

Adaptation, Validation, Reliability and Factorial Equivalence of the Rosenberg Self-Esteem Scale in Colombian and Spanish Population

Mayra Gómez-Lugo¹, José P. Espada², Alexandra Morales², Laurent Marchal-Bertrand¹, Franklin Soler³ and Pablo Vallejo-Medina¹

¹ *Fundación Universitaria Konrad Lorenz (Colombia)*

² *Universidad Miguel Hernandez (Spain)*

³ *Universidad del Rosario (Colombia)*

Abstract. The Rosenberg Self-Esteem Scale is the most widely used instrument to assess self-esteem. In light of the absence of adaptations in Colombia, this study seeks to validate and adapt this scale in the Colombian population, and perform factorial equivalence with the Spanish version. A total of 1,139 seniors (633 Colombians and 506 Spaniards) were evaluated; the individuals answered the Rosenberg Self-Esteem Scale and sexual self-esteem scale. The average score of the items was similar to the questionnaire's theoretical average, and standard deviations were close to one. The psychometric properties of the items are generally adequate with alphas of .83 and .86 and significant ($CI = .95$) and correlations with the sexual self-esteem scale ranging from .31 and .41. Factorial equivalence was confirmed by means of a structural equation model ($CFI = .912$ and $RMSEA = .079$), thus showing a strong level of invariance.

Received 11 August 2015; Revised 2 August 2016; Accepted 4 August 2016

Keywords: Colombia, reliability, Rosenberg, self-esteem, validation.

General self-esteem is one of the most widely studied constructs in social sciences (Bachman, O'Malley, Freedman-Doan, Trzesniewski, & Donnellan, 2011), and its measurement one of the most amply discussed topics in Psychology (Rosenberg, Schooler, Schoenbach, & Rosenberg, 1995). Since its inception with the tenets of William James in 1890, self-esteem has been approached from different perspectives. In 1965, Rosenberg considered a component of self-concept and defined it as a single set of thoughts and feelings about the value or importance of each individual; in other words, it is the overall positive or negative attitude people have of themselves. Rosenberg described a person with good self-esteem as someone who defines him/herself as valuable, and appreciates his/her merits, but also recognizes his/her faults. More recent studies from different theoretical perspectives complement and/or postulate the definition of general self-esteem. For instance, González-Pienda, Núñez, González-Pumariega, and García (1997) report that general self-esteem is the way in which individuals appreciate themselves due to the social feedback they receive in their various roles. Luciano, Gómez, and Valdivia (2002) also define

self-esteem as verbal feedback (or assessment) which the person relates to his/her actions, and which is enhanced with the locus of control and the types of functions they perform.

Self-esteem is relevant in that it is a representative indicator of health and well-being, as well as an explanatory variable of human behavior (Rosenberg, 1965). For example, it has been found that low self-esteem – especially in teens – is a great predictor of deterioration in mental health (Marshall, Parker, Ciarrochi, & Heaven, 2014) and is related to eating disorders (Brechan & Kvaem, 2015), suicide attempts (Wichstrom, 2000), difficulty developing positive support networks (Marshall et al., 2015), among others. In turn, it has been observed that high self-esteem is related, for example, to quality of life (Muñoz & Alonso, 2013), increased academic performance in adolescents (Fiz & Oyon, 1998), adaptation to the social environment (Silbereisen & Wiesner, 2002), and emotional stability (Zeigler-Hill et al., 2015). However, it is important to note that high or low self-esteem is not necessarily related to positive or negative effects on health, respectively (Cheng, Govorun, & Chartrand, 2012).

Given the importance of self-esteem, a number of instruments have been proposed in order to measure it. The Rosenberg Self-Esteem Scale (RSES; Rosenberg, 1965) is the most widely validated scale (Gray-Little, Williams, & Hancock, 1997) and the most widely used scale worldwide to evaluate self-esteem (Robins, Hendin, & Trzesniewski, 2001). The original RSES is a

Correspondence concerning this article should be addressed to Pablo Vallejo-Medina. Phone: +57-31851649.

E-mail: pablo.vallejom@konradlorenz.edu.co

This work has been possible, thanks to "Fundación Universitaria Konrad Lorenz" funding associated to the research project (number: 55270151) granted to last author.

unidimensional scale that reports a global index of self-esteem. This scale has a Likert response scale, and its response options range from “strongly disagree” to “completely agree”. A high score on the RSES is interpreted as high self-esteem. The high level of acceptance of the RSES lies on the fact that it not only shows high levels of validity and reliability in different population groups (Huang & Dong, 2012), the language used in the scale is simple, requires only basic reading skills, can be administered within minutes and the items are closely related to the construct (Gray-Little et al., 1997). In the empirical literature exist a disagreement about the underlying factor structure of the Rosenberg Self-Esteem Scale. Inconsistency main problems are focused on whether or not the scale assesses global self-esteem as one factor (as Rosenberg proposed) composed of 5 items written in negative and 5 items in positive, or as two distinct constructs representing the positive and the negative aspects of the self-esteem (Hyland, Boduszek, Dhingra, Shevlin, & Egan, 2014; McKay, Boduszek, & Harvey, 2014).

To date, the RSES has been translated and adapted in several languages including: Canadian French, Persian, Chinese, Italian, Estonian, German and Portuguese. Additionally, although the RSES has been validated in Spanish for Spain, Chile and Argentina, and although it is a scale which is commonly used in Colombia (see: Harper et al., 2014; Milanés & Gómez-Bustamante, 2011; Perez-Olmos, Tellez, Velez, & Ibáñez-Pinilla, 2012; inter alia), there is currently no validated adaptation in Colombia. Besides, cultural comparisons of self-esteem have been widely used (Li, Delvecchio, Di Riso, Salcuni, & Mazzeschi, 2015). In fact, implicit self-esteem does not allow for cultural comparison, although explicit self-esteem – addressed here – does (Falk, Heine, Takemura, Zhang, & Hsu, 2015). Thus, the use of unbiased evaluations between cultures is recommended for this purpose (Li et al., 2015).

Consequently, the aim of this instrumental research, was to validate and adapt the Rosenberg Self Esteem Scale (1965) in Colombians, and also to show their factorial equivalence with the Spanish version.

Method

Participants

The sample of this study is divided into two groups:

The first group includes the experts; four psychologists with at least a master degree were selected – this is the number of grad-level experts recommended by Lynn (1986), and they performed the cultural adaptation from the Spanish of Spain to Spanish of Colombia. All of these psychologists are Colombian nationals and residents, and they reported having lived at least two years in Spain. In addition, another four Colombian experts in psychology and/or psychometrics partook

in the evaluation of the qualitative properties of the items (see procedure).

