



Validación en español del Índice de Reactividad Interpersonal –IRI- en estudiantes universitarios colombianos

Luisa Arenas-Estevez¹, Henry Rangel-Quiñonez², Alexandra Cortés-Aguilar³ y Luis Palacio-García³

¹Universidad Industrial de Santander y Universidad Pontificia Bolivariana, Colombia

²Universidad Santo Tomás y Universidad Industrial de Santander, Colombia

³Universidad Industrial de Santander, Colombia

Recibido el 24 de marzo de 2020. Aceptado el 12 de septiembre de 2021

RESUMEN: Este artículo explora las dimensiones factoriales del Índice de Reactividad Interpersonal -IRI (Davis, 1980; 1983), una de las medidas de autoinforme más utilizadas para evaluar los componentes de la empatía. El análisis se realiza con la información de 1779 estudiantes universitarios de primer nivel en los años 2018 y 2019 en una universidad pública colombiana. Con el fin de contribuir a una mejor adaptación del instrumento al contexto colombiano, se valida la estructura de cuatro factores propuesta por el autor originalmente. Para esto se recurre a un Análisis Factorial Confirmatorio (AFC) estimado por Mínimos Cuadrados Ponderados Diagonalizados (MCPD). Los resultados muestran que la versión en español aplicada a los estudiantes colombianos posee características psicométricas similares a la versión original. Sin embargo, se encuentran problemas de ajuste para nueve ítems, al parecer por dificultades de comprensión de la versión en español. El modelo final presenta mejores indicadores de ajuste que indican validez en el instrumento para la evaluación de los diferentes componentes de la empatía.

Palabras clave: Empatía, Validación psicométrica, Análisis Factorial Confirmatorio, Mínimos Cuadrados Ponderados Diagonalizados.

Validation in Spanish of the Interpersonal Reactivity Index –IRI- in Colombian university students

ABSTRACT: This article explores the factorial dimensions of the Interpersonal Reactivity Index -IRI (Davis, 1980; 1983), one of the most widely used self-report measures to assess the components of empathy. The analysis is carried out with the information of 1779 first-level university students in the years 2018 and 2019 in a Colombian public university. In order to contribute to a better adaptation of the instrument to the Colombian context, the original four-factor structure proposed by the author is validated. For this purpose, a Confirmatory Factor Analysis (CFA) estimated by Diagonalized Weighted Least Squares (WLSM) is used. The results show that the Spanish version applied to Colombian students has similar psychometric characteristics to the original version. However, fit problems are found for nine items, apparently due to difficulties in understanding the Spanish version. The final model presents better fit indicators that indicate validity in the instrument for the evaluation of the different components of empathy.

Keywords: Empathy, Psychometric validation, Confirmatory Factor Analysis, Diagonalized Weighted Minimum Squares

Introducción

La empatía es considerada un aspecto fundamental en el funcionamiento social de las personas, al ser parte de los procesos que motivan la conducta prosocial e inhiben la conducta agresiva (Bandura, 1991; Eisenberg et al., 2000; Mestre et al., 2002), puede definirse como la capacidad que tienen los individuos para ponerse en el lugar del otro, una respuesta afectiva que se deriva de la percepción y la comprensión (Eisenberg, 2000). Grohn (2014) define la empatía como “mind reading” ligándola al contexto de toma de decisiones estratégicas, su manifestación varía desde un imaginarse en las circunstancias del otro (concepción egoísta) hasta la toma de perspectiva del sentimiento del otro (concepción altruista). Diferentes estudios coinciden en la relación significativa entre la empatía y la competencia social, en donde se espera que un individuo socialmente competente sea sensible y empático con sus pares (Ceconello y Koller, 2000; Gilar et al., 2008; Howes et al., 1994; Junttila et al., 2006). Además se ha estudiado su relación con variables económicas como el pago de impuestos (Calvet Christian y Alm, 2014), la búsqueda de la equidad (Batson et al., 1999; Kirman y Teschl, 2010), las preferencia y utilidad de los agentes (Salehi-Abari et al., 2019), la interacción estratégica (Grohn et al., 2014; Lerner et al., 2015) entre otras.

Pese a las múltiples aproximaciones para su medición, en la actualidad existe cierto consenso a favor de la empatía como un constructo multidimensional, con fenómenos cognitivos y emocionales (Deutsch y Madle, 1975; Kerem et al., 2001). En esta línea se ubica el concepto de empatía propuesto por Davis (1980, 1983), y su Índice de Reactividad Interpersonal (IRI). Para el autor, las reacciones de los individuos a la experiencia de otro pueden ser de dos clases, cognitiva (la habilidad de entender la perspectiva del otro) y/o visceral o emocional.

El IRI se ha convertido en una de las medidas de autoinforme más utilizadas para evaluar la empatía. Se ha empleado en distintos estudios sobre competencias sociales y comportamiento prosocial (Gilar et al., 2008; Hojat et al., 2005; Konrath, 2014; Mestre et al., 2002; Oceja et al., 2009; Vizoso-Gómez, 2019; Yarnold et al., 1996) y sus propiedades psicométricas han sido ampliamente analizadas en diversas poblaciones; por países, se encuentran validaciones en Alemania (Paulus, 2009), China (Siu y Shek, 2005), Chile (Fernández et al., 2011), Colombia (García-Barrera et al., 2017), Corea (Kang et al., 2009), España (Carrasco et al., 2011; Lucas-Molina et al., 2017; Mestre et al., 2004; Pérez-Albéniz et al., 2003), Estados Unidos (Pulo et al., 2004), Francia (Gilet et al., 2013), Holanda (De Corte et al., 2007; Hawk et al., 2013), Italia (Ingoglia et al., 2016), Suecia (Cliffordson, 2001), entre otros.

En el caso colombiano la estructura factorial del IRI solo ha sido validado por García-Barrera et al., (2017) en una muestra de 548 excombatientes participantes del Programa para la Reincorporación a la Vida Civil. Actualmente, el IRI está siendo utilizado por algunas universidades de Colombia para medir en sus estudiantes la empatía, por tal razón el objetivo principal de esta investigación es realizar una revisión de las propiedades psicométricas del instrumento en los estudiantes de primer nivel en los dos periodos de ingreso de la universidad. Además de ello, esta investigación supone un aporte a la discusión metodológica del algoritmo de estimación de parámetros del Análisis Factorial Confirmatorio (AFC) para variables categóricas, ya que la mayoría de las validaciones del IRI carecen o prestan poca importancia a la selección del algoritmo de estimación de parámetros.

