CFA; alpha (fathers)	3	.91	.06	.06	-	.80	.73	.65	.78	.54
		214	CFA; alpha (mothers)	3	.91	.06	.06	-	.81	.78	.60	.76	.49
		90	CFA; alpha(children)	3	1.00	.01	.05	-	.89	.84	.55	.84	.66
Rivas and Pereira (2016)	Spain	516	EFA; alpha	3	-	-	-	-	.85	.85	.62	.80	-
		390	CFA; alpha	3	.96	.06	.04	-	.85	.85	.62	.80	-
Shine et al. (2020)	Korea	433	Concurrent validity with FACES III; alpha	3	-	-	-	-	.92	.89	.73	.87	.80
Vilaça et al. (2015)	Portugal	276	Discriminant validity (non-clinical vs clinical); alpha	3	-	-	-	-	.84/.85	.82/.87	.69/.68	.70/.76	-
		430	Convergent validity with QoL; alpha	3	-	-	-	-	.85/.87	.70/.87	.61/.67	.78/.80	.75
		66	Predictive validity (1st session vs 4th session); alpha	3	-	-	-	-	.78/.84	.78/.84	.62/.64	.72/.76	-
Vilaça et al. (2018)	Portugal	538	CFA; alpha	3	.95	.06	-	-	.88	.86	.73	.78	.72
Zetterqvist et al. (2020)	Sweden	215	CFA/ (non-clinical); alpha	3	.91	.07	-	-	-	.73	.69	.82	.64
		159	CFA (clinical); alpha	3	.87	.07	-	-	-	.75	.71	.82	.72

Note. ^a = number of families; *n* = sample; Ndim = number of dimensions; CFI = Comparative fit index; F1 = Strengths; F2 = Communication; F3 = Difficulties; / = α for clinical sample; Φ = Interfactorial correlation between F2 and F3.

Instruments and Measures

To collect sociodemographic data, three questions were incorporated into the Google form. These were: age, sex and city of residence.

Participants completed the SCORE-15 (Stratton et al., 2010), which is a self-report instrument with three dimensions: strengths, communication, and difficulties, and 15 Likert-scale items with five scores ranging from *Describes Us Very Well* = 1 to *Do Not Describe Us at All* = 5, which objective is to measure family functioning. Each dimension has five items; thus, strengths contains items 1, 3, 6, 10 and 15 (e.g., In my family we trust each other); communication contains items 2, 4, 8, 12 and 13 (e.g., In my family often don't tell each other the truth) and difficulties contains items 5, 7, 9, 11 and 14 (e.g., Things always seem to go wrong for my family). In order to get scores, items 2, 4, 5, 7, 8, 9, 11, 12, 13, and 14 must be inverted (*Describes Us Very Well* = 5 to *Do Not Describe Us at All* = 1). Then, direct scores are added and a total score is obtained. Therefore, lower ratings indicate good family functioning and higher ratings indicate problematic family functioning. The approximate time per application is between 10 and 15 minutes. The original scale demonstrated factorial validity by EFA, which indicated three dimensions. Besides reliability showed a Cronbach's $\alpha = .89$ (Stratton et al., 2010). We used the Peruvian version of SCORE-15, developed in the current study.

APGAR (Smilkstein, 1978) is a unidimensional self-report instrument with five-point Likert-scale ranging from *Never* = 0 to *Always* = 4, which objective is to measure the perception of family functioning (e.g., I am satisfied with the help I receive from my family when something worries me). In this study, we used the Peruvian adaptation (Castilla et al., 2014), which validity was examined by EFA indicating one dimension and its reliability showed an $\alpha = .79$. To obtain a total score, five items are added together and a minimum score of 0 and a maximum of 20 is obtained, so higher scores indicate better perception of family functioning. The approximate time per application is between 10 and 15 minutes.

Patient Health Questionnaire (PHQ-9) is a one-dimensional self-report instrument, that aims to measure depressive symptoms consisting of nine Likert-scale items (e.g., little interest or pleasure in

doing things) and four responses options: *not at all* = 0 to *nearly every day* = 3 with scores ranging from 0 to 27. In this study, we used the Peruvian version (Calderón et al., 2012). Its validity and reliability in Perú were examined by CFA and internal consistency (Villarreal-Zegarra et al., 2019), which showed one dimension (CFI = .94; TLI = .91; SRMR = .04; $\omega = .87$).