The second group was a Spanish and Colombian community sample. This group included an initial sample which consisted of a total of 1,797 participants. These participants were selected by convenience, trying to compensate for sex and age by selection. Of these – and in pursuance of inclusion criteria – a final sample of 1,139 participants that met the inclusion criteria were selected, including 633 Colombians and 506 Spaniards. Inclusion criteria in the study were being of age, accepting to partake in the study and having a Colombian or Spanish nationality, as appropriate.

Statistically significant differences were found between Spain and Colombia in terms of age, years of education, marital status and religion, as expected in community samples. No differences were observed in sex, sexual orientation and relationship. The age range for Colombia ranged between 18 and 72 years (Mean (M) = 32.21; Standard Deviation (SD) = 10.68) and in the case of Spain the range was 18 to 70 years (M = 33.93, SD = 12.71), $t(1137) = 2.49$, $p < .01$, $d = .14$. Following are the values observed with regard to years of study: $t(1130) = 4.22$, $p < .01$, $d = 0.24$; in Colombia (M = 16.72; SD = 2.81) and in Spain (M = 15.86; SD = 3.98). Differences in marital status were $\chi^2(3, 1125) = 20.63$; $p < .01$, $\eta^2 = .016$ (there were more Colombians in marriage by habit and repute – an uncommon choice in Spain). Following are the values for religion: $\chi^2(5, 1139) = 186.71$, $p < .01$, $\eta^2 = .16$ (Colombians attend a greater number of times to religious events). However, no significant differences were observed in terms of sex by country: $\chi^2(1, 1124) = .37$, $p = .57$; sexual orientation $\chi^2(8, 1124) = 10.23$; $p = .24$ or marital status $\chi^2(8, 1124) = 10.23$; $p = .24$.

Instruments

Sociodemographic information

Different sociodemographic characteristics of participants were evaluated using a self-administered semi-structured interview. Age, sex, educational level, sexual orientation (using the scale Kinsley), religion, and marital status were evaluated.

The rosenberg self-esteem scale (RSES; Rosenberg, 1965)

The starting point was the version validated by Martín-Albo, Núñez, Navarro, and Grijalvo (2007) for Spain. The scale consists of 10 items assessing general self-esteem. Items 1, 3, 4, 7 and 10 are positive and items 2, 5, 6, 8, and 9 are negative. The questionnaire has a Likert response scale wherein items are answered on a four-point scale (1 = *Strongly disagree*, 2 = *Disagree*, 3 = *Agree*, 4 = *Strongly agree*). The total score therefore ranges from 10 to 40 points; higher scores indicate good self-esteem.

Sexuality scale (SS; Snell & Papini, 1989)

The *Sexual Self-esteem* dimension of the short 15-item version (Wiedeman & Allgeier, 1993) validated for Colombia (Soler et al., 2016) was used for this study. Sexual Self-esteem is the tendency to positively self-assess the ability to relate sexually with a partner. The Sexual Self-Esteem subscale were answered on a 5-option Likert scale ranging from: *Disagree* to *Agree*. Adequate Cronbach alphas (.82 in Colombia and .87 in Spain) have been observed in this study. A sample item is: *"I think of myself as a very good sexual partner."*

Procedure

Cultural adaptation of the RSES was conducted following the guidelines for the adaptation of scales of the same language and cultures by Vallejo-Medina et al. (2016), following the guidelines of Elosua, Mujika, Almeida, and Hermsilla (2014), Muñiz, Elosua, and Hambleton (2013), as well as AERA, APA, and NCME, (2014). Adaptation was performed by four Colombian grad-level psychologists who reported having lived in Spain for at least two years.

The properties of the items in the adapted version of the scale were evaluated by four experts in psychometrics and/or psychology. The properties evaluated were: *Representativeness*: contribution of the item to the construct (global or general self-esteem). *Comprehension*: this property determines whether the item in its adapted version is properly understood. *Interpretation*: Level of ambiguity of the item. *Clarity*: how direct and concise the item is. Each expert was provided with the operational definition of the construct to assess –general self-esteem – and the battery of items created, and they were asked to evaluate the properties of each item, by means of a Likert scale, ranging from 1 (*nothing ... and "the feature"*) to 4 (*very ... and "the feature"*). Experts also contributed by providing an alternative wording where deemed necessary.

The battery of tests was administered simultaneously in Colombia and Spain between October the 23rd, 2014 and February the 24th, 2015. Sampling was opportunistic. The questionnaires were entered in Typeform[®] and could be filled out in PCs or laptops, tablets and smartphones. Distribution was carried out through personal contacts, Facebook[®] and Twitter[®]. For the convenience sampling, an online survey was used due to the similarity of answers observed between paper and online samples (de Bernardo & Curtis, 2013).

Data analysis

A table of specifications of the items and the ICAiken program were used for qualitative analysis, thus obtaining the confidence interval for Aiken's V. Criteria considered were those by Merino and Livia (2009), taking a score

below .50 at the lower limit (CI = 95%) as the cutoff point for Aiken's V and a criterion of inadequacy of the item. However, all comments of the experts were taken into account.

The SPSS 20.0 was used to analyze the psychometric properties of the items. The EQS 6.1 was used to calculate Factorial Invariance, which was evaluated progressively under a Mean and Covariances Structures (MACS) procedure (Byrne et al., 2009). To begin, the multivariate distribution of data was calculated using Mardia's test, wherein values above 5 are indicative of non-normality. The test indicated results of 36.5 in Colombia and 47.4 in Spain; therefore the Maximum Likelihood Robust method (ML, Robust) was used as the estimation method as the parameter of normal distribution was not met. In terms of progressive factorial invariance, the following steps were evaluated: *configural invariance*: no restrictions in the model; *metric – or weak – invariance*: the factorial weights are restricted, thus evaluating the equivalence of the weight of each item regarding the factor; *strong invariance*: intercepts are restricted; and *strict invariance*: the variances of errors are restricted. The indices taken into account in order to evaluate the fit of the models were as follows: the Root Mean Square Error Approximation (RMSEA), the confidence interval – at 90% – and the Comparative Fit Index (CFI). Values below .08 for the RMSEA and greater than .95 for the CFI were considered indicative of good fit. A decrease no greater than .01 with regard to the least restrictive model was taken as evidence of invariance (Cheung & Rensvold, 2002). The Akaike Information Criterion (AIC) was also taken into account in the model selection, which indicates absence of FI should the increase with regard to the least restrictive model be considerable.

Results

Item analysis

Table 1 presents the qualitative evaluation performed by four experts in psychometrics and/or psychology on the 10 items of the Colombian version. The lower limit of 95% in Aiken's V is always greater than .50. Therefore, it can be observed that there is adequate wording of the items of the Colombian Spanish version of the RSES.

