ISSN en línea 1688-4221 *Ciencias Psicológicas* julio-diciembre 2022; 16(2), e-2810
doi: 10.22235/cp.v16i2.2810

Validez e invarianza factorial del Índice de Reactividad Interpersonal en universitarios peruanos

Validity and factorial invariance of the Interpersonal Reactivity Index in Peruvian university students

Validade e invariância fatorial do Índice de Reatividade Interpessoal em universitários peruanos

Mirian Grimaldo Muchotrigo[1], ORCID 0000-0003-4197-1906
Jossue Correa-Rojas[2], ORCID 0000-0002-4166-7210
Eduardo Manzanares Medina[3], ORCID 0000-0002-3377-311X
Karen Macavilca Milera[4], ORCID 0000-0001-6198-6318

[1] *Universidad Peruana de Ciencias Aplicadas, Perú*
[2] *Universidad Peruana de Ciencias Aplicadas, Perú*
[3] *Universidad Peruana de Ciencias Aplicadas, Perú*
[4] *Universidad Peruana de Ciencias Aplicadas, Perú*

Resumen

El Índice de Reactividad Interpersonal (IRI) constituye una de las medidas de autoinforme más utilizadas en jóvenes y adultos para medir la empatía, ya que considera en su composición aspectos cognitivos y afectivos. En esta investigación se presenta la validez estructural, la invarianza y fiabilidad de la medida en universitarios peruanos. Se seleccionaron intencionalmente 859 estudiantes de Psicología, entre 17 y 39 años, de dos universidades de Lima Metropolitana. El análisis factorial confirmatorio realizado sobre tres modelos: modelo 1, replica la estructura original; modelo 2, analiza la estructura original sin los ítems redactados en forma negativa; modelo 3, se excluyeron los ítems inversos y los que presentaban varianza irrelevante de constructo (ítem 9). Este último modelo es el que presenta un mejor ajuste (CFI = .924; RMSEA = .091; SRMR = 0.058). Sobre este modelo se estableció la invarianza configuracional, métrica, escalar y residual según edad. Adicionalmente, se probó un modelo de segundo orden que demostró la presencia de los factores latentes empatía cognitiva y reacciones emocionales, este modelo presenta un ajuste adecuado (CFI = .957; RMSEA = .078; SRMR = 0.068). Se concluye que la versión reducida del IRI es una medida válida y fiable en universitarios peruanos.

Palabras clave: análisis factorial; propiedades psicométricas; confiabilidad y validez; empatía; universitarios peruanos

Abstract

The Interpersonal Reactivity Index (IRI) is one of the most used self-report measures in young people and adults to measure empathy, it considers cognitive and affective aspects in its composition. This research presents the structural validity, invariance and reliability of the measure in Peruvian university students. 859 psychology students, between 17 and 39 years old, from two universities in Metropolitan Lima were intentionally selected. The confirmatory factor analysis performed on three models: Model 1, replicates the original structure, Model 2, analyzes the original structure, but negatively worded items and model 3 were excluded, inverse items and those with irrelevant variance were excluded. construct (item 9). This last model is the one with the best fit (CFI = .924; RMSEA = .091; SRMR = 0.058). On this model, the configurational, metric, scalar and

Esta obra está bajo una licencia de Creative Commons Reconocimiento 4.0 Internacional

residual invariance according to age was established. In addition, a second order model was tested that demonstrated the presence of the latent factors Cognitive Empathy and Emotional Reactions, this model presents an adequate fit (CFI = .957; RMSEA = .078; SRMR = 0.068). It is concluded that the reduced version of the IRI is a valid and reliable measure in Peruvian university students.
Keywords: factorial analysis; psychometric properties; validity and reliability; empathy; Peruvian university students

Resumo

O Índice de Reatividade Interpessoal (IRI) é uma das medidas de autorrelato mais utilizadas em jovens e adultos para mensurar a empatia, e considera aspectos cognitivos e afetivos em sua composição. Nesta pesquisa apresenta-se a validade estrutural, a invariância e a confiabilidade da medida em estudantes universitários peruanos. Foram selecionados intencionalmente 859 estudantes de psicologia, entre 17 e 39 anos, de duas universidades da região metropolitana de Lima. A análise fatorial confirmatória realizada em três modelos: Modelo 1, replica a estrutura original, Modelo 2, analisa a estrutura original, mas foram excluídos os itens redigidos negativamente, e o Modelo 3, se excluíram os itens inversos e os que apresentavam variância de construto irrelevante (item 9). Este último modelo é o que apresenta melhor ajuste (CFI = .924; RMSEA = .0091; SRMR = 0.058). Sobre este modelo foram estabelecidas as invariâncias configuracional, métrica, escalar e residual segundo a idade. Adicionalmente, foi testado um modelo de segunda ordem que demonstrou a presença dos fatores latentes Empatia Cognitiva e Reações Emocionais, este modelo apresenta um ajuste adequado (CFI = .957; RMSEA = .078; SRMR = 0.068). Conclui-se que a versão reduzida do IRI é uma medida válida e confiável em estudantes universitários peruanos.
Palavras-chave: análise fatorial; propriedades psicométricas; confiabilidade e validade; empatia; estudantes universitários peruanos

Recibido: 31/01/2021 Aceptado: 11/08/2022

Correspondencia: Jossue Correa Rojas, Universidad Peruana de Ciencias Aplicadas, Perú. E-mail: jossue.correa@upc.pe

La empatía es una respuesta emocional y cognitiva procedente de la comprensión de la condición de otra persona (Eisenberg, 2000), permite afrontar exigencias vitales (Cañero et al., 2019), motiva la cooperación (Zaki, 2018) y la moralidad (De Waal, 2010). Ha sido estudiada desde tres perspectivas, la cognitiva, la emocional y la integradora (Lorente, 2014; Palma, 2013). La visión integradora de la empatía cobra fuerza en la década de los 80 (Parra et al., 2012). Un representante importante de esta perspectiva es Davis (1980), quien la define como una reacción a la experiencia observada, y plantea la presencia de diferentes factores que intervienen en el constructo.

Con base en esta estructura teórica, Davis (1980) diseñó el Índice de Reactividad Interpersonal (IRI) que constituye una de las medidas de autoinforme más utilizadas en jóvenes y adultos para medir este atributo (Jordan et al., 2016; Lucas-Molina et al., 2017). El IRI está conformado por cuatro subescalas que miden dos componentes cognitivos: (a) toma de perspectiva, que constituye la tendencia a adoptar el punto de vista del otro y (b) fantasía, considerada como la tendencia a identificarse con los sentimientos y acciones de personajes ficticios, y dos componentes afectivos: (c) preocupación empática, referida a

los sentimientos de preocupación por las personas necesitadas, y (d) malestar personal, referido a los sentimientos de ansiedad personal frente a la angustia de otros.

