

Escala de autoestima de Rosenberg (EAR): Análisis de la estructura factorial y propuesta de una nueva versión de solo ítems positivos

Lindsey W. Vilca¹, Aaron Travezaño-Cabrera², Stephany Santos-García²

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

Correspondencia: 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

© Universidad de Almería and Ilustre Colegio Oficial de la Psicología de Andalucía Oriental (Spain)

Resumen

Introducción. El objetivo fue evaluar la estructura factorial y proponer una nueva versión de la Escala de autoestima de Rosenberg de solo ítems positivos para superar el efecto de método asociado a los ítems negativos.

Método. Se planteó una versión A (ítems positivos y negativos) y una versión B (solo ítems positivos). Se recolectó una muestra de 350 universitarios para cada versión.

Resultados. El AFC evidencia que el modelo unidimensional (A1) de la versión A, no presentan adecuados índices de ajuste. También se encontró que añadir un factor específico para ítems negativos (modelo A2) y otro factor para ítems positivos (modelo A3) no mejora los índices de ajuste. Un modelo bidimensional (modelo A4), tampoco mejorar los índices de ajuste. Respecto al modelo unidimensional de la versión B, este presenta un ajuste superior frente al modelo original (modelo A1)

Discusión y conclusiones. Se concluye que la versión B, una propuesta de solo ítems positivos, mide adecuadamente la autoestima, ya que no tiene ítems negativos y además concuerda con el planteamiento original de Rosenberg.

Palabras Clave: Escala de Autoestima de Rosenberg, análisis factorial confirmatorio, autoestima, validez de estructura, ítems negativos, universitarios.

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.

Introducción

La autoestima es uno de los constructos psicológicos más estudiados en el contexto de la educación superior ya que está relacionado con el logro académico (Hosseini et al., 2016), una mejor percepción del aprendizaje (González-Vázquez, 2019), bajos niveles de estrés académico (Chávez & Peralta, 2019) y procrastinación (Hajloo, 2014).

En este contexto, la escala de Autoestima de Rosenberg (EAR) (Rosenberg, 1965) es uno de los instrumentos más empleados y en consecuencia ha sido adaptada a diferentes países. Las adaptaciones se han realizado en adolescentes, jóvenes, adultos, adultos mayores y a nivel de educación secundaria y superior. Sin embargo, no hay un consenso sobre la estructura factorial de la escala, ya que por una parte varios estudios respaldan un modelo bidimensional. Por ejemplo, un estudio realizado en México, en adultos mayores reportó adecuados índices de ajuste ($CFI=.91$, $RMSEA=.07$, $SRMR=.73$) (De León Ricardi & García, 2016). También en estudiantes universitarios mexicanos se encontró un adecuado soporte al modelo bidimensional ($GFI=.99$, $CFI=.98$, $RMSEA=.04$) (Jurado et al., 2015). En Irlanda se reportó que la estructura bidimensional era adecuada ($GFI=.97$, $CFI=.98$, $RMSEA=.07$) (Kielkiewicz et al., 2020).

Por otra parte, otros estudios respaldan una estructura unidimensional, es así que en una investigación realizada en 53 países mediante un análisis factorial exploratorio se brindó respaldo a la solución unifactorial (Schmitt & Allik, 2005). Adicionalmente, estudios que realizados en Estonia ($GFI=.98$, $RMSEA=.04$) (Pullmann & Allik, 2000), España ($CFI=.99$, $RMSEA=.09$, $SRMR=.05$) (Martín-Albo et al., 2007) y Tailandia ($CFI=.97$, $RMSEA=.05$, $SRMR=.04$) (Tinakon & Nahathai, 2012) evidencian adecuados valores en el modelo unidimensional.

Sin embargo, varios estudios sugieren que la unidimensionalidad del EAR se logra mediante el control del efecto de los ítems negativos. Muestra de ello, es el reciente estudio en adolescentes de Perú en donde se evidenció un mejor ajuste al añadir un factor específico para los ítems negativos ($CFI=.97$, $RMSEA=.047$, $SRMR=.33$) (Sánchez-Villena et al., 2021). De manera similar, en Brasil se reportó adecuados índices de ajuste al modelo cuando se añadió un factor específico para los ítems negativos ($CFI=.97$, $TLI=.94$, $RMSEA=.07$) (de Lima & de Souza, 2019).

