EVALUAR

Revista Evaluar

Laboratorio de Evaluación Psicológica y Educativa

Facultad de Psicología de la Universidad Nacional de Córdoba

2024
VOL 24 - N°1
ISSN 1667-4545

2024, Vol. 24, No. 1
ISSN 1667-4545
Recuperado de https://revistas.unc.edu.ar/index.php/revaluar

Laboratorio de Evaluación Psicológica y Educativa
Facultad de Psicología - Universidad Nacional de Córdoba

Datos normativos para la Escala de Dificultades en la Regulación Emocional (DERS) en población de estudiantes universitarios de Río Cuarto

Normative data for the Difficulties in Emotional Regulation Scale (DERS) in a population of college students from Río Cuarto (Argentina)

Mariano Daniel Villarrubia * [1] , Valeria Estefanía Morán [1] , [2] , María Zoé Natera [1] , Vanessa Giovannini [1] , Valentín Vieyra [1] , Claribel Abbá [1]

1- Universidad Siglo 21, Córdoba, Argentina.
2 - Instituto de Investigaciones Sociales, Territoriales y Educativas (ISTE) CONICET-Universidad Nacional de Río Cuarto, Córdoba, Argentina.

Recibido: 05/10/2023 **Revisado:** 01/09/2023 **Aceptado:** 04/12/2023

Resumen

Se ha demostrado en estudios previos que las dificultades en la regulación emocional se asocian con un aumento del riesgo de problemas de salud mental en estudiantes universitarios. Por ello, en Argentina se ha adaptado la Escala de Dificultades en la Regulación Emocional (DERS, por sus siglas en inglés) específicamente para esta población, sin embargo, aún no existen datos normativos que permitan utilizarla en entornos clínicos o para propósitos de clasificación. El objetivo principal de este estudio fue obtener baremos para la interpretación de la DERS con una muestra de 403 estudiantes universitarios de Río Cuarto mediante percentiles y puntuaciones T. Los puntajes demostraron una alta fiabilidad, lo que indica que son precisos para clasificar a los individuos. Los datos normativos proporcionados permiten utilizar la escala en evaluaciones de screening, en combinación con otras fuentes de información.

Palabras clave: *regulación emocional, dificultades en la regulación emocional, DERS, datos normativos, estudiantes universitarios*

Abstract

Previous studies show that difficulties in emotional regulation are associated with an increased risk of mental health problems in college students. Therefore, in Argentina, the Difficulties in Emotional Regulation Scale (DERS) has been adapted specifically for this population; however, there is still no normative data that allows its use in clinical settings or for classification purposes. The main objective of this study was to obtain scales for the interpretation of the DERS with a sample of 403 university students from Río Cuarto, using percentiles and T-scores. The scores demonstrated high reliability, indicating that they are accurate for classifying individuals. The normative data provided allows the scale to be used in screening evaluations in combination with other sources of information.

Keywords: *emotional regulation, emotional regulation difficulties, DERS, normative data, college students*

***Correspondencia a:** Mariano Daniel Villarrubia, Dirección: Moreno 320, Río Cuarto, Córdoba, Argentina.
E-mail: villarrubiamariano@gmail.com. Teléfono: 358 4177856

Cómo citar este artículo: Villarrubia, D., A., Morán, E., V., Natera Z., M., Giovannini, V., Vieyra, V. & Abbá, C. (2024): Datos normativos para la Escala de Dificultades en la Regulación Emocional (DERS) en población de estudiantes universitarios de Río Cuarto. *Revista Evaluar, 24*(1) 1-13. Recuperado de https://revistas.unc.edu.ar/index.php/revaluar

Participaron en la edición de este artículo: Carla Marianela Lopez Ghisolfi, Andrea Suárez, Eugenia Barrionuevo, Florencia Ruiz, Jorge Bruera.

Introducción

En la actualidad, la Regulación Emocional (RE) se ha posicionado como una temática de creciente interés en diferentes áreas de investigación, dada la evidencia empírica que señala su influencia en el surgimiento y mantenimiento de diversas patologías (Gross & Jazaieri, 2014). No obstante, existen múltiples definiciones que refieren a este constructo sin consenso acerca de las características distintivas del mismo (Bolgeri-Escorza et al., 2020; Gratz & Roemer, 2004; Guzmán-González et al., 2014; Hervás & Jódar, 2008).