Para el presente estudio se usó la adaptación española de Mestre et al. (2004), que analiza la validez del IRI a partir de la relación con otros constructos con los cuales correlacionó positivamente, tales como conducta prosocial y estilos de razonamiento prosocial, y negativamente con conducta agresiva e inestabilidad emocional. En cuanto a la confiabilidad, arrojó los siguientes valores: fantasía .70, preocupación empática .65, malestar personal .64 y toma de perspectiva .56.

El IRI ha sido adaptado y validado en diferentes contextos. En Asia (Siu & Shek, 2005), se obtuvo un índice de confiabilidad que fluctuaba entre .65 y .70 para sus dimensiones; además de definir tres factores mediante el análisis factorial. Wang et al. (2020) señalan que los investigadores a menudo utilizan los puntajes del IRI de manera flexible en sus estudios basándose en las diferentes construcciones de empatía. Así, algunos lo asumen como constructo general y suman las cuatro dimensiones del IRI para obtener una puntuación de empatía (Sun et al., 2018), aunque sin evidencia empírica que justifique dicho proceso (Dueber & Toland, 2021). Además, los investigadores que reconocen la empatía como un constructo dualista combinan las dimensiones de preocupación empática y malestar personal en un factor de empatía afectiva y las dimensiones de toma de perspectiva y fantasía en un factor de empatía cognitiva (Fan & Hu, 2017). En otros casos, combinan las dimensiones de preocupación empática y malestar personal para representar la empatía afectiva, pero solo toma de perspectiva como empatía cognitiva (He & Zhu, 2016). También se ha encontrado que solo consideran la dimensión preocupación empática y toma de perspectiva como representantes de la empatía afectiva y la empatía cognitiva, respectivamente (Luo et al., 2013).

Respecto a lo anterior, en Estados Unidos, Pulos et al. (2004) encontraron que, si se desea una escala de empatía de orden superior al IRI, resultan útiles las subescalas de preocupación empática, toma de perspectiva y fantasía. De esta manera, se deja de lado la dimensión malestar personal, debido a su baja correlación con las otras dimensiones. Recientemente, se ha encontrado que la dimensión malestar personal es deficiente en la validez de constructo (Murphy et al., 2018). En Canadá, en un estudio realizado con jóvenes adultos, consideraron el puntaje de las subescalas toma de perspectiva y preocupación empática como la puntuación general (Nicol & Rounding, 2013). En México, Ahuatzin et al. (2019) encontraron la existencia de cuatro factores, aunque algunos ítems se desplazaron de una dimensión a otra, tal como se halló en el estudio de Pérez-Albéniz et al. (2003).

En Holanda, se examinaron las propiedades psicométricas del IRI mediante un análisis factorial confirmatorio (AFC) y los resultados revelaron que hay necesidad de realizar una adecuación psicométrica de los puntajes en términos de la estructura de los factores. Sin embargo, en general la consistencia interna, y la estructura factorial de las puntuaciones de la versión holandesa sugieren que es un instrumento útil (De Corte et al., 2007).

En América Latina, se encuentran diversos estudios instrumentales en diferentes países. En Colombia, Bernal et al. (2015) determinaron las propiedades psicométricas del IRI y concluyeron que debían eliminar cinco ítems que tenían en común el estar redactados negativamente, para mejorar los índices de confiabilidad de cada una de las subescalas. De la misma manera, García-Barrera et al. (2017) examinaron la estructura interna del IRI en excombatientes, y encontraron que los ítems negativos presentaron cargas factoriales bajas.

En Argentina, Richaud de Minzi (2008) validó el instrumento en población infantil mediante la verificación de la estructura interna a través de análisis factoriales exploratorios realizados consecutivamente en muestras independientes, estos mostraron que el IRI conserva su estructura interna básica, aunque se observó un comportamiento inestable de los ítems redactados de forma negativa; se encontró un alfa de .70 para la versión total. Por su parte, Müller et al. (2015) analizaron la estructura factorial mediante un AFC con una muestra de adultos, y obtuvieron un modelo de cuatro dimensiones ajustado al modelo teórico, con una fiabilidad adecuada para cada dimensión. En Chile, Fernández et al. (2011) aplicaron el AFC demostrando un buen ajuste del modelo en línea a la propuesta de Davis (1980), con una adecuada consistencia interna.

En el Perú (Díaz et al., 2015), la estructura interna no demostró solvencia y en cuanto a la fiabilidad obtuvo coeficientes alfa igual a .41 para la dimensión toma de perspectiva, .49 para fantasía, .48 en preocupación empática y .52 en la dimensión malestar personal. Estos resultados sugieren que los ítems redactados negativamente podrían influir en las respuestas debido a las diferencias culturales e incluso lingüísticas.

Una de las propiedades psicométricas menos exploradas en el IRI es su invarianza respecto a la edad. En este sentido, para establecer la invarianza de la medida implica que la estructura, los ítems, las medidas y los errores de medida sean similares en ambos grupos. Es decir, solo sería posible interpretar de una misma manera las puntuaciones del instrumento en diferentes grupos (Byrne, 2008). Así, dado que en el presente estudio se realizó sobre una muestra de estudiantes de Psicología, es de esperarse que en las primeras etapas de su formación no se establezcan variaciones respecto a la empatía. Sin embargo, es posible que la exposición a actividades y conocimientos propios de la carrera puedan generar cambios en el atributo conforme va avanzando en su formación. Es por ello que se ha considerado la edad como criterio, ya que esta característica está relacionado con el ciclo académico de los universitarios.

La medición de la empatía en estudiantes de Psicología resulta importante puesto que el dominio de la capacidad empática es un elemento esencial en la formación de los profesionales (Hernández, 2019). Además, el IRI constituye una herramienta muy usada en la evaluación de la empatía, por su carácter multidimensional (Müller et al., 2015). En ese sentido, el objetivo del presente estudio fue brindar evidencias de la estructura factorial del IRI. Se verificó la existencia de un factor general subyacente a las dimensiones de esta escala, se analizó la invarianza según edad y, por último, la confiabilidad por consistencia interna.