Frente a la problemática de la estructura factorial, Greenberger et al. (2003) realizó un estudio donde comparó la versión original del EAR con una versión de únicamente ítems positivos y otra versión de solo ítems negativos. Sus resultados evidenciaron un mayor soporte al modelo bidimensional ($CFI=.95$, $TLI=94$, $RMSEA=.08$), encontrándose que las otras versiones no mostraron un buen ajuste a los datos.

Sin embargo, el estudio evidenció algunas limitaciones metodológicas, como el tamaño de muestra inferior a 300 participantes, lo cual no es recomendable cuando se realiza un AFC (Guadagnoli & Velicer, 1988). Además, ambas versiones no evidenciaron ninguna evidencia de validez de contenido de los ítems modificados, este último debiera considerarse pues permite reducir la varianza irrelevante del constructo (Ventura-León, 2019). Realizarla contribuye a obtener una mayor fuente de validez, es decir lograr acumular mayor evidencia que respalde la interpretación de los puntajes del test para un fin determinado (Association American Educational Research et al., 2014).

Por otro lado, en el estudio las categorías de respuesta del EAR fueron modificadas de 4 a 6 categorías, lo cual difiere del planteamiento original de Rosenberg (1965) y de los diversos estudios psicométricos (de Lima & de Souza, 2019; Kielkiewicz et al., 2020; Sánchez-Villena et al., 2021).

Respecto a la confiabilidad del EAR, la mayoría de los estudios utilizaron el coeficiente *Alpha de Cronbach*, el cual no es adecuado porque no cumple los supuestos para su cálculo como la tau-equivalencia, es decir que todos los ítems presenten la misma carga factorial (Cho, 2016). Y cuando se empleó el coeficiente omega de McDonald no se trabajó con las cargas factoriales obtenidas mediante un AFC (Viladrich et al., 2017).

Frente a lo expuesto, el presente estudio tiene como objetivo general evaluar la estructura factorial de la versión de ítems positivos del EAR, ya que una versión ítems positivos es más adecuada debido principalmente a la ausencia de ítems negativos. Se conoce que la inclusión de ítems negativos producen errores de medición que alteran el análisis de los datos, siendo una de las consecuencias principales la aparición de factores adicionales para el conjunto de ítems (de Lima & de Souza, 2019; DiStefano & Motl, 2006). Asimismo, genera una reducción en los valores de la fiabilidad (Menold, 2019; Salazar, 2015; Suárez-Alvarez et al.,

2018), a su vez contribuye en la variabilidad de las respuestas lo que ocasiona una menor varianza (Suárez-Alvarez et al., 2018).

Como objetivos específicos, el estudio pretende (a) evaluar la validez de contenido de la escala original de ítems positivos y negativos y la escala de solo ítems positivos, (b) comparar la versión de solo ítems positivos con tres modelos competidores y (c) evaluar la consistencia interna de ambas versiones de la escala.

Método

Participantes

Para la versión A de la escala, se obtuvo una muestra de 350 universitarios peruanos de ambos sexos 160 varones (46%) y 190 mujeres (54%) entre las edades 18 a 35 años ($M = 21.62$, $SD = 3.24$). Para la versión B, se obtuvo una muestra de 350 universitarios peruanos de ambos sexos 129 varones (37%) y 221 mujeres (63%) entre las edades de 18 a 35 años ($M = 20.87$, $SD = 3.08$). Los criterios de inclusión fueron: participantes que estén cursando estudios universitarios, mayores de 18 años, que acepten el consentimiento informado. En tanto que los criterios de exclusión fueron: estudiantes extranjeros, menores de 18 años, estudiantes universitarios que no estén estudiando actualmente.

Instrumentos

La escala original fue desarrollada por Rosenberg (1965), quien planteó una medida unidimensional de la autoestima. Está conformada por 10 ítems con cuatro opciones de respuesta (Muy de acuerdo=4, De acuerdo=3, En desacuerdo=2, Muy en desacuerdo=1). Para el presente estudio se seleccionaron los ítems con mayor representación de constructo de las versiones adaptadas del EAR en Colombia (Ceballos-Ospino et al., 2017), Chile (Rojas-Barahona et al., 2009) y Perú (Ventura-León et al., 2018) para su utilización en el contexto peruano. El criterio que se utilizó para seleccionar los ítems fue la claridad en el contenido de los reactivos y relevancia para medir el constructo en el contexto peruano. De tal forma se obtuvieron dos versiones de la escala, formas A y B, que fueron evaluadas cada una mediante 6 psicólogos especialistas en el tema.