Finally, it should be noted that both APGAR and PHQ-9 were used to test convergent and divergent validity, respectively.

Procedure

The procedure for carrying out this research was divided into two different moments.

First, permission for translating and validating SCORE-15 was requested to the owners of the test (P. Stratton, personal communication, 21st June 2020). The double translation method was used; the items obtained were compared with the Spanish version (Rivas & Pereira, 2016) and the original version (Stratton et al., 2010). This translation was made with the help of five Peruvian systemic psychotherapists (see Annex), and the checklist by Muñiz et al. (2013) was used to corroborate the achievement of linguistic equivalence with the Aiken V to measure the agreement between the judges. The results were sent to the official SCORE-15 translation team, who reviewed the final version of the items in Spanish.

Second, due to the pandemic and social restrictions, data was collected virtually by using Google forms through snowball sampling. The instructions were displayed within the form and the order in which the instruments were presented was as follows: sociodemographic questions, SCORE-15, APGAR, and PHQ-9. The initial contact with the participants was made through social networks. Before responding the scales, they received an informed consent which indicated that the participation was voluntary and anonymous, and that data obtained would be processed exclusively for academic purposes, without requesting personal information that allows them to be identified. Furthermore, participants under 18 years old had to have the approval and acceptance of their parents. For that reason, 15 people were excluded from the investigation.

Finally, it should be noted that the ethical aspects of this study were approved by the Universidad Privada del Norte, in Cajamarca,

Perú, following the recommendations of the Declaration of Helsinki.

Statistical Analysis

An EFA was performed through the FACTOR software (Lorenzo-Seva & Ferrando, 2006) using unweighted least squares (ULS) as estimation method with polychoric correlation matrix, optimal parallel analysis, and sedimentation graph for the estimation of dimensions together with a Promin rotation. We used ULS method for the following reasons: a) because of the ordinality of the items, b) the sample size for EFA was adequate ($n = 507$), c) because ULS avoids the appearance of Heywood Cases and, finally d) ULS does not require distributional assumptions about normality (Lloret-Segura et al., 2014).

Then, a CFA was performed in the R package Lavaan, where two models were tested: the first one with a three-dimensional structure (corresponding to the original model), which showed convergence problems. The second one was a two-dimensional structure, which followed the model suggested by EFA. For this, the weighted least squares means, and variance adjusted (WLSMV) method of estimation was used due to data ordinality, its robustness to non-normality and because the sample size was adequate for this estimator (Kline, 2015). Robust indices of CFI ($> .95$), TLI ($> .95$), RMSEA ($< .06$) were used to assess the models (Hu & Bentler, 1999).

The measurement equivalence by sex and age was carried out by following the recommendations of Wu and Estabrook (2016). For this, the configural invariance was estimated. Then, a model with equal thresholds was estimated, and a third model equalizing thresholds and factorial loads. Finally, a fourth model was estimated equalizing thresholds, factorial loads and intercepts. To assess the measurement equivalence, the cut-off points $\Delta CFI < .01$ and $\Delta RMSEA < .02$ were considered (Chen, 2007). It is important to mention that to create groups according to age the criteria of Mansilla (2000) were taken. This author argues that in Peru people aged between 12 and 17 are considered adolescents; those aged 18 to 24 are young and those over 25 are adults. The sample of adults was not stratified (for example, primary, intermediate, and older adults), because there were not enough participants in each group to

compare them. The analysis according to age was carried out in order to control, at least partially, the effect of the stages of human development involved in the family life cycle, since age differences create family subsystems within the family system (Minuchin et al., 1998).

Finally, convergent validity was tested based on the relationship with other variables with Pearson correlation, and internal consistency was calculated by using the omega coefficient in JASP (Version 0.15; JASP Team, 2021).

Results

Content validity

The content validity of the Peruvian version of SCORE-15 was made by expert judgement and calculated with Aiken's V. It should be noted that the experts were five Peruvian systemic psychotherapists with experience in research. Aiken's V results were adequate, because the values obtained in the criterion showed a range between .67 (item 7) and 1 in relevance, while the values in representativity had a minimum value of .73 (item 7) with a maximum of 1, concluding that content validity is met. Translation and changes in the items of the Peruvian version, can be seen in Annex.