Método

Participantes

Se seleccionaron intencionalmente 859 estudiantes de psicología provenientes de dos universidades de Lima, 31.8 % estudiaban de una universidad pública y 68.2 % en una privada. El 77.3 % fueron mujeres, con una edad promedio de 20.36 (*DE* = 2.69), y 22.7 % varones, con una edad promedio de 21.29 (*DE* = 2.64), las edades de la muestra total oscilaron entre 17 y 29 años (*M* = 20.57; *DE* = 2.70). Para la distribución de los participantes en los grupos de edad se tuvo en consideración el concepto de adultos emergentes, cuyas edades van desde los 14 a los 29 años; en este caso, se subdividió la muestra en dos grupos de adultos emergentes: el primer grupo de 17 a 20 años (55.50 %), en donde existe un predominio de estudiantes de primeros ciclos, con características más hacia la adolescencia; y un segundo grupo de 21 a 29 años (44.50 %), con un predominio de estudiantes de ciclos intermedios y finales, cuyas características estaban más orientadas a la adultez. En cuanto a su distribución por ciclo año de estudio, se encontró

que el 26.4 % eran de primer año, 14.7 % de segundo año, 23.1 % de tercer año, 17.8 % de cuarto año, y 18 % de quinto año.

Instrumento

El Índice de Reactividad Interpersonal (IRI; Davis, 1983) es una medida que explora la empatía cognitiva y afectiva, mediante un formato de respuesta tipo Likert con cinco opciones: *No me describe bien* (1), *Me describe un poco* (2), *Me describe bien* (3), *Me describe bastante bien* (4) y *Me describe muy bien* (5). Consta de 28 ítems que permiten medir las diferencias individuales del constructo empatía, mediante las siguientes cuatro subescalas (7 ítems cada una): toma de perspectiva (TP) y fantasía (F) (componente cognitivo), y preocupación empática (PE) y malestar personal (MP) (componente emocional). Para la presente investigación se ha empleado la adaptación española de Mestre et al. (2004), quienes analizaron las evidencias de validez basada en un análisis discriminante, encontrando correlaciones canónicas igual a .436 entre las subescalas del IRI que permiten una clasificación correcta de varones en 70.2 % y 69.7 % en mujeres. Estos autores señalan que tanto el sexo como la edad introducen sesgos en la medición de la empatía. La validez en relación con otros constructos se verificó con la conducta prosocial (.312), agresividad (-.171) y razonamiento hedonista (-.142). La consistencia interna de las dimensiones resulta parcialmente aceptable TP (.56), F (.70), PE (.64) y MP (.68).

Procedimiento

La administración del instrumento fue realizada en los meses de abril y mayo del 2019, en diferentes grupos y de forma colectiva. Se consideraron las recomendaciones y normativas para la aplicación de pruebas propuestas por la International Test Commission (2001), con el objetivo de minimizar la varianza irrelevante al constructo proclive a ocurrir durante la administración de pruebas psicológicas. Los participantes completaron las pruebas y firmaron el consentimiento informado, en el que se detallaban aspectos relacionados a su participación, con énfasis en el carácter voluntario y la libertad para participar del estudio, por ello se comunicó que la administración sería anónima y que los resultados serían confidenciales.

Análisis de datos

En primer lugar se analizó la presencia de datos faltantes, al constatar su ausencia se procedió a verificar los *outliers*. Luego, se exploró la frecuencia absoluta de las opciones de respuesta a fin de identificar que todas resulten funcionales, para ello se siguió el criterio de Linacre (2002), quien refiere que frecuencias inferiores a 10 pueden resultar problemáticas. A continuación, se analizaron las medidas descriptivas de los ítems considerando la media (*M*), desviación estándar (*DE*), asimetría (g_1) y curtosis (g_2), a través de estos últimos fue posible evaluar la distribución de los ítems considerando ± 2.0 como valor marginal de normalidad (Reed & Wu, 1974).

Se analizó la estructura interna, mediante un análisis factorial confirmatorio, probándose 3 modelos: modelo 1, incluye 28 reactivos; modelo 2, excluye los ítems inestables; y modelo 3, excluye todos los ítems inversos. Se utilizó el método *Weighted Least Square Mean and Variance Adjusted* (WLSMV), debido a la naturaleza ordinal de los datos. Se revisó el chi-cuadrado entre los grados de libertad, el Error cuadrático medio de aproximación (RMSEA) y el Resido cuadrático medio estandarizado (SRMR), ambos con valores por debajo de .08. También se corroboraron el *Comparative Fit Index* (CFI) y el Índice Tucker-Lewis (TLI), cuyos valores aceptables se encuentran por encima de .90 (Kline, 2016; Xia & Yang, 2018). Adicionalmente, se calculó la *Average Variance*

Extracted (AVE), con valores satisfactorios de alrededor de .50. Asimismo, se exploró la invarianza de la medida, considerando deltas menores a .01 para ΔCFI y ΔRMSEA (Byrne, 2008).

Por último, se analizó el coeficiente omega categórico (ω) con sus intervalos de confianza. Debido al efecto del error aleatorio, es recomendable establecer un rango para estimar el valor real de la fiabilidad (Ventura-León, 2018). Se calculó la *Compositive Reliability* (CR; Zumbo et al., 2007). Los análisis se realizaron con el IBM SPSS, versión 25 y el RStudio versión 3.3.2 (RStudio Team, 2015), empleándose el paquete Lavaan (Rosseel, 2012).

Resultados

En la Tabla 1 se presentan las frecuencias absolutas de las opciones de respuesta de cada uno de los reactivos del IRI, observándose que estas resultan funcionales, con excepción del ítem 27, en el que la opción de respuesta *Me describe muy bien* presenta una frecuencia de respuestas baja. Sin embargo, esto no afectó los cálculos realizados.

