edual editorial UNIVERSIDAD DE ALMERÍA

Electronic Journal of Research in Educational Psychology ejrep

UNIVERSIDAD DE ALMERÍA

Ilustre Colegio Oficial de Psicología DE ANDALUCÍA ORIENTAL

Rosenberg Self-Esteem Scale (RSES): Analysis of the factorial structure and proposal of a new version of only positive items

Lindsey W. Vilca[1], **Aaron Travezaño-Cabrera** [2], **Stephany Santos-Garcia** [2]

[1] South American Center for Education and Research in Public Health, Universidad Norbert Wiener, Lima, Perú
[2] Departamento de Psicología, Universidad Peruana Unión, Lima, Perú

Perú

Correspondence: South American Center for Education and Research in Public Health, Universidad Norbert Wiener, Lima, Perú. E-mail address: lwquiro@gmail.com (Lindsey W. Vilca). Phone number: +51 986357033

Abstract

Introduction. The objective was to evaluate the factor structure and propose a new version of the Rosenberg Self-Esteem Scale with only positive items to overcome the method effect associated with negative items.

Method. A version A (positive and negative items) and a version B (only positive items) were considered. A sample of 350 university students was collected for each version.

Results. The CFA shows that version A's one-dimensional model (A1) does not present adequate fit indices. It was also found that adding a specific factor for negative items (model A2) and another factor for positive items (model A3) does not improve the fit indices. A two-dimensional model (A4 model) does not improve the fit indices either. Regarding the one-dimensional model of version B, it presents a superior fit compared to the original model (model A1)

Discussion and Conclusions. It is concluded that version B, a proposal of only positive items, adequately measures self-esteem since it does not have negative items and agrees with Rosenberg's original approach.

Key Words: Rosenberg Self-esteem Scale, confirmatory factor analysis, self-esteem, construct validity, negative items, university students.

Abstract

Introduction. The objective was to evaluate the factorial structure and propose a new version of the Rosenberg Self-esteem Scale with only positive items to overcome the method effect associated with negative items.

Method. A version A (positive and negative items) and a version B (only positive items) were proposed. A sample of 350 university students was collected for each version.

Results. The CFA shows that version A's one-dimensional model (A1) does not present adequate fit indices. It was also found that adding a specific factor for negative items (model A2) and another factor for positive items (model A3) does not improve the fit indices. A two-dimensional model (A4 model) did not improve the fit indices either. Compared to the one-dimensional model of version B, this presents a better fit compared to the original model (model A1).

Discussion and Conclusions. It is concluded that version B, a proposal with only positive items, adequately measures self-esteem since it does not have negative items and agrees with Rosenberg's original approach.

Key Words: Rosenberg Self-Esteem Scale, confirmatory factor analysis, self-esteem, construct validity, negative items, university students.

Introduction

Self-esteem is one of the most studied psychological constructs in the context of higher education since it is related to academic achievement (Hosseini et al., 2016), a better perception of learning (González-Vázquez, 2019), low levels of academic stress (Chávez & Peralta, 2019) and procrastination (Hajloo, 2014).

In this context, the Rosenberg Self-Esteem Scale (RSES) (Rosenberg, 1965) is one of the most used instruments and has been adapted to different countries. The adaptations have been made in adolescents, young people, adults, older adults, and secondary and higher education levels. However, there is no consensus on the factorial structure of the scale since, on the one hand, several studies support a two-dimensional model. For example, in older adults, a study carried out in Mexico reported adequate adjustment indices (*CFI*=.91, *RMSEA*=.07, *SRMR*=.73) (De León Ricardi & García, 2016). Adequate support for the two-dimensional model was also found in Mexican university students (*GFI*=.99, *CFI*=.98, *RMSEA*=.04) (Jurado et al., 2015). In Ireland, the two-dimensional structure was reported to be adequate (*GFI*=.97, *CFI*=.98, *RMSEA*=.07) (Kielkiewicz et al., 2020).