La RE se refiere a una serie de procesos complejos que engloban comportamientos, respuestas fisiológicas, aspectos cognitivos y estados emocionales que las personas emplean para influir en sus emociones y determinan cuándo las experimentan y cómo las expresan, con el propósito de alcanzar un equilibrio emocional (Koole, 2009). De acuerdo con Gratz y Roemer (2004), la RE se presenta como un concepto multidimensional que facilita el logro de metas relacionadas con el bienestar individual y la adquisición de estrategias funcionales. En otras palabras, cuando se experimentan emociones, se desencadenan cambios en diferentes niveles, como el fisiológico, cognitivo, emocional y motor, que activan procesos internos y externos permitiendo a la persona supervisar, orientar y gestionar sus reacciones emocionales con el fin de actuar de manera adaptativa para alcanzar metas específicas (LeDoux & Brown, 2017; Vargas-Gutiérrez & Muñoz-Martínez, 2013). En consonancia con esta perspectiva, Bolgeri-Escorza et al. (2020) conciben la RE como una habilidad que involucra la conciencia, comprensión y aceptación de las emociones, así como la capacidad de emplear estrategias adecuadas según la situación para modular las respuestas emocionales y adaptarse al entorno.

Además, es importante destacar que la capacidad de un individuo para regular sus propias emociones está estrechamente relacionada con su habilidad para interactuar eficazmente con los demás, ya que depende de su competencia para ajustar sus juicios y su comportamiento de acuerdo con el contexto social presente. Esto implica un constante proceso de evaluación y reevaluación del significado de estímulos socioafectivos tanto externos como internos, las conexiones entre estos estímulos o acciones y los resultados emocionales, y, por último, la capacidad de tomar decisiones equilibradas que consideren tanto las recompensas a corto plazo como las de a largo plazo (Ochsner & Gross, 2008). Este proceso es controlado y demanda esfuerzo, require respuestas basadas en pensamientos y comportamientos conscientes y deliberados (Compas et al., 2017) y abarca una amplia variedad de aspectos, como la emoción, la cognición, el comportamiento y el contexto.

La Desregulación Emocional (DE) refiere a un conjunto de formas poco adaptativas de respuesta hacia las emociones, tales como dificultades para controlar comportamientos en contextos de emociones intensas, ausencia de disposición a la realización de actividades que sean significativas en contextos de angustia y uso inflexible de estrategias adaptativas involucradas en la modulación de la duración o intensidad de experiencias emocionales (Gratz & Roemer, 2004). Es decir, supone un proceso regulatorio disfuncional en el cual se desarrollan respuestas emocionales inadecuadas para realizar actividades, comunicarse e influir en otros. Sin embargo, además de las dificultades en la supresión de estados afectivos considerados negativos, la DE supone una habilidad deficiente para experimentar y diferenciar estados emocionales, por lo tanto, se dificulta la claridad, aceptación y comprensión de ellos en función de los objetivos propuestos en un contexto determinado (de la Rosa-Gómez et al., 2021).

Se ha evidenciado en estudios empíricos recientes que las dificultades en la regulación emo-

cional constituyen un asunto esencial relacionado con el desarrollo y el mantenimiento de múltiples formas de patología psíquica y problemas emocionales, tales como depresión, trastornos de personalidad, trastornos de ansiedad, trastorno obsesivo-compulsivo, síntomas de estrés postraumático, autolesiones (Guzmán-González et al., 2020), consumo problemático de sustancias (Simons et al., 2017; Vilhena-Churchill & Goldstein, 2014; Weiss et al., 2017) así como también, trastornos en las conductas alimentarias, por ejemplo, la ingesta de comida por atracones (Eichen et al., 2017) o la restricción alimenticia (Haynos et al., 2018). Asimismo, las dificultades en la regulación emocional han sido correlacionadas con diversos trastornos de índole física como hipertensión arterial, problemas cardíacos y cáncer (de la Rosa-Gómez et al., 2021; del Valle et al., 2020).

En general, la población de estudiantes universitarios se encuentra transitando el pasaje hacia la adultez propiamente dicha (considerando que la mayoría está en un rango etario de 18 a 25 años). En esta etapa del desarrollo, se experimentan múltiples cambios a nivel físico, cognitivo y social (Arnett et al., 2014), se produce un mayor desarrollo de las funciones cerebrales implicadas en la RE y en la toma de decisiones, se elabora una reconfiguración de los vínculos, a la vez que se participa en nuevas interacciones sociales y se asumen responsabilidades académicas más exigentes o nuevos proyectos laborales (Michelini & Godoy, 2022; Moretti & Medrano, 2014).