Tabla 1
Estadísticos descriptivos del IRI

	Frecuencias absolutas					**Medidas descriptivas**					
	1	2	3	4	5	*Min*	*Max*	*M*	*DE*	*g1*	*g2*
iri1	103	191	260	185	120	1	5	3.03	1.22	0.00	-0.89
iri2	45	108	261	312	133	1	5	3.44	1.06	-0.44	-0.32
iri3	29	73	220	326	211	1	5	3.72	1.03	-0.62	-0.07
iri4	34	83	193	252	297	1	5	3.81	1.13	-0.68	-0.36
iri5	223	235	214	129	58	1	5	2.49	1.22	0.40	-0.81
iri6	121	289	295	126	28	1	5	2.59	1.01	0.24	-0.43
iri7	54	179	294	222	110	1	5	3.18	1.10	-0.06	-0.67
iri8	21	52	239	356	191	1	5	3.75	0.95	-0.60	0.21
iri9	18	75	261	340	165	1	5	3.65	0.96	-0.44	-0.13
iri10	91	291	297	127	53	1	5	2.72	1.04	0.34	-0.35
iri11	15	54	228	359	203	1	5	3.79	0.93	-0.56	0.09
iri12	30	88	234	292	215	1	5	3.67	1.07	-0.51	-0.36
iri13	71	185	229	210	164	1	5	3.25	1.22	-0.12	-0.98
iri14	28	86	189	291	265	1	5	3.79	1.09	-0.66	-0.29
iri15	49	94	260	271	185	1	5	3.52	1.11	-0.45	-0.42
iri16	177	222	245	146	69	1	5	2.66	1.21	0.25	-0.85
iri17	130	254	275	156	44	1	5	2.69	1.09	0.18	-0.66
iri18	27	58	127	190	457	1	5	4.15	1.10	-1.18	0.50
iri19	85	268	356	111	39	1	5	2.71	0.97	0.25	-0.09
iri20	74	208	297	205	75	1	5	3.00	1.08	0.01	-0.64
iri21	15	55	279	338	172	1	5	3.69	0.92	-0.40	-0.04
iri22	58	129	295	211	166	1	5	3.35	1.15	-0.21	-0.68
iri23	96	190	285	182	106	1	5	3.01	1.17	0.02	-0.78
iri24	236	329	224	55	15	1	5	2.17	0.96	0.57	-0.07
iri25	56	146	320	256	81	1	5	3.19	1.04	-0.22	-0.40
iri26	62	166	279	226	126	1	5	3.22	1.13	-0.13	-0.72
iri27	335	315	159	42	8	1	5	1.92	0.92	0.82	0.18
iri28	29	102	251	297	180	1	5	3.58	1.05	-0.41	-0.42

Nota. *M*: Media; *DE*: Desviación estándar; *g1*: Coeficiente de asimetría; *g2*: Coeficiente de curtosis.

Evidencias de validez basada en la estructura interna

En la Tabla 2 se presentan los tres modelos evaluados. El primer modelo corresponde a la propuesta original de cuatro factores. Este modelo presenta cargas factoriales inestables en los ítems redactados en forma negativa (< .40), asimismo, presenta valores AVE que oscilan entre .20 y .36 y sus índices de bondad de ajuste resultan inadecuados (CFI = .521; RMSEA = .133; SRMR = .139). Por esta razón, estos reactivos no fueron considerados para el tercer modelo, que obtuvo cargas factoriales adecuadas e índices de ajuste aceptables (CFI = .924; RMSEA = .073; SRMR = .058). Estos resultados demuestran que la versión reducida del IRI es una medida que sostiene su estructura original, pero que mejora cuando se excluyen los ítems inversos (7, 15, 7, 12, 4 ,14, 18, 13 y 19) y los enunciados redundantes (ítem 9) (ver Tabla 3).

Tabla 2

Análisis factorial confirmatorio del IRI

	Modelo 1					**Modelo 2**					**Modelo 3**				
	F1	F2	F3	F4	Λ^2	F1	F2	F3	F4	Λ^2	F1	F2	F3	F4	Λ^2
iri8	.65				0.43	.60				0.35	.59				0.35
iri11	.74				0.54	.71				0.50	.71				0.50
iri21	.69				0.47	.64				0.40	.64				0.41
iri25	.65				0.43	.64				0.40	.65				0.42
iri28	.65				0.42	.65				0.42	.65				0.42
iri3	.08				0.01	-				-	-				-
iri15	.19				0.04	-				-	-				-
iri1		.62			0.38		.66			0.43		.66			0.43
iri5		.65			0.42		.63			0.39		.63			0.40
iri16		.72			0.51		.75			0.57		.75			0.57
iri23		.78			0.60		.77			0.59		.77			0.59
iri26		.73			0.54		.71			0.50		.71			0.50
iri7		-.03			0.00		-			-		-			-
iri12		.06			0.00		-			-		-			-
iri2			.53		0.28			.56		0.31			.56		0.32
iri9			.51		0.26			.54		0.28			-		-
iri20			.64		0.40			.61		0.37			.65		0.42
iri22			.65		0.42			.70		0.49			.73		0.53
iri4			.03		0.00			-		-			-		-
iri14			.09		0.01			-		-			-		-
iri18			.08		0.01			-		-			-		-
iri6				.66	0.43				.60	0.35				.60	0.35
iri10				.71	0.50				.70	0.48				.69	0.48
iri17				.78	0.61				.77	0.59				.77	0.59
iri24				.73	0.54				.69	0.47				.70	0.48
iri27				.62	0.39				.60	0.36				.61	0.37
iri13				.18	0.03				-	-				-	-
iri19				.12	0.01				-	-				-	-
F1	-					-									
F2	.26	-				.32	-				.32				
F3	.64	.65	-			.63	.63	-			.47	.63			
F4	-.15	.52	.52	-		.04	.54	.58	-		.04	.54	.62		
AVE	0.33	0.35	0.20	0.36		0.42	0.50	0.29	0.46		0.42	0.50	0.42	0.46	
CR	0.74	0.73	0.53	0.76		0.78	0.83	0.70	0.81		0.78	0.83	0.69	0.81	

Nota. Λ^2 = Determinante; AVE: Average Variance Extracted; CR: Composite Reliability.

Tabla 3
Índices de bondad de ajuste

	SB-χ2(*df*)	CFI	TLI	RMSEA [IC 90%]	SMRM	WRMR
Modelo 1	3895,295(344)	.521	.474	.133 [.129-.137]	0.139	3.316
Modelo 2	1042,545(146)	.890	.871	.085 [.080-.089]	0.068	1.920
Modelo 3	714,377(129)	.924	.910	.073 [.068-.078]	0.058	1.633

Nota. SB χ^2 = Satorra-Bentler chi-cuadrado; *df:* grados de libertad.

Invarianza factorial del IRI según edad

Se analizó la invarianza de la versión reducida del IRI (modelo 3) según la edad de los participantes, diferenciando dos grupos: uno de menores de 21 años y otro de mayores de esa edad. Se aprecia que los valores RMSEA y sus variaciones son aceptables (ΔRMSEA < .01), al igual que las variaciones en SRMR (ΔSRMR < .01). En cuanto a los cambios del CFI, se demostró su suficiencia (ΔCFI < .01). Estos hallazgos demuestran que la invarianza de la medida se ha establecido en ambos grupos. En la Tabla 4 se observa que los valores de los índices de ajuste son satisfactorios para la invarianza configuracional, métrica, escalar y estricta, el delta de la variación métrica se encuentra en los márgenes sugeridos por la literatura (Byrne, 2008).