Dentro de las múltiples validaciones psicométricas del IRI realizadas por medio del AFC se encuentran estimaciones por diferentes métodos: Máxima Verosimilitud (MV) (De Corte et al., 2007; Fernández et al., 2011; Hawk et al., 2013; Lucas-Molina et al., 2017; Pérez-Albéniz et al., 2003; Sampaio et al., 2011), Máxima Verosimilitud Robusta (MVR) (Fernández et al., 2011) y Mínimos Cuadrados No-Ponderados (MCNP) (Carrasco et al., 2011). Sin embargo, estos métodos de estimación operan bajo el supuesto de continuidad y de normalidad multivariada de las variables observadas, cualidades que no se cumplen en la mayoría de las pruebas psicológicas que usan constructos con ítems categóricos tipo

Likert, omisión que puede llegar a generar sesgos en las cargas factoriales de los AFC por la incorrecta selección del algoritmo.

De tal forma que las hipótesis de la investigación son H1: el IRI propuesto por Davis no ajusta en su forma original a la población de estudiantes universitarios en Colombia, debido a las diferencias semánticas de los ítems al traducirse del inglés al español y al contexto cultural colombiano. H2: El instrumento del IRI puede ajustar de forma distinta entre las poblaciones de estudiantes admitidos en el primer y segundo periodo académico del año. Debido a dos diferencias fundamentales: puntaje promedio de las pruebas nacionales de ingreso, donde históricamente los estudiantes que ingresan en el segundo periodo tienen puntajes más bajos en comparación con los del primer semestre y los distintos portafolios de programas ofertados entre semestres, pues no todos los programas tienen admisión en ambos semestres

Metodología

Participantes

El tipo de estudio es de carácter instrumental con única medida del IRI para dos grupos de muestra, contó con la participación de estudiantes admitidos en los programas presenciales de pregrado de la sede central de la Universidad Industrial de Santander, Colombia. Las muestras recolectadas fueron de tipo no probabilístico incidental, correspondientes al 35.57% ($n_1 = 413$) de la población del total de admitidos para el segundo semestre del 2018 (M1) y del 84.12% ($n_2 = 1366$) de total del primer semestre del 2019 (M2). Las características de la conformación de las muestras se visualizan en Tabla 1.

Tabla 1.

Descriptivos de las muestras

	<i>Características</i>	<i>M1</i>	<i>M2</i>
	Mujeres	48%	42%
	Edad Promedio	18.1	17.41
	Std. Dev. Edad	1.69	1.60
<i>Facultades</i>	Ciencias	8%	14%
	Ciencias Humanas	25%	25%
	Salud	4%	15%
	Ingenierías Físico-Mecánicas	31%	31%
	Ingenierías Físicoquímicas	32%	15%

Instrumentos

El IRI proporciona evaluaciones separadas de los procesos cognitivos y emocionales, al considerar la empatía como un conjunto de cuatro factores y no como un concepto unidimensional. Su versión original consta de veintiocho ítems en una escala de respuesta tipo Likert con cinco opciones según el grado en que la afirmación describa al sujeto encuestado (0 = No me describe bien; 1 = Me describe un poco; 2 = Me describe bien; 3 = Me describe bastante bien; y 4 = Me describe muy bien).

Los ítems son agrupados en cuatro variables latentes, que conforman dos procesos. El proceso cognitivo incluye las variables Toma de Perspectiva (TP), que refleja la capacidad de los individuos de adoptar la posición de otros sujetos (ítems 3-8-11-15-21-25-28) y Fantasía (FS) que mide la tendencia a trasponerse imaginativamente a los sentimientos y acciones de personajes ficticios (ítems 1-5-7-12-16-23-26). Por otra parte, el proceso emocional contempla las variables Preocupación Empática (PE) que

mide el grado en que el evaluado experimenta sentimientos de calidez y compasión ante situaciones de malestar de otros (ítems 2-4-9-14-18-20-22); y la Angustia Personal (AP) que mide los sentimientos de miedo, aprensión e incomodidad de la persona al presenciar experiencias negativas de los demás (ítems 6-10-13-17-19-24-27).

Análisis de Datos

La validación psicométrica parte de la estructura propuesta por Davis (1980, 1983) en la versión original, con cuatro variables latentes conformadas por veintiocho ítems (Modelo original). Se proponen dos validaciones, la primera de contenido a nivel de expertos, y la segunda una validación estadística a través de un AFC. La validación de contenido tomó como base la traducción al español realizada por Pérez-Albéniz et al. (2003), junto con el juicio de expertos de la región que modificaron palabras y expresiones con el fin de mejorar la comprensión de los ítems en el contexto colombiano. Luego de ello, se realiza un AFC estimando los parámetros con el método de Mínimos Cuadrados Ponderados Diagonalizados, donde se verifica el correcto ajuste de los indicadores como: el error cuadrático de aproximación (RMSEA), el residuo estandarizado cuadrático medio (SRMR), el índice de ajuste comparativo (CFI), el índice Tucker-Lewis (TLI) y el índice de bondad de ajuste (GFI). Si se encuentra que el modelo no presenta un adecuado ajuste, se procede a un refinamiento eliminando ítems que presenten bajo aporte a su conformación factorial (TP, FS, PE y AP), $\lambda < 0.40$, tal como se propone en Pituch y Stevens (2015). Para lo anterior se hará uso del paquete Lavaan 0.6-5 del lenguaje de programación R, versión 3.6.1.

Algoritmo de estimación

Entre las alternativas a los métodos de estimación clásica como MV, se encuentran los algoritmos de estimación de MVR y Mínimos Cuadrados Ponderados (MCP) usados bajo la presencia de asimetría moderada, pero, de igual forma, creados para variables continuas (Brown, 2015). Sin embargo, tratar los datos categóricos como continuos, implica la reducción de la variabilidad de los ítems observados, dando como resultado menores correlaciones, es decir, reduciendo la relación entre las variables latentes y las variables observadas, causando una subestimación del tamaño de las cargas factoriales (Holgado-Tello et al., 2018); además conlleva a cálculos incorrectos de los estadísticos de ajuste y de los errores estándar de las cargas factoriales (Brown, 2015). Estos problemas restan precisión y validez en las estimaciones de los parámetros del modelo, lo que podría llevar a conclusiones erradas (Beauducel y Herzberg, 2006; Flora y Curran, 2004; Li, 2015, 2016).