On the other hand, other studies support a unidimensional structure. For example, in an investigation carried out in 53 countries through exploratory factor analysis, support was given to the unifactorial solution (Schmitt & Allik, 2005). Additionally, studies conducted in Estonia (*GFI*=.98, *RMSEA*=.04) (Pullmann & Allik, 2000), Spain (*CFI*=.99, *RMSEA*=.09, *SRMR*=.05) (Martín-Albo et al., 2007) and Thailand (*CFI*=.97, *RMSEA*=.05, *SRMR*=.04) (Tinakon & Nahathai, 2012) show adequate values in the one-dimensional model.

However, several studies suggest that the unidimensionality of the RSES is achieved by controlling for the effect of negative items. Proof of this is the recent study in adolescents from Peru, where a better fit was evidenced by adding a specific factor for the negative items (*CFI*=.97, *RMSEA*=.047, *SRMR*=.33) (Sánchez-Villena et al., 2021). Similarly, in Brazil, adequate model fit indices were reported when a specific factor was added for negative items (*CFI*=.97, *TLI*=.94, *RMSEA*=.07) (de Lima & de Souza, 2019).

Faced with the problem of the factorial structure, Greenberger et al. (2003) carried out a study where they compared the original version of the RSES with a version of only positive

items and another version of only negative items. Their results showed greater support for the two-dimensional model (*CFI*=.95, *TLI*=94, *RMSEA*=.08), finding that the other versions did not show a good fit to the data.

However, the study evidenced methodological limitations, such as a sample size of fewer than 300 participants, which is not recommended when performing a CFA (Guadagnoli & Velicer, 1988). In addition, both versions did not show any evidence of content validity of the modified items; the latter should be considered as it allows reducing the irrelevant variance of the construct (Ventura-León, 2019). Performing it contributes to obtaining a greater source of validity, that is, to accumulate more evidence that supports the interpretation of the test scores for a specific purpose (Association American Educational Research et al., 2014).

On the other hand, in the study, the response categories of the RSES were modified from 4 to 6 categories, which differs from the original approach of Rosenberg (1965) and the various psychometric studies (de Lima & de Souza, 2019; Kielkiewicz et al., 2020; Sánchez-Villena et al., 2021).

Regarding the RSES reliability, most of the studies used Cronbach's Alpha coefficient, which is not adequate because it does not meet the assumptions for its calculation, such as tau-equivalence, that is, that all items present the same load factorial (Cho, 2016). Furthermore, when McDonald's omega coefficient was used, the factor loads obtained using a CFA were not used (Viladrich et al., 2017).

Given the above, the present study has the general objective of evaluating the factorial structure of the positive items version of the RSES since a positive items version is more suitable, mainly due to the absence of negative items. It is known that the inclusion of negative items produces measurement errors that alter the data analysis, one of the main consequences being the appearance of additional factors for the set of items (de Lima & de Souza, 2019; DiStefano & Motl, 2006). Likewise, it generates a reduction in the reliability values (Menold, 2019; Salazar, 2015; Suárez-Alvarez et al., 2018), which in turn contributes to the variability of the responses, which causes a lower variance (Suárez-Alvarez et al., 2018).

As specific objectives, the study aims to (a) evaluate the content validity of the original scale of positive and negative items and the scale of only positive items, (b) compare the

version of only positive items with three competing models, and (c) evaluate the internal consistency of both versions of the scale.

Método

Participants

For version A, a sample of 350 Peruvian university students of both sexes was obtained: 160 men (46%) and 190 women (54%) between the ages of 18 and 35 (M = 21.62, SD = 3.24). For version B, a sample of 350 Peruvian university students of both sexes was obtained: 129 men (37%) and 221 women (63%) between the ages of 18 and 35 (M = 20.87, SD = 3.08). The inclusion criteria were: participants studying at university, students over 18 years of age, and students who accept informed consent. While the exclusion criteria were: international students, students under 18 years of age, and students who are not currently studying.

Measure

The original scale was developed by Rosenberg (1965), who proposed a one-dimensional measure of self-esteem. It comprises ten items with four response options (Strongly agree=4, Agree=3, Disagree=2, Strongly disagree=1). For the present study, the items with the highest construct representation of the adapted versions of the RSES in Colombia (Ceballos-Ospino et al., 2017), Chile (Rojas-Barahona et al., 2009), and Peru (Ventura-León et al., 2018) were selected for use in the Peruvian context. The criteria used to select the items were clarity in the content of the reagents and relevance to measuring the construct in the Peruvian context. In this way, two versions of the scale were obtained, forms A and B, each evaluated by six psychologists specialized in the subject.