Tabla 4
Diferenciales de los índices de ajuste del IRI

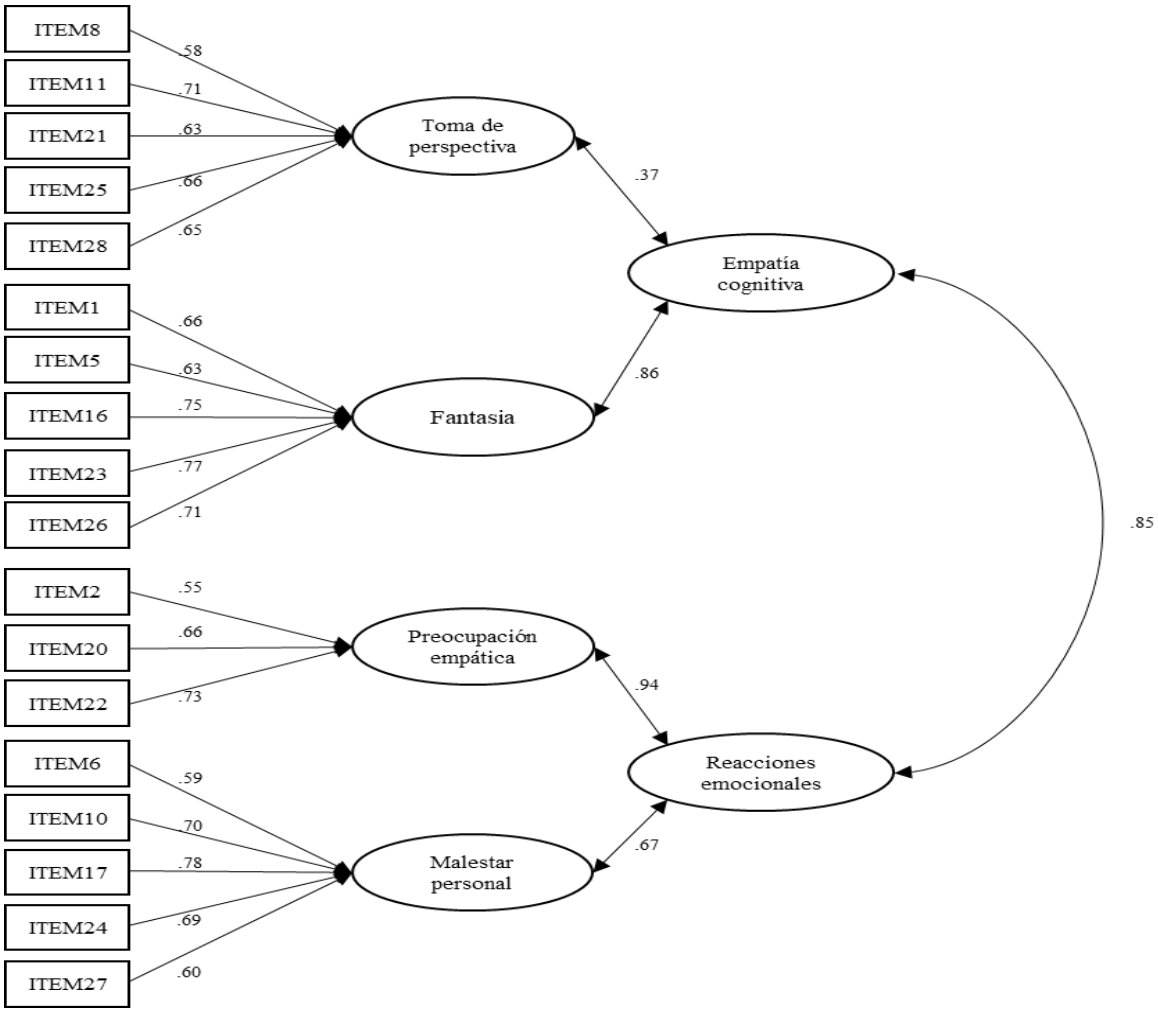
Invarianza	χ2	*df*	CFI	RMSEA	SRMR	Δ *df*	Δ CFI	Δ RMSEA	Δ SRMR
Configuracional	857.311	258.000	.919	.074	0.066				
Métrica	938.165	322.000	.916	.067	0.066	64.000	-0.002	-0.007	0.001
Escalar	1146.619	326.000	.889	.077	0.066	4.000	-0.028	0.010	0.000
Estricta	1169.047	344.000	.888	.075	0.068	18.000	-0.001	-0.002	0.001

En la Figura 1 se presenta un AFC de segundo orden, se establecen las variables latentes sobre las que se agrupan las dimensiones originales del IRI. Se empleó el método de DWLS y se obtuvo un ajuste satisfactorio (CFI = .957; RMSEA = .078 [.073-.083]; SRMR = 0.068; WRMR = 1.902. A través del modelamiento propuesto se confirma la estructura de dos factores latentes (empatía cognitiva y empatía emocional), que contienen a los factores específicos (empatía cognitiva: toma de perspectiva y fantasía, y empatía emocional: preocupación empática y malestar personal) que se describen en el modelo teórico original del IRI.

Consistencia interna del IRI

La consistencia interna se analizó a través del coeficiente omega categórico con sus respectivos intervalos de confianza. Los resultados muestran que las dimensiones del IRI son consistentes en cuanto a sus puntuaciones, para la dimensión toma de perspectiva se obtuvo un omega .757 [.719-.784], en fantasía se encontró un omega igual a .802[.776-.825], la dimensión preocupación empática alcanzó un omega igual a .653[.606-.689] y la dimensión malestar personal un omega de .782[.753-.806]

Figura 1
Análisis factorial confirmatorio de segundo orden del IRI



Discusión

Los hallazgos refieren que la evidencia basada en la validez estructural del instrumento es apropiada luego de la eliminación de los ítems inversos, pues presenta una estructura coherente con el modelo teórico original. Además, se aprecia un incremento favorable de los índices de fiabilidad. Con ello, se demuestra que los ítems redactados negativamente tienen un efecto importante sobre la validez y la fiabilidad (Suárez-Álvarez et al., 2018; Tomás et al., 2012), lo que coincide con lo reportado por Bernal et al. (2015), en una muestra colombiana. Así, se encontraron similares dimensiones. Sin embargo, la presencia de ítems negativos o aquellos que poseen una direccionalidad opuesta a la lógica del constructo (Weijters & Baumgartner, 2012) afectan la estructura del IRI.

Aunque algunos autores, como Nunnally (1978), sugieren la necesidad de incorporar ítems negativos para controlar la aquiescencia; es decir, la tendencia a contestar orientado a un extremo independientemente del contenido (Hidalgo-Rasmussen & Gonzáles-Betanzos, 2019), lo que genera parcialidad y podría producir patrones de respuesta (Van Sonderen et al., 2013). Además, la presencia de ítems inversos puede provocar confusión en los encuestados (Podsakoff et al., 2003), a este fenómeno se le

conoce como *efecto de método* (Abad et al., 2016; Danner et., 2015), y puede producir sesgos en la medición de un atributo (Tomás et. al, 2013).