Preliminary analysis of the items

As seen in Table 2, item 12 (*My family members do not tolerate each other*) and 11 (*Things always seem to go wrong for my family*) have a higher arithmetic mean both in the sample of EFA and CFA. All items showed a value above the .20 recommended in the item-total correlation (Kline, 1986). In fact, the obtained values in the EFA sample ranged from .32 to .67, while in the CFA sample they ranged from .33 to .69.

First study (Exploratory factor analysis)

Before proceeding with EFA, adequacy measures were calculated indicating that EFA could be carried out, since $KMO = .93$ and the Bartlett sphericity test was statistically significant ($X^2_{(105)} = 3330.50, p < .05$). In view of this, ULS was used as estimator along with polychoric correlation matrices due to the ordinal nature of items. Optimal parallel analysis was used as dimension extraction method. Lastly, Promax rotation was applied. The results indicated the presence of two dimensions that explained 42.4% and 11.4% of the total variance (53.8%),

Table 2
Preliminary analysis of the items

Items	EFA				CFA			
	M	SD	rit	ω	M	SD	rit	ω
1. En mi familia hablamos de las cosas que son importantes para nosotros.	2.02	0.93	.56	.86	1.97	0.96	.35	.88
2. En mi familia, muchas veces, no nos decimos la verdad	3.38	1.17	.32	.87	3.26	1.20	.47	.87
3. En nuestra familia todos somos escuchados	2.04	0.97	.57	.86	1.96	1.00	.42	.87
4. Siento que, en nuestra familia, estar en desacuerdo es arriesgado	3.53	1.21	.43	.87	3.38	1.26	.45	.87
5. Como familia, es difícil lidiar con problemas cotidianos	3.52	1.11	.53	.86	3.36	1.19	.54	.87
6. En mi familia nos tenemos confianza	2.14	1.10	.52	.86	2.01	1.09	.33	.88
7. En nuestra familia, uno se siente desanimado	3.76	1.12	.51	.86	3.60	1.20	.62	.86
8. Cuando alguien de mi familia se enoja con otro, se ignoran mutuamente.	3.28	1.18	.43	.87	3.28	1.20	.55	.87
9. Parece que, en mi familia, salimos de un problema para entrar en otro	3.96	1.08	.67	.85	3.85	1.18	.67	.86
10. Cuando algún miembro de la familia está preocupado, lo apoyamos	1.71	0.86	.49	.86	1.71	0.88	.46	.87
11. Las cosas siempre parecen ir mal para mi familia	4.07	0.96	.63	.86	3.91	1.13	.69	.86
12. Los miembros de mi familia no se toleran entre ellos.	4.18	1.02	.60	.86	4.04	1.17	.68	.86
13. Los miembros de mi familia interfieren demasiado en la vida del otro	3.70	1.08	.42	.87	3.68	1.18	.53	.87
14. En mi familia, nos echamos la culpa los unos a los otros cuando las cosas van mal	4.00	1.04	.60	.86	3.87	1.19	.69	.86
15. Somos buenos encontrando nuevas formas de enfrentar situaciones difíciles	2.05	0.93	.48	.86	2.03	1.00	.30	.88

Note. M = mean; SD = standard deviation; rit = item-rest correlation; ω = omega if item dropped

respectively with an interfactorial correlation of $\Phi = -.64$.

Table 3 shows that factorial loads after rotation were grouped in two dimensions. The first factor, called strengths/adaptability, is composed of five items. The second factor, called family difficulties, is composed of 10 items. Saturations of items in both dimensions exceed .50, therefore it was not necessary to eliminate any of them.

Second Study (Confirmatory Factor Analysis)

Two models were tested using WLSMV as estimation method due to data ordinality. The three-dimensional model, corresponding to the original structure, had convergence problems, so it was set aside. The correlated two-dimensional model showed an adequate adjustment (CFI = .98, TLI = .98, RMSEA = .06 [.05- .06], SRMR = .04). Factorial loads are shown in Table 3.

Invariance by sex and age

A multigroup confirmatory factor analysis was carried out to prove the measurement equivalence by sex and age. Table 4 shows that the base

model had adequate goodness-of-fit indices, so parameters are progressively restricted. In this way, it is demonstrated that configurational, metric, strong, and strict invariance had an adequate adjustment of the model in men and women, since changes in values of ΔCFI and $\Delta RMSEA$ were inferior to .01 and .02 (Chen, 2007), respectively. In the same way, the invariance according to age groups showed adequate values in the ΔCFI and $\Delta RMSEA$ in the groups 12 to 17, 18 to 24, and 25 and up. Therefore, it can be concluded that the SCORE-15 measures the same construct in men and women and in the different age groups.