Procedure

The ethics committee of the Universidad Peruana Unión approved the study. In addition, the study strictly followed the standards established in the Helsinki code (World Medical Association, 2013). To study the factorial structure of the RSES, two versions were considered: version A, which contains positive and negative items, and version B, which only contains positive items. The items used in each version were extracted from the RSES adaptations in Latin America. First, both versions were evaluated by twelve psychologists (six for each version) with extensive knowledge of the construct. Second, the corrected scale of both versions was applied to a pilot group of 20 university students (ten for each version) where no

modification was made to the items. The objective of the pilot test was to identify any difficulty in understanding items. Third, the final scale of both versions was applied through a virtual form made on the Google Forms digital platform that was distributed on social networks (Facebook and Whatsapp). Only participants who completed the informed consent could complete the other sections of the virtual form. Additionally, the purpose of the study and the confidentiality of the data collected were indicated on the form.

Statistical Analyses

For content validity, Aiken's V coefficient was used, where values greater than .70 show a greater consensus of the judges on the suitability of the items (Aiken, 1980). For the descriptive analysis, descriptive statistics were calculated, where asymmetry and kurtosis indicate good values when they are less than ±1.5 (Forero et al., 2009). For the CFA, the WLSMV estimator was used, due to the ordinal nature of the items (Brown, 2015) and to evaluate the fit of the model, the coefficients *CFI* (>.95), *TLI* (>.95), *RMSEA* (<.08), *SRMR* (<.08) and *WRMR* (< 1.00), the latter was used because it is adequate when using ordinal data (DiStefano, Liu, Jiang, y Shi, 2018). To assess the fit indices, we worked with the proposal by Hu y Bentler (1999) and DiStefano, Liu, Jiang, y Shi (2018). All statistical analyzes were developed using the R program (version 3.5.0) (R Core Team, 2019) and the RStudio Team (2018).

Results

Content Validity

Table 1 shows that versions A and B present good values in the criteria of clarity, relevance, coherence, and context (V > .71). On the other hand, the items were modified following the judges' suggestions in versions A (item 5) and B (item 3, item 5, item 9, and item 10).

Table 1. *Content validity of version A and B.*

Items	Version A	V (Rele)	V (Cohe)	V (Clar)	V (Cont)	Version B	V (Rele)	V (Cohe)	V (Clar)	V (Cont)
1	Me siento una persona tan valiosa como las otras.	0.94	1.00	0.94	1.00	Me siento una persona tan valiosa como las otras.	0.89	1.00	0.89	0.89
2	Creo que tengo cualidades positivas.	0.94	0.94	0.89	0.94	Creo que tengo cualidades positivas.	0.94	0.94	0.78	0.83
3	En general, pienso que soy un fracaso.	1.00	0.94	1.00	1.00	En general, pienso que soy bueno para **varias** cosas.	0.89	1.00	0.89	0.89
4	Soy capaz de hacer las cosas tan bien como los demás.	1.00	1.00	1.00	1.00	Soy capaz de hacer las cosas tan bien como los demás.	0.83	1.00	0.89	0.89
5	**Pienso que no tengo muchos motivos para sentirme orgulloso/a de mí**.	0.94	0.89	0.83	0.94	**Pienso que tengo muchos motivos para sentirme orgulloso/a de mí.**	1.00	1.00	0.83	0.89
6	Tengo una actitud positiva hacia mí mismo.	1.00	1.00	1.00	1.00	Tengo una actitud positiva hacia mí mismo.	0.89	0.94	0.72	0.94
7	En general, me siento bien conmigo mismo.	1.00	1.00	1.00	1.00	En general, me siento bien conmigo mismo.	0.94	0.94	0.94	0.94
8	Desearía tener una mejor valoración de mí mismo.	0.94	0.94	0.94	0.94	Tengo una buena valoración de mí mismo.	0.83	0.83	0.89	0.94
9	Realmente me siento inútil en algunas ocasiones.	1.00	1.00	1.00	1.00	Realmente me siento útil en **varias** ocasiones.	0.89	0.94	0.89	0.89
10	A veces pienso que no sirvo para nada.	0.89	1.00	1.00	1.00	**Pienso** que soy **capaz** para muchas cosas.	0.89	0.89	0.89	0.89

Descriptive analyzes of the items

Table 2 shows the descriptive statistics of version A, where the mean of the items varies between 2.02 and 3.31. As for the standard deviation, it ranges between .68 and .99. In addition, the kurtosis and asymmetry present values less than ±1.5. Regarding version B, the mean is in a range of 3.22 to 3.36 in the items. Regarding the standard deviation, the value ranged between .68 and .74. Likewise, the kurtosis and asymmetry of most of the items present adequate values (<±1.5).