Por esto, la relevancia de trabajos de simulación que evidencien los efectos de los diferentes métodos de estimación en AFC con ítems categóricos, tal como lo hacen Holgado-Tello, et al. (2018), quienes comparan los métodos de MV, MCP, MVR, MCNP y Mínimos Cuadrados no Ponderados Robusto (MCNPR), deduciendo que MCNPR es la mejor opción para la estimación de parámetros con ítems categóricos. Esta conclusión, dicen los autores, está en la misma línea de otras investigaciones que defienden el uso de métodos que parten de la matriz policórica (Holgado-Tello et al., 2018).

De igual forma, Li (2015) estudia la pertinencia de tratar variables categóricas como continuas en los modelos AFC. Allí se pone a prueba la estimación por MVR, que parte de una matriz de covarianza, y Mínimos Cuadrados Ponderados Diagonalizados (MCPD), originada de una matriz policórica. Dos de sus resultados son 1) sin importar el número de categorías, las cargas factoriales en MCPD tienen menor sesgo que MVR; 2) MCPD sobre estima, moderadamente, las correlaciones entre factores cuando el tamaño de la muestra es pequeño o cuando la distribución de la variable latente es moderadamente asimétrica, etc. Más adelante, el mismo Li (2016) pone a prueba el método MV, con

matriz de covarianza, MCNP y MCPD, con matriz policórica; concluyendo que los dos métodos de estimación que parte de la matriz policórica muestran mejores resultados con respecto a los estadísticos de ajuste.

Los hallazgos de Holgado-Tello et al. (2018) y Li (2015; 2016) motivan a validar el constructo de empatía de Davis (1980) por medio de MCPD, partiendo de una matriz policórica, con lo que se espera obtener un menor sesgo en los parámetros y bajos errores tipo I.

Indicadores de ajuste

Para la evaluación del ajuste del modelo son considerados distintos indicadores, entre ellos la prueba chi-cuadrado (χ^2), en donde la hipótesis nula indica que la matriz de covarianzas o correlaciones estimada es igual a la observada. En este caso rechazar la hipótesis nula revela un mal ajuste del modelo. Sin embargo, se debe considerar que este indicador se basa en un estimador que asume normalidad en los ítems y es sensible al tamaño de la muestra, a medida que se incrementan las observaciones, aumenta la probabilidad de rechazar la hipótesis nula (Brown, 2015; Pituch y Stevens, 2015). Estas falencias en el estadístico chi-cuadrado, llevan a considerar el χ^2 escaldado una propuesta para variables categóricas o variables que no siguen la distribución normal, cuya interpretación es similar a la anterior.

Otros índices de ajuste alternativos que se toman para este estudio son, el RMSEA, que refleja un buen ajuste si es menor o próximo al 0.05 y un ajuste aceptable si se encuentra en el rango de 0.05 a 0.10 (Bentler, 1995), esta misma interpretación aplica para el SRMR; por otra parte, el CFI, el TLI y el GFI pueden oscilar entre cero y uno, donde valores cercanos a uno indican un buen ajuste, siendo 0.9 el valor mínimo requerido para aceptar el modelo (Bentler, 1990; Bentler y Bonett, 1980; Brown, 2015). Se incluye también el índice de validación cruzada esperada (ECVI), para comparar los modelos, el modelo con menor valor del índice indica un mejor ajuste (Schermelleh-Engel et al., 2003).

Coefficientes de fiabilidad

Para el análisis de la fiabilidad del cuestionario se incluye la estimación del alfa (α) de Cronbach de las cuatro escalas que componen el instrumento con el fin de compararlo con los resultados obtenidos por Davis en su estudio original. Sin embargo, se reconoce los limitantes de este coeficiente con variables categóricas, por lo cual se incluyen el coeficiente Omega de McDonald (ω) y el Composite Reliability (C.R), estos coeficientes pueden oscilar entre 0 y 1, y cuanto más próximos estén a 1 más consistente será la estructura.

Resultados

El primer resultado en la revisión de las propiedades psicométricas del instrumento es una mejor adaptación de cada uno de los ítems al contexto colombiano (Tabla 3), con base en la validación de contenido a partir de expertos.

Tabla 1.
Adaptación Ítems del cuestionario al contexto colombiano

Reescritura del ítem	M1					M2						
	M	SD	Sk	K	90%	M	SD	Sk	K	90%		
1 Sueño y fantaseo, con cierta regularidad, acerca de las cosas que me podrían suceder.	2.58	1.10	-0.45	2.55	2.49	2.67	2.68	1.09	-0.43	2.41	2.63	2.72
2 A menudo tengo sentimientos tiernos y de preocupación hacia	2.63	1.00	-0.51	2.84	2.55	2.72	2.61	1.01	-0.40	2.59	2.56	2.65

	la gente menos afortunada que yo.												
(r)3	A menudo encuentro difícil ver las cosas desde el punto de vista de otra persona.	2.78	0.93	-0.56	2.93	2.70	2.85	1.20	1.02	0.68	2.97	1.15	1.24
(r)4	A veces no me siento muy preocupado por otras personas cuando tienen problemas.	2.69	1.10	-0.64	2.72	2.60	2.77	1.25	1.08	0.62	2.65	1.21	1.30
5	Verdaderamente me identifico con los sentimientos de los personajes de una novela.	1.54	1.28	0.37	2.06	1.44	1.65	1.65	1.29	0.28	2.00	1.59	1.71
6	En situaciones de emergencia me siento aprensivo e incómodo.	1.59	1.16	0.31	2.31	1.50	1.69	1.54	1.15	0.36	2.32	1.48	1.59
(r)7	Soy normalmente objetivo cuando veo una película u obra de teatro y no me involucro completamente.	2.31	1.09	-0.25	2.34	2.23	2.40	1.76	1.13	0.18	2.32	1.71	1.81
8	Intento tener en cuenta cada una de las partes (opiniones) en un conflicto antes de tomar una decisión.	3.05	0.82	-0.55	2.84	2.99	3.12	3.13	0.85	-0.74	3.02	3.09	3.17
9	Cuando veo que a alguien se le toma el pelo tiendo a protegerlo.	2.32	1.04	-0.27	2.57	2.24	2.41	2.18	1.04	-0.11	2.50	2.13	2.23
10	Normalmente siento desesperanza cuando estoy en medio de una situación muy emotiva.	1.54	1.07	0.33	2.49	1.45	1.62	1.52	1.13	0.35	2.35	1.47	1.57
11	A menudo intento comprender mejor a mis amigos imaginándome cómo ven ellos las cosas (poniéndome en su lugar).	2.75	0.95	-0.48	2.81	2.67	2.82	2.78	0.99	-0.59	2.83	2.74	2.82
(r)12	Resulta raro para mí implicarme completamente en un buen libro o película.	2.92	1.06	-0.84	3.09	2.84	3.01	1.13	1.10	0.78	0.78	1.08	1.18
(r)13	Cuando veo a alguien herido tiendo a permanecer calmado.	1.81	1.12	0.09	2.34	1.72	1.90	2.35	1.16	0.78	0.78	2.30	2.40
(r)14	Las desgracias de otros normalmente no me molestan mucho.	2.69	1.08	-0.58	2.65	2.60	2.78	1.27	1.05	0.58	0.58	1.22	1.31
(r)15	Si estoy seguro de que tengo la razón en algo no pierdo tiempo escuchando los argumentos de los demás.	2.77	1.20	-0.83	2.86	2.67	2.87	1.09	1.12	0.58	0.58	1.04	1.14
16	Después de ver una obra de teatro o cine me he sentido como si fuera uno de los personajes.	1.90	1.26	0.06	2.03	1.79	2.00	1.88	1.29	0.07	1.94	1.83	1.94
17	Cuando estoy en una situación emocionalmente tensa me asusto.	1.74	1.14	0.20	2.25	1.65	1.84	1.77	1.18	0.26	2.17	1.71	1.82