Validity based on relationship with other variables

To check the convergent and divergent validity, the scores of each dimension were correlated with family APGAR and PHQ-9, which measure perception of family functioning and depression, respectively. Results showed that strengths/adaptability has a statistically significant, positive, and medium effect size correlation with

Table 3
Factorial loadings of SCORE-15 in EFA and CFA

Items	EFA (Rotated matrix)		CFA	
	F1	F2	F1	F2
1	.78	.01	.66	
3	.69	.10	.75	
6	.79	.02	.65	
10	.80	.04	.83	
15	.75	.03	.60	
2	.04	.35		.55
4	.16	.67		.56
8	.03	.50		.63
12	.01	.76		.84
13	.11	.62		.65
5	.06	.69		.64
7	.04	.60		.72
9	.10	.73		.80
11	.13	.66		.83
14	.03	.71		.81
Φ		-.64		-.48
ω	.83	.83	.76	.89

Note. F1 = Strengths; F2 = Difficulties; Φ = interfactor correlation; ω = omega coefficient.

Table 4
Factor invariance by sex and age of SCORE-15

		CFI	RMSEA	ΔCFI	ΔRMSEA
Sex (Men vs. Women)	Configural	.98	.06	-	-
	Metric	.98	.05	.00	.00
	Strong	.98	.05	.00	.00
	Strict	.98	.05	.00	.00
Age (12-17 vs. 18-24 vs. 25 and up)	Configural	.97	.06	-	-
	Metric	.98	.05	.00	-.01
	Strong	.98	.05	.00	.00
	Strict	.98	.05	.00	.00

Note. CFI = Comparative fit index; RMSEA = Root mean squared error of approximation

the APGAR ($r = .65, p < .01$), while the correlation with the PHQ-9 was statistically significant, negative, and with a small effect size ($r = -.37, p < .01$). On the other hand, family difficulties has a statistically significant, negative, and small effect size correlation with the APGAR ($r = -.48, p < .01$). While the correlation between family difficulties and the PHQ-9 showed a statistically significant, positive and with a small effect size ($r = .29, p <$

.01).

Reliability

Reliability was calculated using omega coefficient. In the first study, the results showed an omega of .83 for strengths/adaptability and of .83 for family difficulties. While, in the second study, strengths/adaptability obtained an omega of .76 and .89 for family difficulties.

Discussion

The aim of this research was to examine psychometric properties of SCORE-15 in the Peruvian general population. To achieve it, the original instrument in English was firstly translated into Spanish. Then, the internal structure was identified by exploratory and confirmatory factor analyses. The measurement equivalence and internal consistency of SCORE15 were then analyzed.

Regarding translation and content validity, item 7 presented the lowest values in Aiken's V, besides receiving observations from the judges, mainly for the word *miserable*, which meaning in English refers to someone wretched, downcast, sad, or unhappy. However, in Spanish it can also have other connotations referring, for example, to someone stingy, despicable, or a person in extreme poverty. For that reason, the word *desanimado* (discouraged) was chosen, since it has more sense of someone who, in the family context, feels decayed.

Once the items were translated, the internal structure of SCORE-15 was examined. EFA showed a structure of two factors: the first one, with five items corresponding to the strengths/adaptability dimension; and the second one, with 10 items that make up the family difficulties dimension. Due to this, a CFA was carried out to test the original three-dimensional model (Stratton et al., 2010) and the two-dimensional model suggested by EFA. CFA showed that the three-dimensional model of SCORE-15 presented convergence problems, while the two-dimensional model obtained appropriate goodness of fit indices. These results contrast with previous studies from Europe and Asia (Fay et al., 2013; Hamilton et al., 2015; Józefik et al., 2016; Limsuwan & Prachason, 2020; Paolini & Schepisi, 2020; Shine et al., 2020; Vilaça et al., 2015, 2018; Zetterqvist et al., 2020), even with the Spanish version (Rivas & Pereira, 2016),