Analysis of the items' psychometric properties and reliability

Table 2 shows that the indicators analyzed are generally adequate. Both versions are reliable, and have with very similar values in both countries. Corrected item-total correlations (r_{it}^c) are always greater than .30, except in item 8 of the Colombian version. No significant increase in Cronbach's alpha is observed if an

Table 1. Evaluation of Characteristics of the Items of the Rosenberg Self Esteem Scale by Experts

		EXP. 1	EXP.2	EXP.3	EXP.4	M	95% Aiken's V	95%	
								LL	UL
Item 1	R	3	4	4	4	3.75	.92	.65	.92
	C	3	4	4	4	3.75	.92	.65	.92
	I	4	4	4	4	4	1.00	.76	1.00
	CL	4	4	4	4	4	1.00	.76	1.00
Item 2	R	3	4	4	4	3.75	.92	.65	.92
	C	3	4	4	3	3.50	.83	.55	.83
	I	4	4	4	4	4	1.00	.76	1.00
	CL	4	4	4	4	4	1.00	.76	1.00
Item 3	R	3	4	–	4	3.67	.89	.59	.89
	C	3	4	4	4	3.75	.92	.65	.92
	I	4	4	4	4	4	1.00	.76	1.00
	CL	4	4	4	4	4	1.00	.76	1.00
Item 4	R	3	4	4	4	3.75	.92	.65	.92
	C	3	4	4	4	3.75	.92	.65	.92
	I	4	4	4	4	4	1.00	.76	1.00
	CL	4	4	4	4	4	1.00	.76	1.00
Item 5	R	3	4	–	4	3.67	.89	.59	.89
	C	3	3	4	4	3.50	.83	.55	.83
	I	4	4	4	4	4	1.00	.76	1.00
	CL	4	3	4	4	3.75	.92	.65	.92
Item 6	R	3	4	4	4	3.75	.92	.65	.92
	C	3	4	4	4	3.75	.92	.65	.92
	I	4	4	4	4	4	1.00	.76	1.00
	CL	4	4	4	4	4	1.00	.76	1.00
Item 7	R	3	4	4	4	3.75	.92	.65	.92
	C	3	4	4	4	3.75	.92	.65	.92
	I	4	4	4	4	4	1.00	.76	1.00
	CL	3	4	4	4	3.75	.92	.65	.92
Item 8	R	3	4	4	4	3.75	.92	.65	.92
	C	3	3	4	4	3.50	.83	.55	.83
	I	4	4	4	4	4	1.00	.76	1.00
	CL	4	4	4	4	4	1.00	.76	1.00
Item 9	R	3	4	4	4	3.75	.92	.65	.92
	C	3	4	4	4	3.75	.92	.65	.92
	I	4	4	4	4	4	1.00	.76	1.00
	CL	4	4	4	4	4	1.00	.76	1.00
Item 10	R	3	4	4	4	3.75	.92	.65	.92
	C	3	4	4	3	3.50	.83	.55	.83
	I	4	4	4	4	4	1.00	.76	1.00
	CL	4	4	4	3	3.75	.92	.65	.92

Note: R = Representativeness; C = Comprehension; I = Interpretation; CL = Clarity; Exp = Expert; M = Mean; LL = Lower Limit; UL = Upper Limit.

item is removed, and the mean values of the items are slightly above the theoretical average of response (2.5). Furthermore, SDs near and/or slightly above 1 indicate adequate response variability.

Factorial invariance

Progressive factorial invariance was performed in order to evaluate the scale's construct validity and test

factorial equivalence between versions of Spain and Colombia in four different models (M1, M2, M3 and M4) (Table 4). Figure 1 and Table 3 shows the standardized results of the configural model, associated with factorial weights (λ), errors of the items and item variance of all the models. Analysis began by assessing the model's most basic level: the configural level – without restrictions.

Table 2. Psychometric Properties of Items in the Versions from Spain and Colombia

Country	Item	M	SD	r_{it}^c	α -i	α	Total Sum	M (DT) Total
Colombia	Item 1	3.43	0.70	.48	.82	.83	34.20	3.42 (4.52)
	Item 2	3.38	0.80	.56	.81			
	Item 3	3.61	0.57	.55	.81			
	Item 4	3.67	0.54	.60	.81			
	Item 5	3.72	0.49	.66	.81			
	Item 6	3.22	0.85	.62	.80			
	Item 7	3.71	0.54	.48	.82			
	Item 8	2.39	1.08	.28	.86			
	Item 9	3.52	0.70	.69	.80			
	Item 10	3.56	0.62	.62	.81			
Spain	Item 1	3.17	0.72	.53	.85	.86	32.21	3.22 (5.39)
	Item 2	3.09	0.85	.70	.84			
	Item 3	3.41	0.66	.49	.86			
	Item 4	3.45	0.64	.64	.85			
	Item 5	3.15	0.84	.57	.85			
	Item 6	3.11	0.90	.63	.85			
	Item 7	3.34	0.77	.37	.87			
	Item 8	2.78	1.0	.55	.86			
	Item 9	3.46	0.73	.69	.84			
	Item 10	3.24	0.73	.71	.84			

Note: M = mean; SD = Standard Deviation; r_{it}^c = item-total corrected correlation; α -i = Cronbach alpha if item is deleted; α = Cronbach alpha.

At this level M2 and M4 were not given any further consideration due to the poor adjust observed in this basic level. Model 3 has shown good fit index, but invariance could not be defended in the second level: weak invariance. Thus, M1 has in this sample both the best adjustment and the strongest invariance, and has the lowest starting AIC value. As it can be observed in Table 4, this model reaches good fit indexes for all levels. But item 8 was not constricted in the weak invariance as recommended by the modification indexes. Once this modification was done, we observed a strong invariance level. No strict invariance was observed for this model.

Concurrent validity and percentile ranking scores

An analysis of concurrent validity was performed – using Pearson correlations – between the RSES and the sexual self-esteem dimension of the Sexuality Scale. Positive, moderate and significant correlations were observed with Sexual Self-Esteem $p < .01$; $r(620) = .31$; for Colombia and $p < .01$; $r(494) = .41$ for Spain. Finally, the percentile ranking scores of the adaptation of the RSES for the Colombian population were obtained, as differentiated by sex and age group; 18–30, 31–44 and over 45 years (Table 5).

Discussion

This study sought to adapt and validate the RSES in a sample of Colombian individuals, and perform its

factorial equivalence with a sample from Spain. The qualitative evaluation found that all items fit properly in terms of representativeness, comprehension, interpretation and clarity. Appropriate reliability indicators were obtained in the versions both of Spain and Colombia. The mean of scores of the questionnaire's items is very similar to the theoretical mean, and its standard deviation is close to one. Different unidimensional and bidimensional models reach adequate fit indexes, but the only model that achieved a strict partial invariance was the unidimensional with all errors of the positive and negative items covariate. This indicates that this model is equivalent in terms of factorial coefficients and the values of the intercepts. The psychometric properties of the items are adequate; only item 8 ("I wish I could have more respect for myself") has a corrected item-total correlation below .30. Furthermore, it was observed in the analysis of concurrent validity that there are positive, significant and moderate correlations between the RSES and the sexual self-esteem dimension of the Sexuality Scale.