Table 2. *Exploratory item analysis.*

	Version A				Version B			
Items	M	SD	g1	g2	M	DS	g1	g2
Item 1	3.20	.77	-.78	.27	3.36	.72	-1.28	2.18
Item 2	3.31	.68	.57	.13	3.34	.69	-1.10	1.72
Item 3	2.02	.78	.57	.13	3.22	.74	-.88	.82
Item 4	3.12	.74	-.45	-.27	3.24	.69	-.88	1.33
Item 5	2.22	.94	.29	-.81	3.29	.76	-1.02	.82
Item 6	3.00	.78	-.44	-.21	3.28	.72	-.91	.82
Item 7	2.93	.78	-.48	-.01	3.25	.72	-.79	.54
Item 8	2.99	.88	-.69	-.08	3.26	.71	-.86	.88
Item 9	2.56	.94	-.21	-.85	3.26	.68	-.93	1.56
Item 10	2.11	.99	.45	-.88	3.33	.71	-1.11	1.75

Nota: *M* = Median; *SD* = Standard deviation; *g1* = Asymmetry; *g2* = Kurtosis.

Confirmatory Factor Analysis

As can be seen in Table 3, in version A, Model A1 (original model) presented adequate values in most of the fit indices (*CFI*= .946, *TLI*=.930, *RMSEA*=.165 [IC 90%: .150–.180], *SRMR*=.060). Likewise, Model A2 was proposed, where a specific factor was added for negative items. This model showed some improvements in most of the fit indices (*CFI*= .977, *TLI*=.965, *RMSEA*=.117 [IC 90%: .101–.134], *SRMR*=.038). Also, Model A3 was proposed, where a specific factor was added for the positive items. In this model, it was shown that the adjustment indices were mostly satisfactory (*CFI*= .972, *TLI*=.959, *RMSEA*=.127 [IC 90%: .111–.144], *SRMR*=.039). Finally, the fit of Model A4 was evaluated, with two correlated dimensions, which showed improvements in most of the fit indices, unlike the previous models (*CFI*= .976, *TLI*=.968, *RMSEA*=.111 [IC 90%: .095–.127], *SRMR*=.040). Likewise, it was found that the evaluated models of version A showed value that exceeded the cut-off point in *χ2 /gl*, *RMSEA*, and *WRMR*.

Against these results, the fit of version B was evaluated, which presented good values in most of the fit indices (*CFI*= .975, *TLI*=.968, *RMSEA*=.138 [IC 90%: .123–.154],

SRMR=.049). However, similar to the results of the version A models, the values were higher in *χ2 /gl*, *RMSEA*, and *WRMR*. It is essential to point out that the factorial weights of the items in version B are mostly better than the models proposed in version A.

Table 3. *Model fit indices and standardized parameter estimates.*

Models	X^2	*gl*	*p*	X^2 /*gl*	*RMSEA* [IC 90%]	*SRMR*	*CFI*	*TLI*	*WRMR*	*α*	*ω*
Model A1	366.74	35	.000	10.48	.165 [.150–.180]	.060	.946	.930	1.65	-.093	.036
Model A2	174.03	30	.000	5.80	.117 [.101–.134]	.038	.977	.965	1.04		
Model A3	199.35	30	.000	6.65	.127 [.111–.144]	.039	.972	.959	1.07		
Model A4	180.12	34	.000	5.30	.111 [.095–.127]	.040	.976	.968	1.08		
Model B	267.53	35	.000	7.64	.138 [.123–.154]	.049	.975	.968	1.51	.965	.960