they all found a three-dimensional structure. However, the results coincide with the research of O'Hanrahan et al. (2017), which did not show adequate adjustments for the three-dimensional structure of the instrument. This would indicate that SCORE-15 dimensionality is not replicable in some populations, as indicated by Carr and Stratton (2017). This may be to the fact that correlations between communication and difficulties are large, which indicates that both dimensions have much in common (Reise, 2012). Therefore, it makes sense that they form a single factor. In fact, previous studies have reported high correlations between these dimensions (Hamilton et al., 2015; O'Hanrahan et al., 2017; Zetterqvist et al., 2020), which is theoretically coherent, since communication is directly involved in family difficulties (Vilaça et al., 2018).

When checking the proper adjustment of the two-dimensional model, factor invariance according to sex and age was analyzed, which is important to avoid bias in measurement in both groups and implies that the scores are valid regardless of the sex or age. The results of multigroup CFA showed that the two-dimensional structure is invariant in both males and females of Peruvian general population and, therefore, the construct of family functioning is understood in the same way by both sexes. In other words, SCORE-15 items have been answered and interpreted similarly by men and women.

On the other hand, the invariance analysis was also tested in three groups: from 12 to 17, 18 to 24, and 25 to more years of age, in order to know if the SCORE-15 keeps the equivalence of measurement for people in different ages. For this, we started from the idea that the family, being a system, whose members are considered subsystems, may be affected by the stage of human development of each member, which affects the family cycle (Minuchin et al., 1998). For this reason, SCORE-15 is invariant according to age and can be administered to a population older than 12 years.

Regarding the validity based on the relationship with other variables, correlations were found between the APGAR, the PHQ-9, and the SCORE-15 dimensions. Thus, the results indicate that family functioning, measured by APGAR, is higher when there are more strengths/adaptability; while family functioning is lower when there are more family difficulties. Furthermore, it was found

that depression, measured by the PHQ-9, decreases when there are more strengths/adaptability and increases when there are more difficulties. These findings are similar to previous studies (Carr & Stratton, 2017; Shao et al., 2020).

Regarding SCORE-15 reliability, it was calculated by using the coefficient Omega for strengths/adaptability and family difficulties dimensions, which showed adequate indices of internal consistency ($> .80$). This is consistent with previous studies (Fay et al., 2013; Hamilton et al., 2015; Józefik et al., 2016; Limsuwan & Prachason, 2020; Paolini & Schepisi, 2020; Rivas & Pereira, 2016; Shine et al., 2020; Vilaça et al., 2015, 2018; Zetterqvist et al., 2020) and indicates that instrument scores are reliable in the Peruvian general population.

Finally, the limitations of this study were that there was no diagnostic study of participants' family situation, so no data on the sensitivity or specificity of SCORE-15 scores were obtained, which can be considered in later research. Besides, as the instrument was applied in a single moment, the stability of SCORE-15 scores over time was not studied, thus, longitudinal designs are required to demonstrate longitudinal invariance and reliability by test-retest method. Furthermore, another methodological limitation is related to the use of snowball sampling through virtual platforms, since this method does not allow researchers to control and verify if participants have answered twice, for example. So, we suggest that future psychometric studies about SCORE-15 use sampling methods with a higher level of control such as random sampling or, whenever possible, collect data by telephone or physical ways.

Despite these limitations, this study provides valuable information on the metric properties of the Spanish-translated version of SCORE-15, and its application can be useful to measure family functioning in the Peruvian general population. This is important, considering that it is the first psychometric study of this instrument in Latin America, setting an important precedent in the study of family functioning in this context. In addition, it has practical implications since it allows having an instrument which scores are valid and reliable for psychotherapists who wish to obtain reliable measurements on such an important variable in family and personal well-being.

In conclusion, this research shows that the

Spanish-translated version of SCORE-15 has adequate psychometric properties in the Peruvian general population, demonstrating a two-dimensional model, which is invariant according to sex and age.

Acknowledgments

Este trabajo ha sido realizado con el aporte de Subsidios de la Universidad Privada del Norte, Cajamarca, Perú (Código 20202001).