Homologous samples were obtained in this study in terms of sex, sexual orientation and relationship between the two countries. Also, as expected due to cultural differences between Colombia and Spain (van de Vijver & Leung, 2000) significant differences between countries were observed in marital status and religion. However, the samples from Spain and Colombia

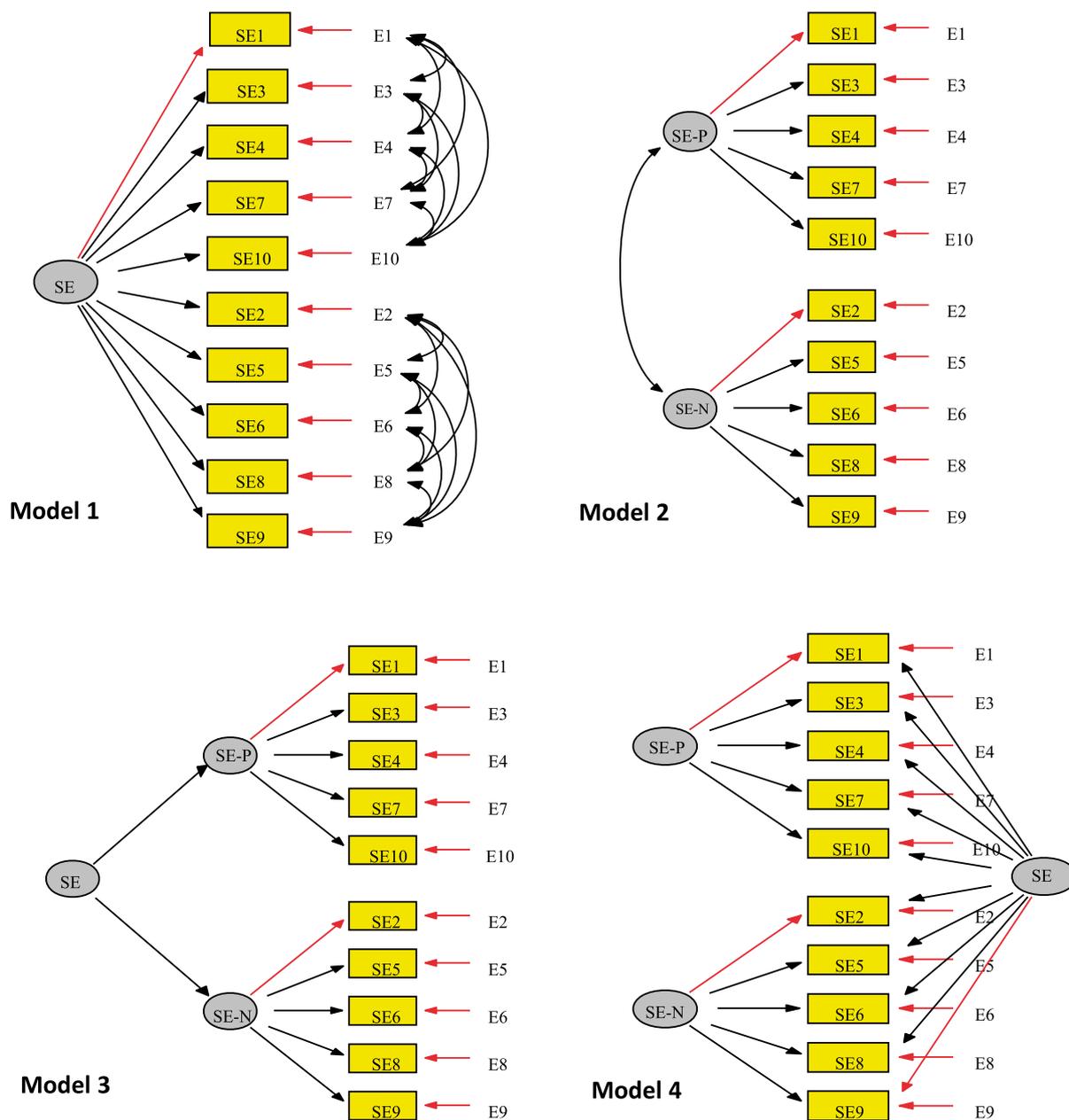


Figure 1. Path diagram of the standardized configural invariance models of the Rosenberg Self-Esteem Scale in Colombia and Spain. SE = General Self-Esteem; SE-P = Positive Self-Esteem; SE-N = Negative Self-Esteem; R^2 = Variance of the item which is explained by the factor. The intercepts have been removed for the sake of visual clarity (V999).

could not be balanced in age and years of study. Nevertheless, the effect size of the differences observed for these two categories was low, therefore it is not expected to have an important impact in the results.

The experts who adapted the scale advised a direct wording of one of the inverse items. Thus, item 5 of the validated Spanish version (“I feel I do not have much to be proud of”), was worded positively (“I feel I have much to be proud of”). Furthermore, it was observed that this adaptation has no consequences in qualitative evaluation or psychometric features.

Qualitative analysis of the scale included a process to designate the table of specifications of the items. This analysis greatly allowed to obtain adequate content validity, as it enabled us to obtain items which were more related to the construct. This procedure is not usually utilized, although it is of utmost importance in processes of validation and/or adaptation of scales, as pointed out by Carretero-Dios and Pérez (2005). The results of the qualitative analysis showed that the items were adequately worded, since the limit below 95% of Aiken’s V was always greater than .50.

Table 3. Standardized factor loadings (λ), errors, and % of explained variance for the Spain and Colombia configural model