Procedimiento

El estudio fue aprobado por el comité de ética de la Universidad Peruana Unión siguió estrictamente las normas establecidas en el código de Helsinki (World Medical Association, 2013). Para estudiar la estructura factorial de la EAR se planteó dos versiones, la versión A que contiene ítems positivos y negativos y la versión B que solo contienen ítems positivos. Los ítems utilizados en cada versión fueron extraídos de las adaptaciones del EAR en Latinoamérica. Primero, ambas versiones fueron evaluadas por doce psicólogos (seis por versión) con amplio conocimiento del constructo. En segundo lugar, la escala corregida de ambas versiones se aplicó a un grupo piloto de 20 universitarios (diez por versión) en donde no se realizó ninguna modificación a los ítems. El objetivo de la prueba piloto fue identificar alguna dificultad en la comprensión de ítems. En tercer lugar, la escala final de ambas versiones se aplicó mediante un formulario virtual realizado en la plataforma digital Google Forms que fue distribuida en redes sociales (Facebook y Whatsapp). En el formulario los participantes debían completar el consentimiento informado previo a su realización, adicionalmente se indicaba el objetivo del estudio y la confidencialidad de los datos recolectados.

Análisis de datos

Para la validez de contenido, se utilizó el coeficiente V de Aiken, donde valores superiores a .70 evidencia un mayor consenso de los jueces sobre la idoneidad de los ítems (Aiken, 1980). Para el análisis descriptivo, se calculó estadísticos descriptivos, donde la asimetría y curtosis indican valores adecuados cuando son menores a ± 1.5 (Forero et al., 2009). Para el AFC se utilizó el estimador WLSMV, debido a la naturaleza ordinal de los ítems (Brown, 2015) y para evaluar el ajuste del modelo se utilizó el coeficiente *CFI* ($>.95$), *TLI* ($>.95$), *RMSEA* ($<.08$), *SRMR* ($<.08$) y *WRMR* (< 1.00), este último se utilizó debido a que es adecuado cuando se utilizan datos ordinales (DiStefano, Liu, Jiang, y Shi, 2018). Para valorar los índices de ajuste se trabajó con la propuesta de Hu y Bentler (1999) y DiStefano, Liu, Jiang, y Shi (2018). Todos los análisis estadísticos fueron desarrollados utilizando el programa R (version 3.5.0) (R Core Team, 2019) y el entorno RStudio Team (2018).

Resultados

Validez basada en el contenido

En la tabla 1, se muestra que la versión A y B presentan valores adecuados en los criterios de claridad, relevancia, coherencia y contexto ($V > .71$). Por otra parte, los ítems fueron

modificados siguiendo las sugerencias de los jueces en la versión A (ítem 5) y en la versión B (ítem 3, ítem 5, ítem 9 y ítem 10).

Tabla 1. *Validez de contenido de la versión A y B.*

Ítems	Versión A	V (Rele)	V (Cohe)	V (Clar)	V (Cont)	Versión 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

Análisis descriptivos de los ítems

En la tabla 2, se muestran los estadísticos descriptivos de la versión A en donde la media de los ítems varía entre 2.02 a 3.31. En cuanto a la desviación estándar, este oscila entre .68 y .99. Además, la curtosis y asimetría presentan valores menores a ± 1.5 . Respecto a la versión B la media se encuentra en un rango de 3.22 a 3.36 en los ítems. En cuanto a la desviación estándar, el valor oscilaba entre .68 y .74. Asimismo, la curtosis y asimetría de la mayoría de los ítems presentan valores adecuados ($< \pm 1.5$).

Tabla 2. Análisis exploratorio de ítems.

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

Nota: *M* = Media; *DS* = Desviación estándar; *g1* = Asimetría; *g2* = Curtosis.