(r)18	Cuando veo a alguien que está siendo tratado injustamente a veces no siento ninguna compasión por él.	3.39	0.91	-1.61	5.16	3.32	3.47	0.64	0.95	1.63	5.24	0.60	0.68
(r)19	Normalmente soy bastante eficaz al ocuparme de emergencias.	1.84	0.95	-0.08	2.74	1.76	1.92	2.21	1.03	-0.10	2.53	2.16	2.25
20	A menudo estoy bastante afectado emocionalmente por cosas que veo que ocurren.	1.79	1.20	0.19	2.13	1.69	1.88	1.92	1.21	0.12	2.09	1.86	1.97
21	Pienso que hay dos partes para cada cuestión e intento tener en cuenta ambas partes.	2.77	0.95	-0.48	2.96	2.70	2.85	2.91	0.93	-0.65	3.13	2.87	2.95
22	Me describiría como una persona bastante sensible.	2.28	1.23	-0.15	2.04	2.18	2.38	2.37	1.22	-0.19	2.03	2.32	2.43
23	Cuando veo una buena película puedo muy fácilmente situarme en el lugar del protagonista.	2.15	1.20	-0.19	2.19	2.05	2.25	2.27	1.18	-0.20	2.18	2.22	2.32
24	Tiendo a perder el control durante las emergencias.	1.03	0.94	0.64	2.83	0.95	1.10	1.01	0.98	0.90	3.44	0.97	1.06
25	Cuando estoy disgustado con alguien normalmente intento ponerme en su lugar por un momento.	1.93	1.05	0.09	2.40	1.84	2.01	2.01	1.12	-0.06	2.28	1.96	2.06
26	Cuando estoy leyendo una historia interesante o una novela imagino cómo me sentiría si los acontecimientos de la historia me sucedieran a mí.	2.49	1.17	-0.43	2.38	2.39	2.58	2.58	1.18	-0.49	2.32	2.53	2.63
27	Cuando veo a alguien que necesita urgentemente ayuda en una emergencia me derrumbo.	0.81	0.86	0.73	2.60	0.74	0.88	0.80	0.90	1.03	3.63	0.76	0.84
28	Antes de criticar a alguien intento imaginar cómo me sentiría si estuviera en su lugar.	2.58	1.08	-0.37	2.48	2.49	2.67	2.54	1.09	-0.38	2.47	2.49	2.59

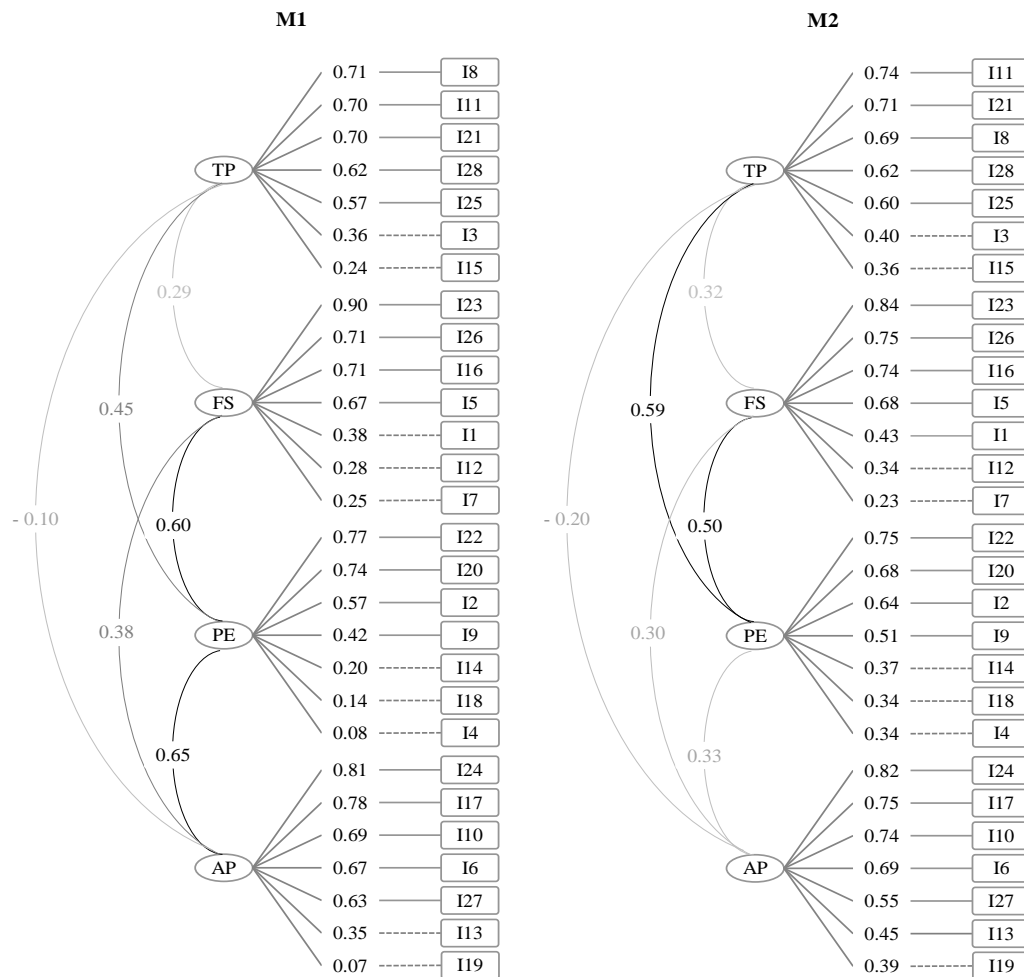
* Las afirmaciones marcadas con una (r) deben tabularse de manera inversa (0 = 4; 1 = 3; 2 = 2; 3 = 1; 4 = 0), eso quiere decir que si el participante marca 0; al momento de tabular se debe escribir 4.