Items	Model A1	Model A2		Model A3			Model A4		Model B
	λ	*λ*		*λ*			*λ*		*λ*
Item 1 (+)	.79	.82		.69	.43	.82			.82
Item 2 (+)	.72	.74		.59	.48	.74			.84
Item 4 (+)	.72	.74		.61	.44	.74			.85
Item 6 (+)	.89	.92		.77	.49	.92			.92
Item 7 (+)	.87	.89		.74	.50	.89			.88
Item 3 (-)	-.82	-.74	.42	-.88			.88		.84
Item 5 (-)	-.61	-.55	.32	-.65			.65		.87
Item 8 (-)	-.47	-.41	.32	-.50			.50		.89
Item 9 (-)	-.74	-.62	.55	-.79			.79		.90
Item 10 (-)	-.84	-.74	.51	-.89			.89		.88

Model A1 = Original one-dimensional model, Model A2 = Model with effect control for negative items, Model A3 = Model with effect control for positive items, Model A4 = Model with two related factors, Model B = One-dimensional model with only items positive

Discussion and Conclusions

The study's general objective was to evaluate the factorial structure of the RSES scale and propose a new version of the RSES with only positive items to overcome the method effect associated with negative items. Two versions of the same scale were proposed: version A, which contains positive and negative items, and version B, which has only positive items.

A CFA was carried out to evaluate the RSES´s factorial structure, where the models of versions A and B were compared. Model A1, which follows the original approach of Rosenberg (1965), presented good values in most of the adjustment indices (*CFI*= .95, *TLI*=.93). However, it obtained an unacceptable value in the *RMSEA*=.165 and *WRMR*. This result does not coincide with that reported in previous studies (Martín-Albo et al., 2007; Pullmann & Allik, 2000; Tinakon & Nahathai, 2012), where a one-dimensional model does fit the data.

About this, the adjustment of models A2 (model to control the method effect of negative items) and A3 (model to control the method effect of positive items) was evaluated, where an improvement in the indices was evidenced of adjustment, since a specific factor was added for the positive and negative items, however, high values were found in the *RMSEA* and *WRMR*. These findings do not agree with other studies where they support that adding factors favors the adjustment of the data (de Lima & de Souza, 2019; Sánchez-Villena et al., 2021).

The A4 model showed a better fit than the previous models (*CFI*=.98, *TLI*=.97, *RMSEA*=.111 [IC 90%: .095 – .127]), despite this, as happened in the previous models from version A, high coefficients of *RMSEA* and *WRMR* were found. Thus, this does not coincide with previous studies (De León Ricardi & García, 2016; Jurado et al., 2015; Kielkiewicz et al., 2020).

Finally, the proposal of model B was evaluated, with only positive items, which showed similar fit indices and values higher than .80 in factor loadings than the models of version A. In addition, model B showed a better fit to the data than the study by Greenberger et al. (2003), where the fit indices were mostly lower in the model of the positive items. On the other hand, model B shows an advantage over the models of version A due to the absence of negative items. Several studies have reported that using negative items decreases reliability values (Menold, 2019; Salazar, 2015; Suárez-Alvarez et al., 2018), influences the variability of the responses obtained, producing less variance (Suárez-Alvarez et al., 2018), and causes lower scores for negative items (Weems et al., 2003). In addition, it generates that the one-dimensional models do not adequately fit the data and cause other factors (Suárez-Alvarez et al., 2018; Woods, 2006). Another advantage of model B is that it is no longer necessary to add a factor in the CFA models to control the method effect of negative items. Thus improv-

ing the practice of adding additional factors when studying the factorial structure of the scale (de Lima & de Souza, 2019; Sánchez-Villena et al., 2021).

It is essential to point out that the problems associated with negative items do not occur in the same way in all cultures (Wong et al., 2003). Thus, in North America, the combined use of positive and negative items on the same scale works well (Wong et al., 2003). In contrast, in Latin American countries, this practice generates inconsistencies in people's responses (Marin et al., 1992). This could explain why the original version of the scale developed in the United States fits well to a one-dimensional model. Moreover, because in the studies carried out in Latin America, the original approach of negative and positive items has generated the appearance of additional factors to control the negative items (de Lima & de Souza, 2019; Sánchez-Villena et al., 2021) or the appearance of two-dimensional models (De León Ricardi & García, 2016; Jurado et al., 2015).