Respecto a lo anterior, Benson y Hocevar (1985, citados en Weems et al., 2003) señalan que resulta complejo determinar la evidencia basada en el constructo, ya que los ítems se organizan más por un aspecto semántico que por el constructo teórico y, por ello, debe realizarse con precaución. Otros estudios revelan que la combinación de ítems positivos y negativos no logra reducir el sesgo de aquiescencia (Sauro & Lewis, 2011), que podría ser característico en muestras que tengan bajos niveles educativos o cognitivos (Meisenberg & Williams, 2008), razón por la cual no pueden establecer diferencias en los ítems redactados de manera positiva y negativa (Solis, 2015). Aunque también podría evidenciar la pereza o indiferencia en el momento de contestar una prueba (Solis, 2015), lo que se demuestra también en la inconsistencia entre los ítems y podría conducir a un error de medición.

Tal como se ha encontrado en este estudio, los ítems redactados de manera negativa o inversa al parecer no fueron comprendidos o, tal como lo señalan Suárez-Álvarez et al. (2018), podrían haber sido mal interpretados. Por ello, algunos autores sugieren utilizar ítems redactados positivamente (DeVellis, 2003), aunque ello depende de la cultura y nivel educativo (Tomás et al., 2012). Así, se ha encontrado que en Estados Unidos el reconocimiento de los ítems negativos es más viable en comparación con los países asiáticos (Wong et al., 2003). De la misma manera, el estudio de Marín et al. (2002) mostró inconsistencias al combinar ítems positivos y negativos.

Finalmente, luego de la eliminación de los ítems negativos o inversos, la evidencia apoya la validez basada en la estructura interna y consistencia de las puntuaciones del IRI, en línea con el modelo teórico propuesto por Davis (1980). Se aprecia un incremento sustancial en todos los índices de bondad de ajuste cuando se excluyen estos reactivos, donde los valores de AVE también se aproximan a un nivel satisfactorio, y las cargas factoriales alcanzan valores satisfactorios (Kline, 2016). Este crecimiento se extiende además a las medidas de fiabilidad. Se concluye que el Índice de Reactividad Interpersonal es una medida válida y fiable del constructo empatía en los estudiantes universitarios de Lima.

Asimismo, se analizó un modelo de segundo orden para identificar la congruencia teórica del IRI respecto de la propuesta original. Los hallazgos fueron favorables, identificándose que el IRI en su versión reducida conserva la estructura teórica propuesta inicialmente por Davis (1983). Estos resultados coinciden con otros estudios (Mestre et al., 2004; Mestre et al., 1999). Estos hallazgos tienen una implicancia práctica muy importante y es que a partir de lo reportado es posible obtener un puntaje total de las dimensiones que componen el IRI y que este puntaje sea interpretable como la presencia del atributo empatía.

A partir de los resultados obtenidos se establece la invarianza de la medida. Las implicancias prácticas de estos hallazgos aluden a que el constructo es comprendido de igual forma por universitarios de ciclos iniciales y finales. Del mismo modo, se establece la equivalencia del constructo en ambos grupos y, por lo tanto, es posible utilizar el instrumento para realizar comparaciones en muestras similares y bajo los mismos criterios (Byrne, 2008). Esto implica que la formación y los conocimientos propios de la carrera no afectan la autopercepción de este atributo, ni su medición con esta escala.

En lo que refiere a las evidencias de fiabilidad, los hallazgos dan cuenta de la consistencia de las puntuaciones del IRI, que a pesar de la reducción de los ítems redactados en forma negativa, todas las dimensiones alcanzaron coeficientes omega aceptables. Estos resultados coinciden con otros estudios en los que también se señala la estabilidad del IRI (Fernández et al., 2011; Müller et al., 2015). No obstante, es necesario

mencionar que si bien la dimensión preocupación empática obtuvo un omega ligeramente inferior a .70, que a su vez resulta inferior al obtenido en las otras dimensiones, ello no inhabilita el uso de la medida, pues tal como lo señalaron Oviedo y Campos-Arias (2005) estos resultan son valores referenciales. Además, de acuerdo con Cortina (1993), su utilización es pertinente en la medida en que no exista otra más eficiente. En línea con ello, estos resultados son sustancialmente mejores a los valores reportados por Mestre et al. (2004), en la validación realizada en población española.

Una limitación de este estudio fue que los participantes no fueron seleccionados mediante un muestreo probabilístico, lo que restringe la capacidad de generalización de los resultados. Tampoco fue posible calcular la estabilidad de las puntuaciones del IRI, siendo esta propiedad necesaria pues es evidencia complementaria a la consistencia interna, denota la presencia del atributo y su susceptibilidad a ser modificado (Correa-Rojas, 2021).

Se sugiere replicar el estudio en distintos grupos urbanos de diferentes ciudades y carreras, pues se ha encontrado que los niveles de empatía pueden variar en función al sexo, carrera y grado académico (Kidron et al., 2018).

Financiamiento

La Dirección de Investigación de la Universidad Peruana de Ciencias Aplicadas apoyó para la realización de este trabajo de investigación a través del incentivo UPC-IP065-2015.

Referencias

Abad, F. J., Sorrel, M. A., Garcia, L. F., & Aluja, A. (2016). Modeling general, specific, and method variance in personality measures: results for ZKA-PQ and NEO-PI-R. *Assessment, 25*(8), 959-977. https://doi.org/10.1177/1073191116667547

Ahuatzin, A., Martínez-Velázquez, E., García, G., & Vázquez-Moreno, A. (2019). Propiedades psicométricas del Interpersonal Reactivity Index (IRI) en mexicanos universitarios. *Revista Iberoamericana de Psicología, 12*(1), 111-122.

Bernal, E. L., Cetina, M. J. & Romero, L. F. (2015). *Propiedades psicométricas del Interpersonal Reactivity Index (IRI) diseñado por Davis (1980), en una muestra de residentes de la ciudad de Bogotá* [Tesis de Licenciatura inédita]. Universidad Piloto de Colombia.

Byrne, B. M. (2008). Testing for multigroup equivalence of a measuring instrument: a walk through the process. *Psicothema, 20*(4), 872-882.