Análisis factorial confirmatorio

Como se aprecia en la tabla 3, en la versión A, el Modelo A1 (modelo original) presentó adecuados valores en la mayoría de los índices de ajuste ($CFI=.946$, $TLI=.930$, $RMSEA=.165$ [IC 90%: .150–.180], $SRMR=.060$). Asimismo, se planteó el Modelo A2, donde se agregó un factor específico para los ítems negativos. Este modelo presentó algunas mejoras en la mayor parte de los índices de ajuste ($CFI=.977$, $TLI=.965$, $RMSEA=.117$ [IC 90%: .101–.134], $SRMR=.038$). También, se planteó el Modelo A3, donde se agregó un factor específico para los ítems positivos. En este modelo se evidenció que los índices de ajuste fueron satisfactorios en su mayoría ($CFI=.972$, $TLI=.959$, $RMSEA=.127$ [IC 90%: .111–.144], $SRMR=.039$). Finalmente, se evaluó el ajuste del Modelo A4, de dos dimensiones correlacionadas, el cual mostró mejoras en la mayor parte de los índices de ajuste a diferencia de los modelos anteriores ($CFI=.976$, $TLI=.968$, $RMSEA=.111$ [IC 90%: .095–.127], $SRMR=.040$). Asimismo, se encontró que los modelos evaluados de la versión A mostraron valores que superaron el punto de corte en el $\chi^2/g1$, $RMSEA$ y $WRMR$.

Frente a estos resultados se evaluó el ajuste de la versión B, el cual presentó adecuados valores en la mayoría de los índices de ajuste ($CFI=.975$, $TLI=.968$, $RMSEA=.138$ [IC 90%: .123–.154], $SRMR=.049$). No obstante, similar a lo resultados de los modelos de la ver-

sión A los valores fueron superiores en el χ^2 / gl , *RMSEA* y *WRMR*. Es importante señalar que los pesos factoriales de los ítems de la versión B en su mayoría son mejores que los modelos planteados en la versión A.

Tabla 3. *Índices de ajuste de los modelos y las estimaciones estandarizadas de los parámetros.*

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

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

Modelo A1 = Modelo unidimensional original, Modelo A2 = Modelo con control del efecto para los ítems negativos, Modelo A3 = Modelo con control del efecto para los ítems positivos, Modelo A4 = Modelo de dos factores relacionados, Modelo B = Modelo unidimensional de solo ítems positivos

Discusión y Conclusiones

El estudio tuvo como objetivo general evaluar la estructura factorial de la escala EAR y proponer una nueva versión de la EAR de solo ítems positivos para superar el efecto de método asociado a los ítems negativos. Para lo cual se planteó dos versiones de la misma escala:

la versión A que contiene ítems positivos y negativos y la versión B que tiene solo ítems positivos.

Para evaluar la estructura factorial del EAR se realizó un AFC, donde se compararon los modelos de la versión A y B. El modelo A1 que sigue el planteamiento original de Rosenberg (1965) presentó adecuados valores en la mayoría de índices de ajuste ($CFI=.95$, $TLI=.93$), sin embargo obtuvo valor no aceptable en el $RMSEA=.165$ y $WRMR$. Este resultado no coincide a lo reportado en estudios previos (Martín-Albo et al., 2007; Pullmann y Allik, 2000; Tinakon y Nahathai, 2012), donde un modelo unidimensional si se ajusta a los datos.

En relación a ello, se evaluó el ajuste de los modelos A2 (modelo para controlar el efecto de método de los ítems negativos) y A3 (modelo para controlar el efecto de método de los ítems positivos), donde se evidenció una mejora en los índices de ajuste, ya que se añadió un factor específico para los ítems positivos y negativos, no obstante, se encontró valores altos en el $RMSEA$ y $WRMR$. Estos hallazgos no concuerdan con otros estudios en donde respaldan que añadir un factor adicional favorece en el ajuste de los datos (de Lima y de Souza, 2019; Sánchez-Villena et al., 2021).

El modelo A4 mostró un mejor ajuste respecto a los modelos anteriores ($CFI=.98$, $TLI=.97$, $RMSEA=.111$ [IC 90%: .095 – .127]), pese a ello similar a los resultados de los modelos de la versión A, se encontraron un excesivo $RMSEA$ Y $WRMR$. De tal forma, esto no coincide con estudios previos (De León Ricardi y García, 2016; Jurado et al., 2015; Kielkiewicz et al., 2020).