Taking the above into account, another advantage of model B is that its unidimensionality is consistent with the theoretical proposal of Rosenberg (1965), who classifies people within a single continuous range that varies from very low to very high self-esteem. Due to the above, it is decided to choose the proposal of model B of only positive items for the evaluation of self-esteem.

Regarding the reliability of the models of version A, lower values were found in Cronbach's alpha and the omega coefficient; these values are mainly because the negative items decrease the reliability values in the scale (Menold, 2019; Salazar, 2015; Suárez-Alvarez et al., 2018). On the other hand, when evaluating the reliability of model B, it was found that the values in the indices were acceptable (α= .965, ω= .960).

The study shows some limitations. In the first place, convenience sampling was carried out, which prevents generalizing the results obtained in the study. Secondly, a factorial invariance study was not carried out according to the sex of the participants, since in the sample of model B, there was a significant difference between the number of men and women, which would generate altered results; likewise, there are several studies that have provided evidence that this variable can affect the scale scores (Caballo & Salazar, 2018; Ruiz-González et al., 2018). Thirdly, we worked with university students from some public and private universities; therefore, it is suggested for future studies to expand the sample size.

It is concluded that version B, a proposal of only positive items, supposes an improvement in measuring self-esteem since it does not have negative items and agrees with Rosenberg's unidimensional approach. This version was called the Rosenberg Self-Esteem Scale-P (RSES-P).

References

Aiken, L. R. (1980). Content Validity and Reliability of Single Items or Questionnaires. *Educational and Psychological Measurement*, *40*(4), 955–959. https://doi.org/10.1177/001316448004000419

Association American Educational Research, [AEA], Association American Psychological, [APA], & Education National Council on Measurement in education, [NCME]. (2014). *Standards for educational and psychological testing*. American Educational Research Association.

Brown, T. A. (2015). *Confirmatory Factor Analysis for Applied Research* (Second Edi).

Caballo, V. E., & Salazar, I. C. (2018). La autoestima y su relación con la ansiedad social y las habilidades sociales. *Psicología Conductual*, *26*(1), 23–53.

Ceballos-Ospino, G. A., Paba-Barbosa, C., Suescún, J., Oviedo, H. C., Herazo, E., & Campo-Arias, A. (2017). Validez y dimensionalidad de la escala de autoestima de Rosenberg en estudiantes universitarios. *Pensamiento Psicológico*, *15*(2), 29–39. https://doi.org/http://dx.doi.org/10.11144/Javerianacali.PPSI15-2.vdea

Chávez, J., & Peralta, R. (2019). Estrés académico y autoestima en estudiantes de enfermería , Arequipa-Perú. *Revista de Ciencias Sociales*, *25*(1), 384–399.

Cho, E. (2016). Making Reliability Reliable. *Organizational Research Methods*, *19*(4), 651–682. https://doi.org/10.1177/1094428116656239

De León Ricardi, C., & García, M. (2016). Escala de Rosenberg en población de adultos mayores. *Ciencias Psicológicas*, *10*(2), 127. https://doi.org/10.22235/cp.v10i2.1245

de Lima, T. J. S., & de Souza, L. E. C. (2019). Rosenberg self-esteem scale: Method effect and gender invariance. *Psico-USF*, *24*(3), 517–528. https://doi.org/10.1590/1413-82712019240309

DiStefano, C., Liu, J., Jiang, N., & Shi, D. (2018). Examination of the Weighted Root Mean Square Residual: Evidence for Trustworthiness? *Structural Equation Modeling*, *25*(3), 453–466. https://doi.org/10.1080/10705511.2017.1390394

DiStefano, C., & Motl, R. W. (2006). Further investigating method effects associated with negatively worded items on self-report surveys. *Structural Equation Modeling*, *13*(3), 440–464. https://doi.org/10.1207/s15328007sem1303_6

Forero, C. G., Maydeu-Olivares, A., & Gallardo-Pujol, D. (2009). Factor Analysis with Ordinal

Indicators: A Monte Carlo Study Comparing DWLS and ULS Estimation. *Structural Equation Modeling: A Multidisciplinary Journal*, *16*(4), 625–641. https://doi.org/10.1080/10705510903203573