En cuanto a la validación psicométrica, se parte de la aplicación del AFC por medio de MCPD a la estructura original propuesta por Davis (cuatro variables latentes conformadas por veintiocho ítems). Sin embargo, de acuerdo con los indicadores de ajuste (Tabla 4), se acepta la H1, pues el modelo en su forma original no converge adecuadamente en ninguna de las dos muestras: RMSEA y SRMR > 0.10 y CFI, TLI, GFI < 0.90. En las cargas factoriales del modelo original para las dos muestras (Figura 1) se encuentra ítems con un bajo aporte a su conformación factorial. Estos ítems se repiten en su mayoría en las dos muestras. En la variable TP los ítems con menor carga son el 3 y el 15; en FS los ítems 12 y 7, en esta variable el ítem 1 presenta una carga baja de 0.38 en la M1 y una carga moderada de 0.43 en la M2; en la variable PE los ítems son el 14, 18 y 4; y en la variable AP se presentan con bajas cargas, los ítems 13 y 19 en la M1 y en la M2 únicamente el ítem 19. Se evidencia que los ítems cuestionables, a excepción del ítem 1, se caracterizan por ser aquellos que, tal como fueron señalados en la Tabla 3, deben puntuarse de manera inversa.

A partir de la revisión de las cargas de la estructura factorial del modelo original, se especifican tres nuevas estructuras: Propuesta 1, eliminar los ítems que se encuentran por debajo del umbral en la M1 del modelo original (1, 3, 4, 7, 12, 13, 14, 15, 18, 19), quedando en total 18 ítems. Propuesta 2, extraer aquellos ítems que presentan inconvenientes en la M2 del modelo original (3, 4, 7, 12, 14, 15, 18, 19) quedando en total 20 ítems. Propuesta 3, retirar los ítems que puntúan de manera inversa (3, 4, 7, 12, 13, 14, 15, 18, 19) los cuales en su mayoría tienen bajas cargas factoriales en ambas muestras quedando en total 19 ítems.

Figura 1.

Estructura factorial Modelo Original



Nota: La intensidad de las líneas indican el nivel de correlación de los factores. Líneas oscuras indican alta correlación.

Los indicadores de ajuste (Tabla 4) evidencian que la propuesta 1 y 3 son las que presentan mejores resultados. Sin embargo, se opta por la propuesta 3 al evidenciarse que todas sus cargas factoriales en M1 y M2 están por encima de 0.40 (Figura 2) contrario a la propuesta 2. La propuesta 3 presenta indicadores de ajuste más favorables respecto al modelo original, un RMSEA y SRMR < 0.10 y CFI, TLI, GFI > 0.90; además, logra reducir el ECVI en ambas muestras.

Tabla 2.*Indicadores de ajuste de los modelos*

	Modelo Original		Propuesta 1		Propuesta 2		Propuesta 3	
	M1	M2	M1	M2	M1	M2	M1	M2
χ^2	2842.9	7741.8	441	1505.1	549	1866.8	477.3	1590.9
gl	344	344	129	129	164	164	146	146
p-v	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
χ^2 escalado	2118.2	6272.9	492.4	1679.4	576.9	1963.5	520.1	1751.9
p-v.	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
RMSEA	0.133	0.126	0.077	0.088	0.076	0.087	0.074	0.085
SRMR	0.125	0.115	0.075	0.075	0.076	0.077	0.074	0.074
CFI	0.782	0.806	0.963	0.95	0.957	0.942	0.962	0.95
TLI	0.761	0.787	0.957	0.941	0.95	0.933	0.956	0.941
GFI	0.894	0.912	0.976	0.974	0.972	0.971	0.975	0.974
ECVI	7.604	5.886	1.532	1.243	1.843	1.523	1.644	1.313

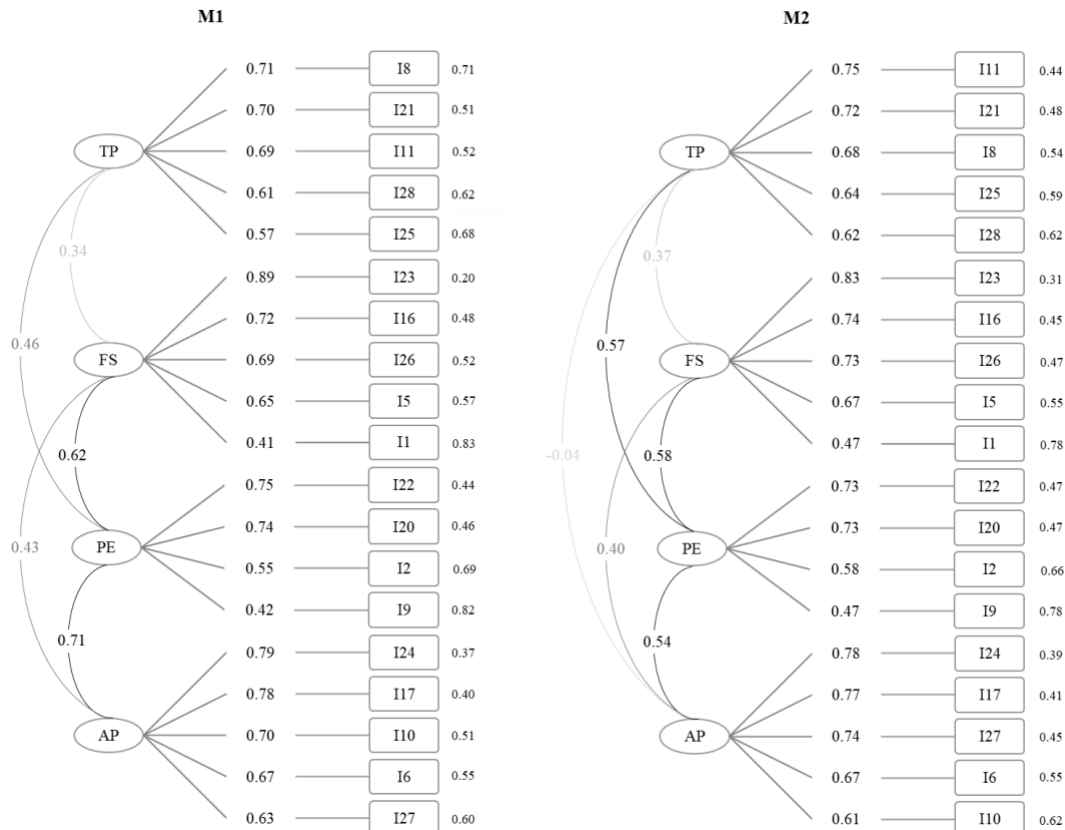
La fiabilidad de cada una de las variables latentes es mostrada en la Tabla 5, comparando los resultados con la validación original de Davis (1980), el modelo original y la propuesta 3. Pese a que Davis solo usó el α de Cronbach para consistencia interna, se incluyeron el ω de MacDonal y C.R como indicadores más apropiados para variables categóricas. Los modelos propuestos presentan resultados satisfactorios y cercanos a los obtenidos por el creador del instrumento, en especial los de la propuesta 3.