References

- Arnold, M., & Osorio, F. (1998). Introducción a los conceptos básicos de la teoría general de sistemas. *Cinta de Moebio, Revista De Epistemología De Ciencias Sociales*, 3, 40-49. <https://cintademoebio.uchile.cl/index.php/CDM/artic le/view/26455>
- Cahill, P., O'Reilly, K., Carr, A., Dooley, B., & Stratton, P. (2010). Validation of a 28-item version of the Systemic Clinical Outcome and Routine Evaluation in an Irish context: the SCORE-28. *Journal of Family Therapy*, 32(3), 210–231. <https://doi.org/10.1111/j.1467-6427.2010.00506.x>
- Calderón, M., Gálvez-Buccollini, J. A., Cueva, G., Ordoñez, C., Bromley, C., & Fiestas, F. (2012). Validación de la versión peruana del PHQ-9 para el diagnóstico de depresión. *Revista Peruana de Medicina Experimental y Salud Pública*, 29(4), 578–579. <https://rpmesp.ins.gob.pe/index.php/rpmesp/article/ view/409/409>
- Cancrini, L., & De Gregorio, F. R. (2018). Borderline Personality Disorder Storyboard from the Systemic Family Therapist's Perspective. In R. Pereira & J. Linares (Eds.), *Clinical Interventions in Systemic Couple and Family Therapy* (pp. 15–29). Springer. https://doi.org/10.1007/978-3-319-78521-9_2
- Carr, A., & Stratton, P. (2017). The Score Family Assessment Questionnaire: A Decade of Progress. *Family Process*, 56(2), 285–301. <https://doi.org/10.1111/famp.12280>
- Castilla, H. A., Caycho, T. P., Shimabukuro, M., & Valdivia, A. A. (2014). Percepción del funcionamiento familiar: Análisis psicométrico de la Escala APGAR-familiar en adolescentes de Lima. *Propósitos y Representaciones*, 2(1), 49–78. <https://doi.org/10.20511/pyr2014.v2n1.53>
- Cerniglia, L., Cimino, S., Tafà, M., Marzilli, E., Ballarotto, G., & Bracaglia, F. (2017). Family profiles in eating disorders: family functioning and psychopathology. *Psychology Research and Behavior Management*, 10, 305–312. <https://doi.org/10.2147/PRBM.S145463>
- Chen, F. F. (2007). Sensitivity of Goodness of Fit Indexes to Lack of Measurement Invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, 14(3), 464–504. <https://doi.org/10.1080/10705510701301834>
- Epstein, N. B., Baldwin, L. M., & Bishop, D. S. (1983). The McMaster Family Assessment Device. *Journal of Marital and Family Therapy*, 9(2), 171–180. <https://doi.org/10.1111/j.1752-0606.1983.tb01497.x>
- Fay, D., Carr, A., O'Reilly, K., Cahill, P., Dooley, B., Guerin, S., & Stratton, P. (2013). Irish norms for the SCORE-15 and 28 from a national telephone survey. *Journal of Family Therapy*, 35(1), 24–42. <https://doi.org/10.1111/j.1467-6427.2011.00575.x>
- Finan, L. J., Schulz, J., Gordon, M. S., & Ohannessian, C. M. (2015). Parental problem drinking and adolescent externalizing behaviors: The mediating role of family functioning. *Journal of Adolescence*, 43(1), 100–110. <https://doi.org/10.1016/j.adolescence.2015.05.001>
- González, F., Gimeno, A., Meléndez, J. C., & Córdoba, A. (2012). La percepción de la funcionalidad familiar: Confirmación de su estructura bifactorial. *Escritos de Psicología*, 5(1), 34–39. <https://doi.org/10.24310/espsiescpsi.v5i1.13294>
- Hamilton, E., Carr, A., Cahill, P., Cassells, C., & Hartnett, D. (2015). Psychometric Properties and Responsiveness to Change of 15- and 28-Item Versions of the SCORE: A Family Assessment Questionnaire. *Family Process*, 54(3), 454–463. <https://doi.org/10.1111/famp.12117>
- Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: A multidisciplinary journal*, 6(1), 1–55. <https://doi.org/10.1080/10705519909540118>
- International Test Commission (2017). *ITC Guidelines for Translating and Adapting Tests* (2nd ed.). International Test Commission. https://www.intestcom.org/files/guideline_test_adap tation_2ed.pdf
- JASP Team. (2021). JASP (Version 0.15) [Computer software]. <https://jasp-stats.org/>
- Józefik, B., Matusiak, F., Wolska, M., & Ułasińska, R. (2016). Family therapy process – works on the Polish version of SCORE-15 tool. *Psychiatria Polska*, 50(3), 607–619. <https://doi.org/10.12740/PP/OnlineFirst/42894>
- Kim, S., Gil, M., & Kim-Godwin, Y. (2021). Development and Validation of the Family Relationship Assessment Scale in Korean College Students' Families. *Family Process*, 60(2), 586–601. <https://doi.org/10.1111/famp.12559>
- Kline, P. (1986). *A Handbook of test construction: introduction to psychometric design*. Routledge.
- Kline, R. B. (2015). *Principles and practice of structural equation modeling* (4th ed.). Guilford Press.
- Limsuwan, N., & Prachason, T. (2020). The reliability