Cañero, M., Mónaco, G., & Montoya, I. (2019). La inteligencia emocional y la empatía como factores predictores del bienestar subjetivo en estudiantes universitarios. *European Journal of Investigation in Health, Psychology and Education, 9*(1), 19-29. https://dialnet.unirioja.es/servlet/articulo?codigo=7109640

Correa-Rojas, J. (2021). Coeficiente de Correlación Intraclase: Aplicaciones para estimar la estabilidad temporal de un instrumento de medida. *Ciencias Psicológicas, 15*(2), e–2318. https://doi.org/10.22235/cp.v15i2.2318

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology, 78*(1), 98-104. https://doi.org/10.1037/0021-9010.78.1.98

Danner, D., Aichholzer, J., & Rammstedt, B. (2015). Acquiescence in personality questionnaires: Relevance, domain specificity, and stability. *Journal of Research in Personality, 57,* 119-130. https://doi.org/10.1016/j.jrp.2015.05.004

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology, 10*, 85.

Davis, M. H. (1983). Measuring Individual Differences in Empathy: Evidence for a Multidimensional Approach. *Journal of Personality and Social Psychology, 44*(1), 113-126. https://doi.org/10.1037/0022-3514.44.1.113

De Corte, K., Buysse, A., Verhosfstadt, L., & Roeyers, H. (2007). Measuring empathic tendencies: Reliability and validation of the dutch versión of the Interpersonal Reactivity Index, *Psychologica Belgica*, *47*(4), 235-260. https://doi.org/10.5334/pb-47-4-235

De Waal, F. (2010). *The age of empathy: Nature's lessons for a kinder society*. Random House.

DeVellis, R. F. (2003). *Scale development: Theory and applications*. Sage Publications.

Díaz, B., Rodríguez, K., & Santa Cruz, H. (2015). Propiedades del Índice de Reactividad Interpersonal en alumnos del ISTP San Luis de Trujillo. *Revista JANG*, *2*(1), 2-14.

Dueber, D. M. & Toland, M. D. (2021). A bifactor approach to subscore assessment. *Psychological Methods.* https://doi.org/10.1037/met0000459

Eisenberg, N. (2000). Emotion, regulation and moral development. *Annual Review of Psychology, 51*, 665-697. https://doi.org/10.1146/annurev.psych.51.1.665

Fan, M. & Hu, Y. (2017). Status and related factors of empathy in adolescents. *Chinese Mental Health Journal*, *31*(11), 879-884. https://doi.org/10.3969/j.issn.1000-6729.2017.11.009

Fernández, A. M., Dufey, M., & Kramp, U. (2011). Testing the psychometric properties of the Interpersonal Reactivity Index (IRI) in Chile. Empathy in a different cultural context. *European Journal of Psychological Assessment, 27*(3), 179-185. https://doi.org/10.1027/1015-5759/a000065

García-Barrera, M. A., Karr, J. E., Trujillo-Orrego, N., Trujillo-Orrego, S., & Pineda, D. A. (2017). Evaluating empathy in Colombian ex-combatants: examination of the internal structure of the Interpersonal Reactivity Index (IRI) in Spanish. *Psychological Assessment*, *29*(1), 116-122. https://doi.org/10.1037/pas0000331

He, N. & Zhu, Y. (2016). Self-love and other-love: research on the relationships among narcissism, empathy and implicit altruism. *Acta Psychologica Sinica*, *48*, 199-210. https://doi.org/10.3724/SP.J.1041.2016.00199

Hernández, R. M. (2019). La empatía: una necesidad en la educación superior. *Educación Médica*, *20*(Sup. 1), 190. https://doi.org/10.1016/j.edumed.2017.07.026

Hidalgo-Rasmussen, C. A. & Gonzáles-Betanzos, F. (2019). El tratamiento de la Aquiescencia y la Estructura Factorial de la Escala Breve de Resiliencia (BRS) en estudiantes universitarios mexicanos y chilenos. *Anales de Psicología*, *35*(1), 26-32. http://dx.doi.org/10.6018/analesps.35.1.297781

International Test Commission. (2001). *Guidelines on Test Use: Spanish Version*. https://www.intestcom.org/files/guideline_test_use.pdf

Jordan, M. R., Amir, D., & Bloom, P. (2016). Are empathy and concern psychologically distinct? *Emotion, 16*(8), 1107-1116. https://doi.org/10.1037/emo0000228

Kidron, R., Kaganovsky, L., & Baron-Cohen, S. (2018). Empathizing-systemizing cognitive styles: Effects of sex and academic degree. *PLoS ONE, 13*(3). https://doi.org/10.1371/journal.pone.0194515

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4ª ed.). The Guilford Press.

Linacre, J. (2002). What do Infit and Outfit, Mean-square and Standardized mean? *Rasch Measurement Transactions, 16*(2), 878.

Lorente, S. (2014). Efecto de la competencia social, la empatía y la conducta prosocial en adolescentes [Disertación doctoral inédita]. Universitat de València.

Lucas-Molina, B., Perez-Albeniz, A., Ortuño, J., & Fonseca, P. (2017). Dimensional structure and measurement invariance of the Interpersonal Reactivity Index (IRI) across gender. *Psicothema, 29*(4), 590-595. https://doi.org/10.7334/psicothema2017.19

Luo, X., Zheng, D., Liu, W., & Liu, Y. (2013). The relationship between higher vocational school students' empathy and forgiveness: the mediating effect of trait anger. *Chinese Journal of Special Education, 5,* 88-92.

Marín, G., Gamba, R. & Marin, B. (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

Meisenberg, G. & Williams, A. (2008). Are acquiescent and extreme response styles related to low intelligence and education? *Personality and Individual Differences, 44*(7), 1539-1550. https://doi.org/10.1016/j.paid.2008.01.010

Mestre, V., Pérez Delgado, E., Frías, D., & Samper, P. (1999). Instrumentos de evaluación de la empatía. En E. Pérez Delgado & V. Mestre (Coords.), *Psicología moral y crecimiento personal* (pp. 181-190). Ariel.

Mestre, V.; Frías; M. D. & Samper, P. (2004). La medida de la empatía: análisis del Interpersonal Reactivity Index. *Psicothema, 16*(2), 255-260.

Müller, M., Ungaretti, J., & Etchezahar, E. (2015). Evaluación multidimensional de la empatía: adaptación del Interpersonal Reactivity Index (IRI) al contexto argentino. *Revista de Investigación en Psicología Social, 3*(1), 42-53.