		Colombia			Spain		
		λ	Error	R^2	λ	Error	R^2
Model 1	SE1	.48	.87	.23	.55	.83	.30
	SE3	.68	.72	.46	.46	.88	.21
	SE4	.72	.69	.52	.59	.80	.35
	SE7	.65	.75	.42	.32	.94	.10
	SE10	.69	.72	.48	.72	.68	.53
	SE2	.55	.83	.30	.75	.65	.57
	SP5	.81	.58	.65	.64	.76	.41
	SP6	.60	.79	.36	.69	.72	.47
	SP8	.27	.96	.07	.64	.76	.41
Model 2	SE9	.71	.70	.50	.76	.64	.58
	SE-P1	.48	.87	.23	.00	1.0	0
	SE-P3	.69	.72	.48	.59	.80	.35
	SE-P4	.73	.68	.53	.74	.66	.55
	SE-P7	.66	.74	.43	.49	.86	.24
	SE-P10	.69	.71	.48	.76	.6	.42
	SE-N2	.57	.82	.32	.79	.61	.62
	SP-N5	.80	.59	.64	.65	.75	.42
	SP-N6	.62	.78	.39	.72	.69	.52
Model 3*	SP-N8	.28	.95	.08	.60	.79	.36
	SE-N9	.72	.68	.52	.76	.64	.58
	SE-P1	.48	.87	.23	.59	.80	.34
	SE-P3	.69	.72	.48	.57	.81	.33
	SE-P4	.72	.68	.53	.76	.68	.53
	SE-P7	.66	.74	.43	.47	.88	.22
	SE-P10	.69	.71	.48	.77	.63	.60
	SE-N2	.57	.81	.32	.79	.61	.62
	SP-N5	.80	.60	.64	.64	.76	.41
Model 4	SP-N6	.62	.77	.39	.72	.69	.52
	SP-N8	.28	.95	.08	.61	.79	.37
	SE-N9	.73	.68	.53	.76	.64	.58
	SE-P1	.11	.89	.20	.12	.81	.33
	SE-P3	.26	.63	.72	.45	.76	.41
	SE-P4	.31	.65	.56	.61	.51	.73
	SE-P7	.36	.71	.49	.36	.86	.25
	SE-P10	.16	.73	.45	.14	.65	.57
	SE-N2	.00	.83	.29	.67	.03	.99
	SP-N5	.43	.00	1.0	.11	.77	.39
	SP-N6	.24	.72	.48	.15	.70	.50
	SP-N8	.19	.93	.12	.04	.76	.41
	SE-N9	.27	.58	.65	.14	.63	.57
	SE1	.44			.56		
	SE3	.63			.45		
	SE4	.68			.59		
	SE7	.59			.34		
	SE10	.65			.74		
	SE2	.54			.74		
	SP5	.90			.61		
	SP6	.64			.69		
	SP8	.29			.64		
	SE9	.76			.76		

Note: SE = Self-Esteem; SE-P = Positive Self-Esteem; SE-N = Negative Self-Esteem; λ = Standardized weight item-factor; R^2 = explained variance.

*Second order λ all above .85.

Table 4. Goodness-of-Fit Indexes for the Different Invariance Models and Model Selection

		S-B χ^2	df	p	AIC	CAIC	CFI	Δ CFI	RMSEA	90% CI for RMSEA	
										LL	UL
M1	Configural	85.18	30	< .001	25.18	-155.97	.982	-	.057	.043	.071
	Weak	121.33	39	< .001	43.33	-192.18	.970	-.012	.061	.049	.073
	Weak (-item 8)	112.59	38	< .001	36.59	-192.87	.972	-.010	.059	.046	.071
	Strong	418.62	48	< .001	322.62	32.76	.963	-.009	.075	.063	.087
M2	Configural	306.86	68	< .001	170.86	-239.76	.912	-	.079	.070	.087
	Weak	335.35	76	< .001	183.35	-275.59	.904	-.008	.077	.069	.086
	Strong	664.47	85	< .001	494.47	-18.82	.904	.000	.109	.102	.117
M3	Configural	163.72	57	< .001	49.72	-294.48	.962	-	.057	.047	.068
	Weak	65.57	67	.52	-68.42	-473.01	1	.038	.000	.000	.021
	Strong	213.77	77	< .001	59.75	-405.23	.980	-.020	.056	.047	.065
M4	Configural	613.69	94	< .001	425.69	-141.95	.996	-	.099	.091	.106
	Weak	242.53	67	< .001	108.53	-296.06	.935	-.061	.068	.059	.077
	Strong	537.59	77	< .001	383.59	-81.39	.934	-.001	.103	.094	.111

Note: S-B χ^2 = Santorra-Bentler Scaled Chi-Square; df = degree of freedom; CAIC: Consistent Akaike Information Criterion; AIC: Akaike Information Criterion; CFI = Comparative Fit Index; Δ CFI = Increment of the Comparative Fit Index; RMSEA = Root Mean Square Error of Approximation; LL = Lower Limit; UL = Upper Limit; (-item 8) = item 8 was not constricted. M1, M2, M3 and M4 are represented in Figure 1. and Table 3.

The reliability indices found (.83 for Colombia and .86 for Spain) are similar to those described by Rosenberg (1965), and subsequently by other authors (Gray-Little et al., 1997; Martin-Albo et al., 2007; Prezza, Trombaccia, & Armento, 1997; inter alia). These reliability indices are suitable for this study, since – according to Nunnally and Bernstein (1995) – minimum recommended reliability by means of internal consistency reliability is .80 when the aim of the research is to find a diagnosis or classification. The mean value of scores in the questionnaire's items is very similar to the theoretical mean, which is expected for a communitarian population, and its standard deviations near one.

Regarding the items' psychometric properties, it was observed that all except item 8 ("I wish I could have more respect for myself"), had good psychometric properties. Although this item has corrected item-total correlation below .30, and other problems related with the invariance we consider it is not enough to dismiss this item. However, it would be important in forthcoming research to determine whether this item is actually representative of the general self-esteem construct, or it is rather an item which is representative of a specific construct of the Self, for instance "Self-respect" (Rosenberg, 1979; Wang, Siegal, Falck, & Carlson, 2001), because this construct has been described as a parallel – though related – dimension to General Self-Esteem. This would explain the differences in participants' responses, since – according to various authors (Markus & Wurf, 1987) – individuals tend to

evaluate differently various aspects of the Self. In fact, we are not the first authors to find problems with item 8 (Leung & Wong, 2008). And if our goal is to compare scores between Spain and Colombia this item should be excluded due to invariance issues.

Progressive evaluation of factorial invariance is relevant, as this is the only type of analysis that allows to lay the psychometric foundations for justifications of comparisons between groups (Elosua, 2005). As for the factorial invariance analysis conducted in the study, a strong level of partial invariance was found. This implies the possibility to compare general self-esteem scores between Spain and Colombia, with minimum bias in the measurement (Dimitrov, 2010), but as said before item 8 should be excluded for the comparison. Model 1, where all negatives and positives items were covariate within them, has the best invariance here tested. The fact that all errors were covariate implies that all of them had some similar way to be answered. Thus, these covariances suggest that all positive items have a common component (probably its positive redaction) as negative ones does. And at least in our sample this model has showed the best fit, invariance indicators and the lower AIC value in the configural level (Table 4). This data will help to include more information to the debate of the dimensionality of the RSES. As for factorial weights (λ), it was observed that all items show values which are clearly higher than .30, except for item 8 (.32), which is consistent with the

Table 5. Validated RSSE Scales for the Colombian and Spanish Population, by Sex and Age Group