Tabla 3.*Coefficientes de fiabilidad por escalas*

	Validación Original Davis (1980)		Modelo con la estructura Original		Propuesta 3		
	Mujeres	Hombres	M1	M2	M1	M2	
TP	α	0.78	0.75	0.70	0.74	0.74	0.76
	ω			0.77	0.82	0.83	0.80
	C.R			0.77	0.79	0.79	0.81
FS	α	0.75	0.78	0.73	0.73	0.77	0.78
	ω			0.80	0.80	0.77	0.80
	C.R			0.78	0.78	0.81	0.82
PE	α	0.70	0.72	0.62	0.68	0.66	0.68
	ω			0.72	0.73	0.72	0.73
	C.R			0.69	0.75	0.71	0.72
AP	α	0.78	0.78	0.74	0.77	0.79	0.79
	ω			0.83	0.83	0.83	0.84
	C.R			0.80	0.82	0.84	0.84

Las relaciones entre los factores del modelo final, propuesta 3, oscilaron en M1 entre 0.34 a 0.71 y en la M2 entre -0.04 a 0.58. En ambas muestras FS-PE y PE-AP, son las que presentan una mayor correlación. La relación entre fantasía y preocupación empática fueron también las más fuertes en la validación original (Davis, 1980, 1983), Mestre et al. (2004) y en Lucas-Molina et al (2017). Por el contrario, las escalas TP y AP mostraron una relación prácticamente nula, en la M1 y no significativa en la M2.

Figura 2.
Estructura factorial propuesta



Nota: La intensidad de las líneas indican el nivel de correlación de los factores. Líneas oscuras indican alta correlación.

Finalmente, se comprueba la invarianza de estructura factorial de la propuesta 3 por sexo. La revisión de la invarianza factorial se evaluó progresivamente: configuracional, métrica, fuerte y estricta. Los resultados (Tabla 6) demuestran que el modelo final presenta buenos indicadores de ajuste, sin sobrepasar los valores de $\Delta RMSEA$ y ΔCFI el 0.01, lográndose una invarianza factorial en sentido estricto en las dos muestras. Por lo tanto, a partir de los resultados en la Tabla 5 y Tabla 6 se rechaza la H2, al comprobarse que el modelo ajusta de forma adecuada en los estudiantes admitidos en el primer (M1) y segundo (M2) periodo académico del año.

Tabla 4.
Índices de ajuste para Invarianza Factorial

	Muestra 1						
	<i>RMSEA</i>	<i>SRMR</i>	<i>CFI</i>	<i>TLI</i>	<i>GFI</i>	Δ <i>RMSEA</i>	Δ <i>CFI</i>
<i>Configuracional</i>	0.040	0.071	0.974	0.970	0.995		
<i>Métrica</i>	0.046	0.075	0.964	0.959	0.995	0.006	-0.011
<i>Escalar</i>	0.048	0.078	0.958	0.956	0.994	0.002	-0.005
<i>Estricto</i>	0.049	0.082	0.954	0.954	0.994	0.001	-0.004
	Muestra 2						
	<i>RMSEA</i>	<i>SRMR</i>	<i>CFI</i>	<i>TLI</i>	<i>GFI</i>	Δ <i>RMSEA</i>	Δ <i>CFI</i>
<i>Configuracional</i>	0.061	0.065	0.940	0.930	0.996		
<i>Métrica</i>	0.060	0.065	0.939	0.932	0.996	-0.001	-0.001
<i>Escalar</i>	0.060	0.066	0.938	0.933	0.996	0.000	-0.001
<i>Estricto</i>	0.059	0.068	0.937	0.935	0.996	-0.001	-0.001

Discusión

Este artículo proporciona una nueva revisión de las propiedades psicométricas del constructo IRI aplicado a estudiantes recién admitidos en el primer y segundo periodo de ingreso de la Universidad Industrial de Santander. El análisis de su dimensionalidad revela una estabilidad factorial compatible con la propuesta original de Davis (1980; 1983), por lo tanto, se le considera un buen instrumento para medir las dimensiones de la empatía para ambas muestras, evidenciado que el semestre en que se fue admitido no afecta la validación del instrumento pese a que las poblaciones tengan algunas diferencias históricas en los puntajes de las pruebas de ingreso y que el portafolio de programas ofertados sea distintos entre semestres. Sin embargo, la estructura original no ajusta en totalidad, nueve ítems presentan bajas cargas factoriales y coinciden ser los que se puntúan de manera inversa. Al parecer, el IRI al tener mayoritariamente ítems con puntuaciones en una escala ordenada ascendente, induce a un patrón de respuestas que, al encontrarse con ítems invertidos, se desajusta y hace que el participante pierda la referencia de la escala que había desarrollado de las sentencias no invertidas. Por ello, los ítems inversos de una misma dimensión poco aportan a la conformación de la variable latente. Por ejemplo, los ítems del factor PE, en su mayoría, son calificados de forma ascendente ordenada, donde “4” indica una mayor preocupación empática, tal como: “A menudo tengo sentimientos tiernos y de preocupación hacia la gente menos afortunada que yo”; sin embargo, en el ítem 3, inverso y también perteneciente a PE, “A veces no me siento muy preocupado por otras personas cuando tienen problemas” tiene una escala inversa donde “4” indica una menor preocupación empática.