Por último, se evaluó la propuesta del modelo B, de solo ítems positivos, que mostró similares índices de ajuste y valores superiores a .80 en las cargas factoriales que los modelos de la versión A. Además, el modelo B mostró un mejor ajuste a los datos en comparación con el estudio de Greenberger et al. (2003) en donde los índices de ajuste en su mayoría fueron inferiores en el modelo de ítems positivos. Por otra parte, el modelo B muestra una ventaja frente a los modelos de la versión A, lo que se debe a la ausencia de ítems negativos. Varios estudios han reportado que utilizar ítems negativos disminuye los valores en la confiabilidad (Menold, 2019; Salazar, 2015; Suárez-Alvarez et al., 2018), influye en la variabilidad de las respuestas obtenidas produciendo una menor varianza (Suárez-Alvarez et al., 2018) y provoca puntajes mas bajos para los items negativos (Weems et al., 2003). Además, genera que los

modelos unidimensionales no ajusten adecuadamente a los datos y originen otros factores (Suárez-Alvarez et al., 2018; Woods, 2006). Otra ventaja del modelo B, es que ya no es necesario añadir un factor adicional en los modelos de AFC para controlar el efecto de método de los ítems negativos. Mejorando así, la práctica de añadir factores adicionales al estudiar la estructura factorial de la escala (de Lima y de Souza, 2019; Sánchez-Villena et al., 2021).

Es importante señalar que los problemas asociados a los ítems negativos no se dan de la misma manera en todas las culturas (Wong et al., 2003). Así, en Norteamérica el uso combinado de ítems positivos y negativos en una misma escala si funciona bien (Wong et al., 2003). En cambio, en países de Latinoamérica, esta práctica genera inconsistencias en la respuesta de las personas (Marin et al., 1992). Esto podría explicar porque la versión original de la escala desarrollado en Estados Unidos si se ajusta bien a un modelo unidimensional. Y porque en los estudios realizados en Latinoamérica el planteamiento original de ítems negativos y positivos ha generado la aparición de factores adicionales para controlar los ítems negativos (de Lima y de Souza, 2019; Sánchez-Villena et al., 2021) o la aparición de modelos bidimensionales (De León Ricardi y García, 2016; Jurado et al., 2015).

Tomando en cuenta lo anterior, otra ventaja del modelo B es que su unidimensionalidad concuerda con la propuesta teórica de Rosenberg (1965) quien clasifica a las personas dentro de un único rango continuo que varía desde una muy baja o muy alta autoestima. Por lo expuesto se opta por escoger la propuesta del modelo B de solo ítems positivos para la evaluación de la autoestima.

Respecto a la fiabilidad de los modelos de la versión A, se encontraron valores inferiores en el alfa de Cronbach y el coeficiente omega, dichos valores se deben principalmente a que los ítems negativos disminuyen los valores de la confiabilidad en la escala (Menold, 2019; Salazar, 2015; Suárez-Alvarez et al., 2018). Por otra parte, al evaluar la fiabilidad del modelo B se encontraron que los valores en los índices fueron aceptables ($\alpha = .965$, $\omega = .960$).

El estudio muestra algunas limitaciones. En primer lugar, se realizó un muestreo intencional el cual impide generalizar los resultados obtenidos en el estudio. En segundo lugar, no se realizó un estudio de invarianza factorial según sexo de los participantes, ya que en la muestra del modelo B había una diferencia significativa entre la cantidad de varones y mujeres lo que generaría resultados alterados, así mismo hay varios estudios que han aportado evi-

dencias que esta variable puede afectar las puntuaciones de la escala (Caballo y Salazar, 2018; Ruiz-González et al., 2018). En tercer lugar, se trabajó con universitarios provenientes de algunas universidades públicas y privadas, por tanto, se sugiere para futuros estudios ampliar el tamaño de muestra.

Se concluye que la versión B, una propuesta de solo ítems positivos, supone una mejora en la medición de la autoestima ya que no tiene ítems negativos y además concuerda con el planteamiento unidimensional de Rosenberg. Esta versión se denominó Escala de Autoestima de Rosenberg-P (EAR-P).

Referencias

- 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](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](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 adolescentes 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 Psicología*, 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 of Helsinki Ethical Principles for Medical Research Involving Human Subjects. *Clinical Review & Education*, 310(20), 2191–2194. <https://doi.org/10.1001/jama.292.11.1359>

Recibido: 06-05-2021

Aceptado: 10-08-2021