		Colombia					
Sex		Women			Men		
Percentiles	Age	18–30	31–44	+45	18–30	31–44	+45
	N	201	142	49	98	87	49
	M	33.26	34.68	35.39	33.73	34.64	35.57
	SD	4.77	4.40	3.93	4.27	4.78	3.13
	Skewness	–0.61	–0.92	–0.74	–0.76	–0.93	–0.19
	Kurtosis	–0.30	0.41	–0.61	0.33	–0.12	–1.01
	Minimum	20	20	27	21	22	29
	Maximum	40	40	40	40	40	40
	1	21	21	27	21	22	29
	5	24	26	27	25	24	30
	15	29	30	30	29	28	31
	25	30	32	33	31	31	33
	35	32	34	34	33	34	34
	50	34	36	37	34	36	35
	65	36	37	38	36	38	37
	75	37	38	38	37	38	38
	85	38	39	39	38	39	39
	95	40	40	40	40	40	40
	99	40	40				
		Spain					
Sex		Women			Men		
Percentiles	Age	18–30	31–44	+45	18–30	31–44	+45
	N	183	72	49	77	77	41
	M	31.09	31.70	35.38	32.75	33.36	33.39
	SD	5.57	5.47	3.92	5.74	4.78	4.66
	Skewness	–0.55	–0.70	–0.74	–0.92	–0.29	–0.3
	Kurtosis	0.25	–0.06	–0.60	0.61	–0.97	–0.41
	Minimum	12	18	27	15	23	22
	Maximum	40	40	40	40	40	40
	1	13	18	27	15	23	22
	5	21	20	27	21	24	23
	15	26	24	30	27	28	29
	25	28	29	33	29	29	30
	35	29	30	34	31	31	30
	50	31	33	37	34	34	33
	65	34	35	38	36	36	37
	75	36	35	38	38	37	37
	85	37	37	39	39	39	39
	95	39	40	40	40	40	40
	99	40					

Note: M = Media; SD = Standard Deviation *Caution when interpreting for man +45 in Colombia and women in the same age for Spain because non-continuous and non-normality problems issues we could not reach a good variance for scores for these groups.

results observed in this and other studies (Rosenberg, 1979; Wang et al., 2001) and continues to be indicated as a problematic item.

Model 1, as the others, has been also tested in other studies. For instance, studies which have systematically reviewed the factorial structure of the Rosenberg

scale, as is the case of Marsh (1996) or Tomas and Oliver (1999) show how the RSES has been evaluated by means of different ways. Unidimensional structures have been proposed wherein covariances of the errors are performed in both positive and negative items of the scale – whether altogether or independently – or which

have been evaluated through two-factor models, whether *Positive Self-Esteem* and *Negative Self-Esteem* or *General Evaluation* and *Transient Evaluation*, amongst other models.

Regarding the analysis of concurrent validity, it was found that there are significant correlations between the RSES and the dimension of sexual self-esteem of the Sexuality Scale. These correlations are positive and moderate. This is related to the statement above, in that general self-esteem is not equivalent or interchangeable with the specific self-esteem (Rosenberg et al., 1995), since individuals usually relate to one another and evaluate themselves differently in all aspects of their lives. In addition, the findings are consistent with the postulates of Zeanah and Schwarz (1996), who report that people “*have a sense of the sexual self which is distinct from, but contributes to, the global sense of self*” (p. 2). In addition, Swenson, Houck, Barker, Zeanah, and Brown, (2012) state that Sexual Self-Esteem and General Self-Esteem are two constructs that should be differentiated, among other reasons because the clinical implications of having a low general self-esteem, are not similar to the implications of having a low Sexual Self-Esteem. This explains how some people may have low general self-esteem and high sexual self-esteem or *vice versa*.

Lastly, presenting percentile ranking scores differentiated by country, sex, and age ranges can be useful in evaluating general self-esteem in these populations. We ask some caution when interpreting for man +45 in Colombia and women in the same age for Spain because some non-continuous and non-normality limitations we could not reach a good variance for scores for these groups.

In conclusion, we present a reliable scale which was rigorously adapted, in pursuance of the guidelines for adapting tests to other cultures described above, which has a stable performance in two countries with different cultures (Colombia and Spain). Moreover, this is a reliable questionnaire with a correspondence between the theoretical assumption and the dimensionality exposed, as well as with internal consistency indicators and adequate concurrent validity. This, it can be concluded that there is sufficient evidence supporting the construct validity of the scale. In addition, clear indications are presented in that the adapted scale can be used in Colombian population, both men and women of age.

However, this study is not without limitations. A non representative convenience sampling does not allow for generality of results for all Spanish and Colombian populations. Furthermore, using an internet-based sampling method does not allow for equal access to all socioeconomic strata of the population due to the reality of internet access for people. Item 8 is being problematic and future research should determine any changes on it. Therefore, it would be interesting to extend the

evaluation sample in the future to different clinical realities or more complex collectives. This would allow to continue this work with the evaluation of the implications of general self-esteem in the health of the Colombian population – as has been done so far – but with a valid and reliable scale.

References

- AERA, APA., & NCME. (2014). *Standards for educational and psychological tests*. Washington DC: American Educational Research Association.
- Bachman J. G., O'Malley P. M., Freedman-Doan P., Trzesniewski K. H., & Donnellan M. B. (2011). Adolescent self-esteem: Differences by race/ethnicity, gender, and age. *Self and Identity*, 10, 445–473. <http://dx.doi.org/10.1080/15298861003794538>
- Brechan I., & Kvale I. L. (2015). Relationship between body dissatisfaction and disordered eating: Mediating role of self-esteem and depression. *Eating Behaviors*, 17, 49–58. <http://dx.doi.org/10.1016/j.eatbeh.2014.12.008>
- Byrne B. M., Oakland T., Leong F. T. L., van de Vijver F. J. R., Hambleton R. K., Cheung F. M., & Bartram D. (2009). A critical analysis of cross-cultural research and testing practices: Implications for improved education and training in psychology. *Training and Education in Professional Psychology*, 3, 94–105. <http://dx.doi.org/10.1037/a0014516>
- Carretero-Dios H., & Pérez C. (2005). Normas para el desarrollo y revisión de estudios instrumentales [Norms to development and to review instrumental studies]. *International Journal of Clinical and Health Psychology*, 5, 521–551.
- Cheng C. M., Govorun O., & Chartrand T. L. (2012). Effect of self-awareness on negative affect among individuals with discrepant low self-esteem. *Self and Identity*, 11, 304–316. <http://dx.doi.org/10.1080/15298868.2011.567022>
- Cheung G. W., & Rensvold R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, 9, 233–255. http://dx.doi.org/10.1207/S15328007SEM0902_5
- de Bernardo D. H., & Curtis A. (2013). Using online and paper surveys: The effectiveness of mixed-mode methodology for populations over 50. *Research on Aging*, 35, 220–240. <http://dx.doi.org/10.1177/0164027512441611>
- Dimitrov D. M. (2010). Testing for factorial invariance in the context of construct validation. *Measurement and Evaluation in Counseling and Development*, 43, 121–149. <http://dx.doi.org/10.1177/0748175610373459>
- Elosua P., (2005). Evaluación progresiva de la invarianza factorial entre las versiones original y adaptada de una escala de autoconcepto [Progressive way in strict factorial invariance]. *Psicothema*, 17, 356–362.
- Elosua P., Mujika J., Almeida L. S., & Hermosilla D. (2014). Procedimientos analítico-razonales en la adaptación de test. Adaptación al español de la batería de pruebas de razonamiento [Judgmental-analytical procedures for adapting tests: Adaptation to spanish of the reasoning tests battery]. *Revista Latinoamericana de Psicología*, 46, 117–126.