Estos cambios en las escalas pueden provocar confusión al lector, quien, al tener un conjunto grande de enunciados, predispone sus respuestas a marcar alto o bajo, en una misma dirección de acuerdo con lo que él intuya que se mide; esta tendencia se ve perturbada por los ítems inversos, haciendo que pierdan su referencia de auto reporte. Este inconveniente se acentúa aún más, en los ítems con negaciones en donde el adjetivo “no” niega el verbo, pues estos presentan las cargas factoriales más bajas en el AFC (ítems 4, 7, 14, 15, 18), lo que reafirma que el inadecuado ajuste puede ser debido a problemas de comprensión.

Estos resultados son compatibles con la validación del instrumento con excombatientes en Colombia realizado por García-Barrera et al. (2017). En este caso, los autores argumentaron que debido al bajo nivel de educación promedio de los individuos se hacía difícil la comprensión de los ítems

inversos. Sin embargo, los resultados de la validación que se realiza en este artículo reafirman que la estructura gramatical de estos ítems, al ser traducidos al español, puede ser de difícil comprensión incluso en estudiantes universitarios.

Otro posible agravante de la versión en español es la traducción ligera de los adverbios de cantidad y de frecuencia como “muy”, “usualmente”, “a menudo”, etc., pues las traducciones pueden hacer perder la escala de la intensidad de la acción en la frase. Un adverbio como “usually” puede reflejar diferente intensidad para un hablante del inglés; en comparación con su traducción al español como “normalmente”. Esto puede provocar distorsión en el auto reporte, lo cual es un punto crucial en la validación de las traducciones de los constructos, que poco ha sido analizada.

Limitaciones y futuras líneas

Se reconoce que este artículo se limita a valorar las propiedades psicométricas. Sin embargo, no hace mayores aportes al cuerpo teórico sobre cómo medir la empatía y otros constructos creados para ello. Se recomienda realizar un estudio de validez de criterio con otro instrumento que pretenda medir la empatía.

Conclusiones

A manera de conclusión, futuras aplicaciones del IRI en estudiantes universitarios en Colombia pueden seguir el modelo final en el presente documento (propuesta 3), el cual contempla las cuatro dimensiones originales y diecinueve ítems. Este modelo presenta buenos indicadores de ajuste y mejores indicadores de fiabilidad en comparación con el original en ambas muestras, lo que implica que podría ser usado en estudiantes que ingresen al primer y segundo semestre. Esta estructura factorial comprobó ser invariante por sexo. Además, se sugiere que futuras validaciones de este instrumento usen el algoritmo de estimación sea MCPD, pues se evidencia que es el más adecuado para el tipo de ítem del constructo.

Agradecimientos

Este trabajo hace parte de los resultados de los proyectos de investigación “Formación en competencias ciudadanas en educación superior”, financiado por el Instituto Colombiano para la Evaluación de la Educación y la Universidad Industrial de Santander (8265-ICFES- UIS), y “Filosofía moral experimental para la formación de competencias ciudadanas” financiado por la Vicerrectoría de Investigación y Extensión de la Universidad Industrial de Santander (2403-VIE-UIS).

Referencias

- Bandura, A. (1991). Social Cognitive Theory of Moral Thought and Action. En Kurtines, W., Gewirtz, J., Lamb, J. L. *Handbook of Moral Behavior and Development* (pp.45–103). Lawrence Erlbaum Associates.
- Beauducel, A., y Herzberg, P. Y. (2006). On the performance of maximum likelihood versus means and variance adjusted weighted least squares estimation in CFA. *Structural Equation Modeling*, 13(2), 186–203. https://doi.org/10.1207/s15328007sem1302_2
- Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin*, 107(2), 238–246. <https://doi.org/10.1037/0033-2909.107.2.238>

- Bentler, P. M., y Bonett, D. G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, 88(3), 588–606. <https://doi.org/10.1037/0033-2909.88.3.588>
- Brown, T. A. (2015). *Confirmatory Factor Analysis for Applied Research* (T. D. Little, Ed.) (2nd ed.). The Guilford Press.
- Carrasco, M. Á., Begoña, E., Barbero, M., Holgado, F., y Del Barrio, M. (2011). Propiedades psicométricas del Interpersonal Reactivity Index (IRI) en población infantil y adolescente española. *Psicothema*, 23(4), 824–831.
- Cecconello, A., y Koller, S. (2000). Competência social e empatia: um estudo sobre resiliência com crianças em situação de pobreza. *Estudos de Psicologia*, 5(1), 71–93.
- Cliffordson, C. (2001). Parent's judgments and students' self-judgments of empathy: the structure of empathy and agreement of judgment based on the Interpersonal Reactivity Index (IRI). *European Journal of Psychological Assessment*, 17(1), 36–47. <https://doi.org/10.1027//1015-5759.17.1.36>
- Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology*, 10, (85), 1-17
- Davis, M. H. (1983). Measuring individual differences in empathy: evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, 44, 113-126. <https://doi.org/10.1037/0022-3514.44.1.113>
- De Corte, K., Buysse, A., Verhofstadt, L. L., Roeyers, H., Ponnet, K., y Davis, M. H. (2007). Measuring empathic tendencies: reliability and validity of the dutch version of the interpersonal reactivity index. *Psychologica Belgica*, 47(4), 235. <https://doi.org/10.5334/pb-47-4-235>
- Deutsch, F., y Madle, R. A. (1975). Empathy: historic and current conceptualizations, measurement, and a cognitive theoretical perspective. *Human Development*, 18(4), 267–287. <https://doi.org/10.1159/000271488>
- Eisenberg, N. (2000). Emotion, regulation, and moral development. *Annu. Rev. Psychol*, 51, 665–697.
- Eisenberg, N., Fabes, R. A., Guthrie, I. K., y Reiser, M. (2000). Dispositional emotionality and regulation: their role in predicting quality of social functioning. *Journal of Personality and Social Psychology*, 78(1), 136–157. <https://doi.org/10.1037/0022-3514.78.1.136>
- Fernández, A. M., Dufey, M., y 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>
- Flora, D. B., y Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods*, 9(4), 466–491. <https://doi.org/10.1037/1082-989X.9.4.466>
- García-Barrera, M. A., Karr, J. E., Trujillo-Orrego, N., Trujillo-Orrego, S., y 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>