González-Vázquez, B. (2019). ¿Cómo mejorar el aprendizaje? Influencia de la autoestima en el aprendizaje del estudiante universitario. *Revista Complutense de Educación*, *30*(3), 781–795. https://doi.org/10.5209/rced.58899

Greenberger, E., Chen, C., Dmitrieva, J., & Farruggia, S. P. (2003). Item-wording and the dimensionality of the Rosenberg Self-Esteem Scale: Do they matter? *Personality and Individual Differences*, *35*(6), 1241–1254. https://doi.org/10.1016/S0191-8869(02)00331-8

Guadagnoli, E., & Velicer, W. F. (1988). Relation of Sample Size to the Stability of Component Patterns. *Psychological Bulletin*, *103*(2), 265–275. https://doi.org/10.1037/0033-2909.103.2.265

Hajloo, N. (2014). Relationships between self-efficacy, self-esteem and procrastination in undergraduate psychology students. *Iranian Journal of Psychiatry and Behavioral Sciences*, *8*(3), 42–49.

Hosseini, S. N., Mirzaei Alavijeh, M., Matin, B. K., Hamzeh, B., Ashtarian, H., & Jalilian, F. (2016). Locus of control or self-esteem; Which one is the best predictor of academic achievement in Iranian college students. *Iranian Journal of Psychiatry and Behavioral Sciences*, *10*(1). https://doi.org/10.17795/ijpbs-2602

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Jurado, D., Jurado, S., López, K., & Querevalú, B. (2015). Validez de la Escala de Autoestima de Rosenberg en universitarios de la Ciudad de México. *Revista Latinoamericana de Medicina Conductual*, *5*(1), 18–22. http://www.revistas.unam.mx/index.php/rlmc/article/view/55207/50020

Kielkiewicz, K., Mathúna, C., & McLaughlin, C. (2020). Construct Validity and Dimensionality of the Rosenberg Self-Esteem Scale and Its Association with Spiritual Values Within Irish Population. *Journal of Religion and Health*, *59*(1), 381–398. https://doi.org/10.1007/s10943-019-00821-x

Kwek, A., Bui, H. T., Rynne, J., & So, K. K. F. (2013). The Impacts of Self-Esteem and Resilience on Academic Performance: An Investigation of Domestic and International Hospitality and Tourism Undergraduate Students. *Journal of Hospitality and Tourism Education*, *25*(3), 110–122. https://doi.org/10.1080/10963758.2013.826946

Marin, G., Gamba, R. J., & Marin, B. V. (1992). Extreme Response Style and Acquiescence among Hispanics: The Role of Acculturation and Education. *Journal of Cross-Cultural Psychology*, *23*(4), 498–509. https://doi.org/10.1177/0022022192234006

Martín-Albo, J., Núñez, J. L., Navarro, J. G., & Grijalvo, F. (2007). The Rosenberg self-esteem scale: Translation and validation in university students. *Spanish Journal of Psychology*, *10*(2), 458–467. https://doi.org/10.1017/S1138741600006727

Menold, N. (2019). How Do Reverse-keyed Items in Inventories Affect Measurement Quality and Information Processing? *Field Methods*, *32*(2), 140–158. https://doi.org/10.1177/1525822X19890827

Pullmann, H., & Allik, J. üri. (2000). The Rosenberg Self-Esteem Scale: Its dimensionality, stability and personality correlates in Estonian. *Personality and Individual Differences*, *28*(4), 701–715. https://doi.org/10.1016/S0191-8869(99)00132-4

R Core Team. (2019). *A language and environment for statistical computing*. R Foundation for Statistical Computing.

Rojas-Barahona, C. A., Zegers P, B., & Förster M, C. E. (2009). La escala de autoestima de Rosenberg: Validación para Chile en una muestra de jóvenes adultos, adultos y adultos mayores. *Revista Médica de Chile*, *137*(6), 791–800. https://doi.org/10.4067/s0034-98872009000600009

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton University Press.