- Falk C. F., Heine S. J., Takemura K., Zhang C. X. J., & Hsu C. W. (2015). Are implicit self-esteem measures valid for assessing individual and cultural differences? *Journal of Personality*, 83, 56–68. <http://dx.doi.org/10.1111/jopy.12082>
- Fiz M. R., & Oyón M. J. (1998). Influencia de la autoestima y su repercusión en el rendimiento académico de los adolescentes [Influence of the self-esteem and its repercussion in the academic performance of teenagers]. *Estudios de Pedagogía y Psicología*, 10, 27–74.
- González-Pienda J., Núñez J. C., González-Pumariega S., & García M. S. (1997). Autoconcepto, autoestima y aprendizaje escolar [Self-concept, self-esteem and school learning]. *Psicothema*, 9, 271–289.
- Gray-Little B., Williams V. S. L., & Hancock T. D. (1997). An item response theory analysis of the Rosenberg Self-Esteem Scale. *Personality and Social Psychology Bulletin*, 23, 443–451. <http://dx.doi.org/10.1177/0146167297235001>
- Harper L. A., Coleman J. A., Perrin P. B., Olivera S. L., Perdomo J. L., Arango J. A., & Arango-Lasprilla J. C. (2014). Comparison of mental health between individuals with spinal cord injury and able-bodied controls in Neiva, Colombia. *Journal of Rehabilitation Research and Development*, 51, 127–136. <http://dx.doi.org/10.1682/JRRD.2013.04.0086>
- Hyland P., Boduszek D., Dhingra K., Shevlin M., & Egan A. (2014). A bifactor approach to modelling the Rosenberg Self Esteem Scale. *Personality and Individual Differences*, 66, 188–192. <http://dx.doi.org/10.1016/j.paid.2014.03.034>
- Huang C., & Dong N. (2012). Factor structures of the Rosenberg Self-Esteem Scale: A meta-analysis of pattern matrices. *European Journal of Psychological Assessment*, 28, 132–138. <http://dx.doi.org/10.1027/1015-5759/a000101>
- Leung S. O., & Wong P. M. (2008). A study on reliability and validity of Chinese Rosenberg Self-Esteem Scale (in Chinese). *New Horizons in Education*, 56, 62–69.
- Li J. B., Delvecchio E., Di Riso D., Salcuni S., & Mazzeschi C. (2015). Self-esteem and its association with depression among Chinese, Italian, and Costa Rican adolescents: A cross-cultural study. *Personality and Individual Differences*, 82, 20–25. <http://dx.doi.org/10.1016/j.paid.2015.02.036>
- Luciano M. C., Gómez I., & Valdivia S. (2002). Consideraciones acerca del desarrollo de la personalidad desde un marco funcional contextual [Some considerations about personality development from a functional-contextual point of view]. *International Journal of Psychology Therapy*, 2, 173–197.
- Lynn M. R. (1986). Determination and quantification of content validity. *Nursing Research*, 35, 382–386. <http://dx.doi.org/10.1097/00006199-198611000-00017>
- McKay M. T., Boduszek D., & Harvey S. A. (2014). The Rosenberg Self-Esteem Scale: A bifactor answer to a two-factor question? *Journal of Personality Assessment*, 96, 654–660. <http://dx.doi.org/10.1080/00223891.2014.923436>
- Markus H., & Wurf E. (1987). The dynamic self-concept: A social psychological perspective. *Annual Review of Psychology*, 38, 299–337. <http://dx.doi.org/10.1146/annurev.ps.38.020187.001503>
- Marsh H. M. (1996). Positive and negative global self-esteem: A substantial meaningful distinction or artifacts? *Journal of Personality and Social Psychology*, 70, 810–819. <http://dx.doi.org/10.1037/0022-3514.70.4.810>
- Marshall S. L., Parker P. D., Ciarrochi J., & Heaven P. C. L. (2014). Is self-esteem cause or consequence of social support? A four year longitudinal study. *Child Development*, 85, 1275–1291.
- Marshall S. L., Parker P. D., Ciarrochi J., Sahdra B., Jackson C. J., & Heaven P. C. L. (2015). Self-compassion protects against the negative effects of low self-esteem: A longitudinal study in a large adolescent sample. *Personality and Individual Differences*, 74, 116–121. <http://dx.doi.org/10.1016/j.paid.2014.09.013>
- Martín-Albo J., Núñez J. L., Navarro J. G., & Grijalvo F. (2007). The Rosenberg Self-Esteem Scale: Translation and validation in university students. *The Spanish Journal of Psychology*, 10, 458–467. <http://dx.doi.org/10.1017/S1138741600006727>
- Merino C., & Livia J. (2009). Intervalos de confianza asimétricos para el índice la validez de contenido: Un programa visual basic para la Aiken's V [Some considerations about personality development from a functional-contextual point of view]. *Anales de Psicología*, 25, 169–171.
- Milanés Z. C., & Gómez-Bustamante E. (2011). Prevalencia y factores asociados al consumo de sustancias ilegales en adultos de Cartagena, Colombia [Prevalence and associated factors of illegal substance use in Cartagena, Colombia]. *Revista Colombiana de Psiquiatría*, 40, 723–733. [http://dx.doi.org/10.1016/S0034-7450\(14\)60160-8](http://dx.doi.org/10.1016/S0034-7450(14)60160-8)
- Muñiz J., Elosua P., & Hambleton R. K. (2013). Directrices para la traducción y adaptación de los tests: Segunda edición [International test commission guidelines for test translation and adaptation: Second edition]. *Psicothema*, 25, 151–157. <http://dx.doi.org/10.7334/psicothema2013.24>
- Muñoz J. M., & Alonso F. M. (2013). Life satisfaction: Its relationship with prejudice, national identity, self-esteem and material well-being in immigrants. *Acta Colombiana de Psicología*, 15, 99–108.
- Nunnally J. C., & Bernstein I. J. (1995). *Teoría psicométrica* [Psychometric theory]. Madrid, Spain: McGraw-Hill.
- Pérez-Olmos I., Téllez Cruz D. L., Vélez Traslaviña Á. L., & Ibáñez-Pinilla M. (2012). Caracterización de factores asociados con comportamiento suicida en adolescentes estudiantes de octavo grado, en tres colegios bogotanos [Risk factors associated with the suicide attempt and the criteria on it happened in the adolescents]. *Revista Colombiana de Psiquiatría*, 41, 26–47. [http://dx.doi.org/10.1016/S0034-7450\(14\)60067-6](http://dx.doi.org/10.1016/S0034-7450(14)60067-6)
- Prezza M., Trombaccia F. R., & Armento L. (1997). La scala dell'autostima di Rosenberg: Traduzione e validazione italiana [The Rosenberg Self-Esteem Scale: Italian translation and Validation]. *Bollettino di Psicologia Applicata*, 223, 35–44.
- Robins R. W., Hendin H. M., & Trzesniewski K. H. (2001). Measuring global self-esteem: Construct validation of a single-item measure and the Rosenberg Self-Esteem Scale. *Personality and Social Psychology Bulletin*, 27, 151–161. <http://dx.doi.org/10.1177/0146167201272002>
- Rosenberg M. (1965). *Society and the adolescent self-image*. Princeton, NJ: Princeton University Press.