En este marco, la RE cobra relevancia porque, si la persona es capaz de regular sus emociones de manera adaptativa, posiblemente tenga un afrontamiento eficaz de la vida académica y aumente su bienestar psicológico (del Valle et al., 2020; Bolgeri-Escorza et al., 2020). Desde esta perspectiva, un estudio realizado por Cabanach et al. (2018) presentó correlaciones positivas y estadísticamente significativas entre RE y el uso de estrategias de afrontamiento activo, lo que permitió concluir que aquella constituye un componente relevante del afrontamiento eficaz de la vida académica que favorece el bienestar psicológico de los estudiantes universitarios.

Gratz y Roemer (2004) desarrollaron la Escala de Dificultades en Regulación Emocional (DERS, por sus siglas en inglés), cuyo principal objetivo es evaluar la capacidad de RE percibida a nivel de rasgo e identificar dificultades clínicamente relevantes (de la Rosa-Gómez et al., 2021). El cuestionario evalúa las dificultades en la regulación emocional y consiste en una medida de autoinforme conformado por 36 ítems que indagan seis factores: la no aceptación de respuestas emocionales, la dificultad conductual dirigida a metas, la dificultad en el control de impulsos, el acceso limitado a estrategias efectivas de regulación emocional y la falta de claridad emocional. Este instrumento ha evidenciado propiedades psicométricas óptimas en muestras de diferentes edades, como la alta consistencia interna ($\alpha = .93$) (de la Rosa-Gómez et al., 2021), una buena fiabilidad test-retest dentro de un periodo de entre cuatro y ocho semanas ($p = .88$; $p < .01$) y una adecuada validez predictiva y de criterio (Hervás & Jódar, 2008; Marín-Tejeda et al., 2012).

Además, la DERS ha sido adaptada, traducida y validada en diferentes países. En Argentina se realizaron dos adaptaciones en población de estudiantes universitarios, la primera, elaborada por Medrano y Trógolo (2014), consta de 28 ítems agrupados en seis factores: impulsos, estrategias, aceptación, metas, conciencia y claridad. Esta adaptación obtuvo valores óptimos y aceptables de confiabilidad por medio del análisis de consistencia interna. Sin embargo, para las diferentes subescalas, en general, los valores fueron inferiores a los señalados en estudios previos. Según los autores, una posible razón de esta baja confiabilidad sería la reducción significativa en el número

de ítems, por lo que recomiendan efectuar estudios de análisis factorial confirmatorio con el fin de determinar con mayor certeza la cantidad de dimensiones subyacentes y de ítems para cada dimensión. Asimismo, postulan que se requeriría contar con más estudios psicométricos que examinen la estabilidad de la escala y su poder predictivo.

La Segunda fue realizada en la provincia de Córdoba por Michelini y Godoy (2022). Esta versión conservó los 36 ítems de la escala original y el formato de respuesta Likert de cinco puntos (*casi nunca* a *casi siempre*). Los ítems se agrupan en seis dimensiones de regulación emocional: aceptación (falta de aceptación emocional), claridad (falta de claridad emocional), conciencia (falta de conciencia emocional), estrategias (acceso limitado a estrategias efectivas de regulación emocional), impulsos (dificultades en el control de impulsos) y, por último, metas (interferencia en conductas dirigidas a metas frente a un estado emocional alterado). Esta adaptación reportó adecuados valores de consistencia interna para las subescalas (valores α de Cronbach entre .74 y .89) y para la escala total (α = .93).

A pesar de que el instrumento se ha adaptado para estudiantes universitarios en Córdoba y de que el interés en investigar la RE y la DE ha aumentado en los últimos años, debido a su relevancia para la salud mental y el bienestar psicológico, no se han desarrollado hasta el momento baremos que posibiliten la interpretación normativa de la escala. Como afirmaron Tornimbeni et al. (2008), la creación de estos posibilita que, en el contexto de la evaluación psicométrica, cada persona pueda ser comparada con su verdadero grupo de referencia, lo cual favorece una adecuada interpretación según las características sociodemográficas de la población y otras variables relacionadas con lo que el test evalúa. En este marco el presente trabajo se propuso elaborar valores normativos para la DERS en la población estudiantil de Río Cuarto, lo que permitirá una evaluación más precisa y contextualizada de las dificultades en la regulación emocional en este contexto específico.