- and validity of the 15-item Systemic Clinical Outcome and Routine Evaluation (SCORE-15) Thai version. *Journal of Family Therapy*, 42(1), 119–128. <https://doi.org/10.1111/1467-6427.12248>
- Linares, J. L. (2018). The Family Therapy of Psychosis: A Reconfirmation Process. In R. Pereira & J. Linares (Eds.), *Clinical Interventions in Systemic Couple and Family Therapy* (pp. 3–13). Springer. https://doi.org/10.1007/978-3-319-78521-9_1
- Lloret-Segura, S., Ferreres-Traver, A., Hernández-Baeza, A., & Tomás-Marco, I. (2014). El análisis factorial exploratorio de los ítems: una guía práctica, revisada y actualizada. *Anales de Psicología*, 30(3), 1151–1169. <https://doi.org/10.6018/analesps.30.3.199361>
- Lorenzo-Seva, U., & Ferrando, P. J. (2006). FACTOR: A computer program to fit the exploratory factor analysis model. *Behavior Research Methods*, 38, 88–91. <https://doi.org/10.3758/BF03192753>
- Mansilla, M. E. (2000). Etapas del desarrollo humano. *Revista de Investigación en Psicología*, 3(2), 105–116. <https://doi.org/10.15381/rinvp.v3i2.4999>
- Minuchin, S., Lee, W.-Y., & Simón, G. (1998). *El arte de la terapia familiar*. Paidós.
- Muñiz, J., Elosua, P., & Hambleton, R. K. (2013). Directrices para la traducción y adaptación de los tests: segunda edición. *Psicothema*, 25(2), 151–157. <https://doi.org/10.7334/psicothema2013.24>
- O’Hanrahan, K., Daly White, M., Carr, A., Cahill, P., Keenleyside, M., Fitzhenry, M., Harte, E., Hayes, J., Noonan, H., O’Shea, H., McCullagh, A., McGuinness, S., Rodgers, C., Whelan, N., Sheppard, N., & Browne, S. (2017). Validation of 28 and 15 item versions of the SCORE family assessment questionnaire with adult mental health service users. *Journal of Family Therapy*, 39(1), 4–20. <https://doi.org/10.1111/1467-6427.12107>
- Olson, D. H. (1986). Circumplex Model VII: Validation Studies and FACES III. *Family Process*, 25(3), 337–351. <https://doi.org/10.1111/j.1545-5300.1986.00337.x>
- Olson, D. H., Sprenkle, D. H., & Russell, C. S. (1979). Circumplex Model of Marital and Family Systems: I. Cohesion and Adaptability Dimensions, Family Types, and Clinical Applications. *Family Process*, 18(1), 3–28. <https://doi.org/10.1111/j.1545-5300.1979.00003.x>
- Paolini, D., & Schepisi, L. (2020). The Italian Version of SCORE-15: Validation and Potential Use. *Family Process*, 59(4), 1789-1800. <https://doi.org/10.1111/famp.12495>
- Reise, S. P. (2012). The Rediscovery of Bifactor Measurement Models. *Multivariate Behavioral Research*, 47(5), 667–696. <https://doi.org/10.1080/00273171.2012.715555>
- Rivas, G., & Pereira, R. (2016). Validación de una escala de evaluación familiar: adaptación del SCORE-15 con normas en español. *Mosaico: Revista de la Federación Española de Asociaciones de Terapia Familiar*, 63, 16-28. <https://dialnet.unirioja.es/servlet/articulo?codigo=6942907>
- Roelofse, R., & Middleton, M. R. (1985). The family functioning in adolescence questionnaire: a measure of psychosocial family health during adolescence. *Journal of Adolescence*, 8(1), 33–45. [https://doi.org/10.1016/S0140-1971\(85\)80005-1](https://doi.org/10.1016/S0140-1971(85)80005-1)
- Schleider, J. L., Ginsburg, G. S., Keeton, C. P., Weisz, J. R., Birmaher, B., Kendall, P. C., Piacentini, J., Sherrill, J., & Walkup, J. T. (2015). Parental psychopathology and treatment outcome for anxious youth: Roles of family functioning and caregiver strain. *Journal of Consulting and Clinical Psychology*, 83(1), 213–224. <https://doi.org/10.1037/a0037935>
- Shao, R., He, P., Ling, B., Tan, L., Xu, L., Hou, Y., Kong, L., & Yang, Y. (2020). Prevalence of depression and anxiety and correlations between depression, anxiety, family functioning, social support and coping styles among Chinese medical students. *BMC Psychology*, 8, 38. <https://doi.org/10.1186/s40359-020-00402-8>
- Shine, B.-K., Park, Y.-J., Ahn, B.-W., Bae, J., Park, J.-S., & Han, S.-H. (2020). Validation Study of Korean Translated Systemic Clinical Outcome Routine Evaluation-15 as Self-Report Family Assessment Measure: Focusing on Adolescent in Daegu and North Gyeongsang Province. *Korean Journal of Family Medicine*, 41(6), 398-403. <https://doi.org/10.4082/kjfm.19.0076>
- Smilkstein, G. (1978). The family APGAR: A proposal for a family function test and its uses by physicians. *The Journal of Family Practice*, 6(6), 1231–1239. https://cdn.mdedge.com/files/s3fs-public/jfp-archived-issues/1978-volume_6-7/JFP_1978-06_v6_i6_the-family-apgar-a-proposal-for-a-family.pdf
- Stratton, P., Bland, J., Janes, E., & Lask, J. (2010). Developing an indicator of family function and a practicable outcome measure for systemic family and couple therapy: the SCORE. *Journal of Family Therapy*, 32(3), 232–258. <https://doi.org/10.1111/j.1467-6427.2010.00507.x>
- Stratton, P., Lask, J., Bland, J., Nowotny, E., Evans, C., Singh, R., Janes, E., & Peppiatt, A. (2014). Detecting therapeutic improvement early in therapy: validation of the SCORE-15 index of family functioning and change. *Journal of Family Therapy*, 36(1), 3–19. <https://doi.org/10.1111/1467-6427.12022>
- Vilaça, M., de Sousa, B., Stratton, P., & Relvas, A. P. (2015). The 15-item Systemic Clinical Outcome and Routine Evaluation (SCORE-15) Scale: Portuguese Validation Studies. *The Spanish*