Rosli, Y., Othman, H., Ishak, I., Lubis, S. H., Saat, N. Z. M., & Omar, B. (2012). Self-esteem and Academic Performance Relationship Amongst the Second Year Undergraduate Students of Universiti Kebangsaan Malaysia, Kuala Lumpur Campus. *Procedia - Social and Behavioral Sciences*, *60*, 582–589. https://doi.org/10.1016/j.sbspro.2012.09.426

Ruiz-González, P., Medina-Mesa, Y., Zayas, A., & Gómez-Molinero, R. (2018). Relación entre la autoestima y la satisfacción con la vida en una muestra de estudiantes universitarios. *Revista INFAD de Psicología*, *2*(1), 67. https://doi.org/10.17060/ijodaep.2018.n1.v2.1170

Saadat, M., Ghasemzadeh, A., & Soleimani, M. (2012). Self-esteem in Iranian university students and its relationship with academic achievement. *Procedia - Social and Behavioral Sciences*, *31*, 10–14. https://doi.org/10.1016/j.sbspro.2011.12.007

Salazar, M. S. (2015). El dilema de combinar ítems positivos y negativos en escalas. *Psicothema*, *27*(2), 192–199. https://doi.org/10.7334/psicothema2014.266

Sánchez-Villena, A. R., De La Fuente-Figuerola, V., & Ventura-León, J. (2021). Modelos factoriales de la Escala de Autoestima de Rosenberg en adolescentes peruanos. *Revista de Psicopatología y Psicología Clínica*, *26*(1), 47. https://doi.org/10.5944/rppc.26631

Schmitt, D. P., & Allik, J. (2005). Simultaneous administration of the Rosenberg self-esteem scale in 53 nations: Exploring the universal and culture-specific features of global self-esteem. In *Journal of Personality and Social Psychology*, *89*(4), 623–642. https://doi.org/10.1037/0022-3514.89.4.623

Suárez-Alvarez, J., Pedrosa, I., Lozano, L. M., García-Cueto, E., Cuesta, M., & Muñiz, J. (2018). Using reversed items in likert scales: A questionable practice. *Psicothema*, *30*(2), 149–158.

https://doi.org/10.7334/psicothema2018.33

Team, Rs. (2020). *R, RStudio: Integrated Development Environment for R* (RStudio, I). http://www.rstudio.com/

Tinakon, W., & Nahathai, W. (2012). A comparison of reliability and construct validity between the original and revised versions of the Rosenberg Self-Esteem Scale. *Psychiatry Investigation*, *9*(1), 54–58. https://doi.org/10.4306/pi.2012.9.1.54

Ventura-León, J. (2019). De regreso a la validez basada en el contenido. *Adicciones*, *20*(01). https://doi.org/10.20882/adicciones.1213

Ventura-León, J., Caycho-Rodriguez, T., & Barboza-Palomino, M. (2018). Evidencias psicométricas de la escala de autoestima de Rosenberg en adolecentes limeños. *Revista Interamericana de Psicología/Interamerican Journal of Psychology*, *52*(1), 44–60. https://doi.org/10.30849/RIP/IJP.V52I1.363

Viladrich, C., Angulo-Brunet, A., & Doval, E. (2017). Un viaje alrededor de alfa y omega para estimar la fiabilidad de consistencia interna. *Anales de Psicologia*, *33*(3), 755–782. https://doi.org/10.6018/analesps.33.3.268401

Weems, G. H., Onwuegbuzie, A. J., Schreiber, J. B., & Eggers, S. J. (2003). Characteristics of respondents who respond differently to positively and negatively Worded items on Rating Scales. *Assessment and Evaluation in Higher Education*, *28*(6), 587–606. https://doi.org/10.1080/0260293032000130234

Wong, N., Rindfleisch, A., & Burroughs, J. E. (2003). Do Reverse-Worded Items Confound Measures in Cross-Cultural Consumer Research? The Case of the Material Values Scale. *Journal of Consumer Research*, *30*(1), 72–91. https://doi.org/10.1086/374697

Woods, C. M. (2006). Careless responding to reverse-worded items: Implications for confirmatory factor analysis. *Journal of Psychopathology and Behavioral Assessment*, *28*(3), 189–194. https://doi.org/10.1007/s10862-005-9004-7

World Medical Association. (2013). World Medical Association Declaration ofHelsinki Ethical Principles for Medical Research Involving Human Subjects. *Clinical Review & Education*, *310*(20), 2191–2194. https://doi.org/10.1001/jama.292.11.1359

Received: 06-05-2021
Accepted: 10-08-2021