Murphy, B. A., Costello, T. H., Watts, A. L., Cheong, Y. F., Berg, J. M., & Lilienfeld, S. O. (2018). Strengths and weaknesses of two empathy measures: a comparison of the measurement precision, construct validity, and incremental validity of two multidimensional indices. *Assessment, 27,* 246-260. https://doi.org/10.1177/1073191118777636

Nicol, A. A. & Rounding, K. (2013). Alienation and empathy as mediators of the relation between social dominance orientation, right-wing authoritarianism and expressions of racism and sexism. *Personality and Individual Differences*, 55, 294-299. https://doi.org/10.1016/j.paid.2013.03.009

Nunnally, J. C. (1978). *Psychometric theory* (2ª ed.). McGrawHill.

Oviedo, H. & Campo-Arias, A. (2005). Aproximación al uso del coeficiente alfa de Cronbach. *Revista Colombiana de Psiquiatría, 34*(4), 572-580.

Palma, J. (2013). *Componentes cognitivos y afectivos en la configuración de la personalidad* [Disertación doctoral inédita]. Universitat de València.

Parra, A. M., Carvajal, N. M., & Moreno, J. H. (2012). *Acoso escolar y empatía en un grupo de adolescentes con trastornos disocial de la conducta* [Proyecto de Grado]. Universidad de La Sabana.

Pérez-Albéniz, A., De Paúl, J., Etxeberría, J., Paz, M., & Torres, E. (2003). Adaptación de Interpersonal Reactivity Index (IRI) al español. *Psicothema, 15*(2), 267-272. https://www.psicothema.com/pdf/1056.pdf

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*(5), 879-903. http://dx.doi.org/10.1037/0021-9010.88.5.879

Pulos, S., Elison, J., & Lennon, R. (2004). The hierarchical structure of the Interpersonal Reactivity Index. *Social Behavior and Personality, 32*, 355-360. https://doi.org/10.2224/sbp.2004.32.4.355

Reed, A. H. & Wu, G. T. (1974). Estimation of a transformation method for estimation of normal range. *Clinical Chemistry, 20*(5), 576-581.

Richaud de Minzi, M. C. (2008). Estudio del IRI de Davis en población infantil argentina. *Revista de Investigación en Psicología, 11*(1), 101-115. https://doi.org/10.15381/rinvp.v11i1.3880

Rosseel, Y. (2012). Lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, *48*(2), 1-36. https://doi.org/10.18637/jss.v048.i02

RStudio Team. (2015). *RStudio: Integrated Development for R*. http://www.rstudio.com/

Sauro, J. & Lewis, J. R. (2011). When designing usability questionnaires, does it hurt to be positive? En *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)* (pp. 2215-2224), Association for Computing Machinery, Nueva York. https://doi.org/10.1145/1978942.1979266

Siu, A. & Shek, D. (2005). Validation of the Interpersonal Reactivity Index in a Chinese Context. *Research on Social Work Practice*, *15*(2), 118-126. http://dx.doi.org/10.1177/1049731504270384

Solis, M. (2015). The dilemma of combining positive and negative items in scales. *Psicothema, 27*(2), 192-199. http://dx.doi.org/10.7334/psicothema2014.266

Suárez-Álvarez, 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

Sun, J., Liu, P., & Li, D. (2018). Attentional characteristics of high-empathy people in processing emotional stimuli and evidence of eye movements. *Journal of Psychological Science, 41*, 1084-1089.

Tomás, J. M., Galiana, L., Hontangas, P., Oliver, A., & Sancho, P. (2013). Evidencia acumulada sobre los efectos de método asociado a ítems invertidos. *Psicológica, 34*, 365-381.

Tomás, J. M., Sancho, S., Oliver, A., Galiana, L., & Melendez, J. (2012). Efecto de método asociado a ítems invertidos vs ítems en negativo. *Revista Mexicana de Psicología*, *29*(2), 105-115.

Van Sonderen, E., Sanderman, R., & Coyne, J. C. (2013). Correction: ineffectiveness of reverse wording of questionnaire items: let's learn from cows in the rain. *PLoS ONE, 8*(9). https://doi.org/10.1371/annotation/af78b324-7b44-4f89-b932-e851fe04a8e5

Ventura-León, J. (2018). Intervalos de confianza para coeficiente omega: propuesta para el cálculo. *Adicciones*, *30*(1). https://doi.org/10.20882/adicciones.962

Wang, Y., Li, Y., Xiao, W., Fu, Y., & Jie, J. (2020). Investigation on the rationality of the extant ways of scoring the Interpersonal Reactivity Index based on confirmatory factor analysis. *Frontiers in Psychology, 11*, 1086. https://doi.org/10.3389/fpsyg.2020.01086

Weems, G., Onwuegbuzie, A., Schreiber, J. & Eggers, S. (2003). Characteristics of respondents who respond differently to positively and negatively worded items on rating scales. *Assessment & Evaluation in Higher Education, 28*(6), 587-607. https://doi.org/10.1080/0260293032000130234

Weijters, B. & Baumgartner, H. (2012). Misresponse to reversed and negated items in surveys: a review. *Journal of Marketing Research, 49*(5), 737-747. https://doi.org/10.1509/jmr.11.0368

Wong, N., Rindfleisch, A., & Burroghs, J. E. (2003). Do reversed worded items confound measures in cross cultural consumer research? The case of material values scale. *Journal of Consumer Research, 30*(1), 72-91. http://dx.doi.org/10.1086/374697

Xia, Y. & Yang, Y. (2018). RMSEA, CFI, and TLI in structural equation modeling with ordered categorical data: The story they tell depends on the estimation methods. *Behavior Research Methods, 51,* 409-428. http://dx.doi.org/10.3758/s13428-018-1055-2

Zaki, J. (2018). Empathy is a moral force. En K. G. Gray, J. (Ed.), *The atlas of moral psychology* (pp. 49-58). Guilford Press.

Zumbo, B. D., Gadermann, A. M., & Zeisser, C. (2007). Ordinal version of coefficients Alpha and Theta for Likert rating scales. *Journal of Modern Applied Statistical Methods, 6*, 21-29.

Cómo citar: Grimaldo Muchotrigo, M., Correa-Rojas, J., Manzanares Medina, E., & Macavilca Milera, K. (2022). Validez e invarianza factorial del Índice de Reactividad Interpersonal en universitarios peruanos. *Ciencias Psicológicas*, *16*(2), e-2810. https://doi.org/10.22235/cp.v16i2.2810

Contribución de los autores: a) Concepción y diseño del trabajo; b) Adquisición de datos; c) Análisis e interpretación de datos; d) Redacción del manuscrito; e) revisión crítica del manuscrito.
M. G. M. ha contribuido con a, b, d; J. C. R. con c, d, e; E. M. M. con b, d, e; K. M. M. con a, b, d.

Editora científica responsable: Dra. Cecilia Cracco.