Journal of Psychology, 18, E87.

<https://doi.org/10.1017/sjp.2015.95>

- Vilaça, M., Relvas, A. P., & Stratton, P. (2018). A Portuguese translation of the Systemic Clinical Outcome and Routine Evaluation (SCORE): the psychometric properties of the 15- and 28-item versions. *Journal of Family Therapy*, 40(4), 537–556. <https://doi.org/10.1111/1467-6427.12197>
- Villarreal-Zegarra, D., Copez-Lonzoy, A., Bernabé-Ortiz, A., Melendez-Torres, G. J., & Bazo-Alvarez, J. C. (2019). Valid group comparisons can be made with the Patient Health Questionnaire (PHQ-9): A measurement invariance study across groups by demographic characteristics. *PLoS ONE*, 14(9), e0221717. <https://doi.org/10.1371/journal.pone.0221717>
- Wu, H., & Estabrook, R. (2016). Identification of Confirmatory Factor Analysis Models of Different Levels of Invariance for Ordered Categorical Outcomes. *Psychometrika*, 81, 1014–1045. <https://doi.org/10.1007/s11336-016-9506-0>
- Zetterqvist, M., Hånell, H. E., Wadsby, M., Cocozza, M., & Gustafsson, P. A. (2020). Validation of the Systemic Clinical Outcome and Routine Evaluation (SCORE-15) self-report questionnaire: index of family functioning and change in Swedish families. *Journal of Family Therapy*, 42(1), 129–148. <https://doi.org/10.1111/1467-6427.12255>