- Gilar, R., Miñano, P., y Castejón, J. (2008). Inteligencia emocional y empatía: su influencia en la competencia social en educación secundaria obligatoria. *SUMMA Psicológica UST*, 5(1), 21–32.
- Gilet, A. L., Mella, N., Studer, J., Griehn, D., y Labouvie-Vief, G. (2013). Assessing dispositional empathy in Adults: A french validation of the interpersonal reactivity index (IRI). *Canadian Journal of Behavioural Science*, 45(1), 42–48. <https://doi.org/10.1037/a0030425>
- Hawk, S. T., Keijsers, L., Branje, S. J. T., Van Der Graaff, J., De Wied, M., y Meeus, W. (2013). Examining the Interpersonal Reactivity Index (IRI) among early and late adolescents and their mothers. *Journal of Personality Assessment*, 95(1), 96–106. <https://doi.org/10.1080/00223891.2012.696080>
- Hojat, M., Mangione, S., Kane, G. C., y Gonnella, J. S. (2005). Relationships between scores of the Jefferson Scale of Physician Empathy (JSPE) and the Interpersonal Reactivity Index (IRI). *Medical Teacher*, 27(7), 625–628. <https://doi.org/10.1080/01421590500069744>
- Holgado-Tello, F. P., Morata-Ramírez, M. Á., y Barbero García, M. I. (2018). Confirmatory factor analysis of ordinal variables: A simulation study comparing the main estimation methods. *Avances En Psicología Latinoamericana*, 36(3), 601–618. <https://doi.org/10.12804/revistas.urosario.edu.co/apl/a.4932>
- Howes, C., Matheson, C. C., y Hamilton, C. E. (1994). Maternal, teacher, and child care history correlates of children's relationships with peers. *Child Development*, 65(1), 264–273. <https://doi.org/10.1111/j.1467-8624.1994.tb00749.x>
- Ingoglia, S., Lo Coco, A., y Albiero, P. (2016). Development of a brief form of the Interpersonal Reactivity Index (B-IRI). *Journal of Personality Assessment*, 98(5), 461–471. <https://doi.org/10.1080/00223891.2016.1149858>
- Junttila, N., Voeten, M., Kaukiainen, A., y Vauras, M. (2006). Multisource assessment of children's social competence. *Educational and Psychological Measurement*, 66(5), 874–895. <https://doi.org/10.1177/0013164405285546>
- Kang, I., Kee, S.-W., Kim, S.-E., Jeong, B.-S., Hwang, J.-H., Song, J.-E., y Kim, J.-W. (2009). Reliability and validity of the korean-version of Interpersonal Reactivity Index. *Journal Korean Neuropsychiatr Association*, 48(5), 352–358.
- Kerem, E., Fishman, N., y Josselson, R. (2001). The experience of empathy in everyday relationships: Cognitive and affective elements. *Journal of Social and Personal Relationships*, 18(5), 709–729. <https://doi.org/10.1177/0265407501185008>
- Konrath, S. (2014). The empathy paradox: Increasing disconnection in the age of increasing connection. In *Digital Arts and Entertainment: Concepts, Methodologies, Tools, and Applications*, 2, 967–991. <https://doi.org/10.4018/978-1-4666-6114-1.ch047>
- Li, CH. (2015). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavior Research Methods*, 48(3), 936–949. <https://doi.org/10.3758/s13428-015-0619-7>
- Li, CH. (2016). The performance of ML, DWLS, and ULS estimation with robust corrections in structural equation models with ordinal variables. *Psychological Methods*, 21(3), 369–387. <https://doi.org/10.1037/met0000093>

- Lucas-Molina, B., Pérez-Albéniz, A., Ortuño-Sierra, J., y Fonseca-Pedrero, E. (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>
- Mestre, V., Frías, M. D., y Samper, P. (2004). La medida de la empatía: análisis del Interpersonal Reactivity Index. *Psicothema*, 16(2), 255–260.
- Mestre, V., Samper, P., y Frías, M. D. (2002). Procesos cognitivos y emocionales predictores de la conducta prosocial y agresiva: la empatía como factor modulador. *Psicothema*, 14(2), 227–232.
- Oceja, L., López-Pérez, B., Ambrona, T., y Fernández, I. (2009). Measuring general dispositions to feeling empathy and distress. *Psicothema*, 21(2), 171–176.
- Paulus, C. (2009). Der Saarbrücker Persönlichkeitsfragebogen SPF (IRI) zur Messung von Empathie. Psychometrische Evaluation der deutschen Version des Interpersonal Reactivity Index.
- Pérez-Albéniz, A., Paúl, J. de, Etxebarria, J., Montes, M. P., y Torres, E. (2003). Adaptación de Interpersonal Reactivity Index (IRI) al español. *Psicothema*, 15(2), 267–272.
- Pituch, K. A., y Stevens, J. (2015). Applied multivariate statistics for the social sciences (6th ed.). Routledge. <https://doi.org/10.4324/9781315814919>
- Pulos, S., Elison, J., y Lennon, R. (2004). The hierarchical structure of the Interpersonal Reactivity Index. *Social Behavior and Personality*, 32(4), 355–360. <https://doi.org/10.2224/sbp.2004.32.4.355>
- Sampaio, L. R., Guimarães, P., Camino, C., Formiga, N., y Menezes, I. (2011). Estudos sobre a dimensionalidade da empatia tradução e adaptação do Interpersonal Reactivity Index (IRI). *Psico*, 42(1), 67–76.
- Schermelleh-Engel, K., Moosbrugger, H., y Müller, H. (2003). Evaluating the fit of structural equation models: Test of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online*, 8(2), 23–74.
- Siu, A. M. H., y Shek, D. T. L. (2005). Validation of the interpersonal reactivity index in a Chinese context. *Research on Social Work Practice*, 15(2), 118–126. <https://doi.org/10.1177/1049731504270384>
- Vizoso-Gómez, C. M. (2019). Empatía en futuros maestros: diferencias por género. *International Journal of Developmental and Educational Psychology. Revista INFAD de Psicología*, 5(1), 541. <https://doi.org/10.17060/ijodaep.2019.n1.v5.1636>
- Yarnold, P. R., Bryant, F. B., Nightingale, S. D., y Martin, G. J. (1996). Assessing physician empathy using the Interpersonal Reactivity Index: A measurement model and cross-sectional analysis. *Psychology, Health and Medicine*, 1(2), 207–221. <https://doi.org/10.1080/13548509608400019>