- Rosenberg M. (1979). *Conceiving the self*. New York, NY: Basic Books.
- Rosenberg M., Schooler C., Schoenbach C., & Rosenberg F. (1995). Global self-esteem and specific self-esteem: Different concepts, different outcomes. *American Sociological Review*, 60, 141–156. <http://dx.doi.org/10.2307/2096350>
- Silbereisen R. K., & Wiesner M. (2002). Lessons from research on the consequences of German unification: Continuity and discontinuity of self-efficacy and the timing of psychosocial transitions. *Applied Psychology*, 51, 291–317. <http://dx.doi.org/10.1111/1464-0597.00093>
- Snell W. E., & Papini D. R. (1989). The sexuality scale: An instrument to measure sexual-esteem, sexual-depression, and sexual-preoccupation. *Journal of Sex Research*, 26, 256–263. <http://dx.doi.org/10.1080/00224498909551510>
- Soler F., Gómez-Lugo M., Espada J. P., Morales S., Sierra J. C., Marchal-Bertrand L., & Vallejo-Medina P. (2016). Translation, adaptation and validation of the brief sexuality scale in Colombian and Spanish populations. *International Journal of Psychology and Psychological Therapy*, 16, 343–356.
- Swenson R. R., Houck C. D., Barker D., Zeanah P. D., & Brown L. K. (2012). Prospective analysis of the transition to sexual experience and changes in sexual self-esteem among adolescents attending therapeutic schools. *Journal of Adolescence*, 35, 77–85. <http://dx.doi.org/10.1016/j.adolescence.2011.06.002>
- Tomas J. M., & Oliver A. (1999). Rosenberg's self-esteem scale: Two factors or method effects. *Structural Equation Modeling*, 6, 84–98. <http://dx.doi.org/10.1080/10705519909540120>
- Vallejo-Medina P., Gómez-Lugo M., Marchal-Bertrand L., Saavedra-Roa A., Morales A., & Soler F. (2016). *Guidelines to translate and Adapt Scales in intra-language but inter-cultural contexts*. (Manuscript submitted for publication).
- van de Vijver F. J. R., & Leung K. (2000). Methodological issues in psychological research on culture. *Journal of Cross-Cultural Psychology*, 31, 33–51. <http://dx.doi.org/10.1177/0022022100031001004>
- Wang J., Siegal H. A., Falck R. S., & Carlson R. G. (2001). Factorial structure of Rosenberg's Self-Esteem Scale among crack-cocaine drug users. *Structural Equation Modeling*, 8, 275–286. http://dx.doi.org/10.1207/S15328007SEM0802_6
- Wichstrom L. (2000). Predictors of adolescent suicide attempts: A nationally representative longitudinal study of Norwegian adolescents. *Journal of the American Academy of Child & Adolescent Psychiatry*, 39, 603–610. <http://dx.doi.org/10.1097/00004583-200005000-00014>
- Wiedeman M. W., & Allgeier E. R. (1993). The measurement of sexual-esteem: Investigation of Snell and Papini's (1989) Sexuality Scale. *Journal of Research in Personality*, 27, 88–102. <http://dx.doi.org/10.1006/jrpe.1993.1006>
- Zeigler-Hill V., Holden C. J., Enjaian B., Southard A. C., Besser A., Li H., & Zhang Q. (2015). Self-Esteem instability and personality: The connections between feelings of self-worth and the five dimensions of personality. *Personality and Social Psychology Bulletin*, 41, 183–198. <http://dx.doi.org/10.1177/0146167214559719>
- Zeanah P. D., & Schwarz J. C. (1996). Reliability and validity of the sexual self-esteem inventory for women. *Assessment*, 3, 1–15.

Appendix

Rosenberg Self-Esteem Scale (RSES; Rosenberg, 1965) – Version Validated in Colombia.

1 = Totalmente en desacuerdo; 2 = De acuerdo; 3 = En desacuerdo; 4 = Totalmente en desacuerdo
(1 = Strongly Disagree; 2 = Agree; 3 = Disagree; 4 = Strongly Disagree)

1. En general estoy satisfecho/a conmigo mismo/a. (On the whole, I am satisfied with myself.)	1	2	3	4
2. A veces pienso que no soy bueno/a para nada.* (At times I think I am no good at all.)	1	2	3	4
3. Creo que tengo buenas cualidades. (I feel that I have a number of good qualities.)	1	2	3	4
4. Soy capaz de hacer las cosas tan bien como la mayoría de las personas. (I am able to do things as well as most other people.)	1	2	3	4
5. Pienso que tengo muchas cosas de las cuales sentirme orgulloso/a. (I feel I have much to be proud of.)	1	2	3	4
6. A veces me siento inútil.* (I certainly feel useless at times.)	1	2	3	4
7. Creo que soy una persona igual de valiosa a la mayoría de la gente. (I feel that I'm a person of worth, at least on an equal plane with others.)	1	2	3	4
8. Quisiera respetarme más a mí mismo/a.* (I wish I could have more respect for myself.)	1	2	3	4
9. Tiendo a pensar que soy un fracasado/a.* (All in all, I am inclined to feel that I am a failure.)	1	2	3	4
10. Tengo una visión positiva sobre mí mismo/a. (I take a positive attitude toward myself.)	1	2	3	4

Note: Items with an asterisk are inversely worded. English items presented here are not validated and should not be used. To read the Spanish–Spain– version see: Martín-Albo, J. Núñez, J. L., Navarro, J. G., & Grijalvo, F. (2007). The Rosenberg Self-Esteem Scale: Translation and validation in university students. *The Spanish Journal of Psychology*, 10, 458–467. <http://dx.doi.org/10.1017/S1138741600006727>