Método

Participantes

La muestra se conformó por 403 estudiantes universitarios de Río Cuarto, de los cuales el 74.7% se identificaron como mujeres y el 25.3% como varones. La edad varió entre los 18 y 52 años (M = 22,5; DE = 4,51). Aproximadamente, la mitad de los participantes estudiaba Licenciatura en Psicología (46%), seguida de Contador Público (11.4%); en total, se representaron 51 carreras de nivel superior. Se utilizó un muestreo no probabilístico accidental (León & Montero, 2002) y el tamaño de la muestra se estableció con un error muestral del 5% para poblaciones infinitas, según la fórmula de Arkin y Colton (Sierra-Bravo, 2007).

Instrumentos

Se aplicó la Escala de Dificultades en la Regulación Emocional (DERS; Gratz & Roemer, 2004). El cuestionario evalúa las dificultades en la regulación emocional y consiste en una medida de autoinforme conformado por 36 ítems que indagan seis factores: la no aceptación de respuestas emocionales, la dificultad conductual dirigida a metas, la dificultad en el control de impulsos, el acceso limitado a estrategias efectivas de regulación emocional y la falta de claridad emocional (de la Rosa-Gómez et al., 2021). Para responder, los participantes deben indicar con qué frecuencia experimentan las dificultades descritas en cada ítem en una escala Likert de cinco puntos (1 = *casi nunca* a 5 = *casi siempre*). Para obtener el

puntaje total de la escala se suman todos los ítems y para el puntaje por subescala se suman los ítems de cada factor.

La escala ha evidenciado una alta consistencia interna (α = .93 para la escala total y entre .80 y .89 para las subescalas) y una buena confiabilidad test-retest (p =.88, p <.01) (Gratz & Roemer, 2004; Hervás & Jódar, 2008). Asimismo, mostró evidencia de validez basada en la estructura interna y una adecuada validez predictiva y de criterio (Hervás & Jódar, 2008; Marín-Tejeda et al., 2012).

En este estudio se utilizó la adaptación de Michelini y Godoy (2022) para adultos emergentes universitarios de la ciudad de Córdoba, Argentina. La misma mostró valores adecuados de consistencia interna tanto para las subescalas, con valores α entre .74 y .89, como para la escala total (α = .93); y valores de confiabilidad compuesta (ρ) entre .81 a .93 y 0,98 para la escala total. Con respecto a la validez, la adaptación argentina obtuvo evidencias para la estructura original de seis factores; asimismo, se observaron correlaciones directas con las estrategias poco adaptativas y correlaciones inversas con las estrategias adaptativas de regulación emocional cognitiva. Respecto de la evidencia de validez concurrente de la escala DERS, se observó que mayores dificultades en la regulación emocional se relacionaron con mayor afecto negativo, urgencia positiva y urgencia negativa y con menor afecto positivo.

Complementariamente, se administró un cuestionario *ad hoc* para recolectar información sociodemográfica de los participantes.

Procedimiento

La investigación se ajustó los lineamientos éticos de la declaración de Helsinki de la Asociación Médica Mundial (World Medical Association, 2013, Estándar 25) y la ley argentina 25.326 de Protección de los datos personales (2000). Antes de la administración del cuestionario, se informó a los estudiantes sobre el objetivo del estudio, sus derechos y el carácter anónimo y confidencial de sus respuestas. Además, se solicitó el consentimiento de ellos para participar de la investigación y se dejó constancia.

Para la aplicación del instrumento se contó con dos formatos, *online* (formulario de *Google*) y lápiz y papel. Los cuestionarios impresos fueron para, en primera instancia, estudiantes de cuarto año de psicología de la Universidad Siglo 21, en el marco de la asignatura Técnicas e Instrumentos de Exploración II, en el período de marzo a julio de 2022. En una segunda instancia, con la aprobación de las autoridades, se administraron en espacios comunes de la institución mencionada y de la Universidad de Mendoza (sede Río Cuarto) en el período de marzo a abril de 2023. Paralelamente, se difundió el enlace del formulario a través de *Instagram*, *Facebook* y *WhatsApp*.

Posteriormente a la recolección de los datos, se realizó una prueba de diferencias de medias en función de los dos formatos del cuestionario y no se encontraron resultados estadísticamente significativos (p > .05). Esto posibilitó unificar la muestra para el análisis de los datos.

Análisis de datos

En primer lugar, se evaluó la normalidad y la homocedasticidad de la variable y las subescalas a través de las pruebas de Kolmogorov-Smirnov y Levene, respectivamente. Luego, se obtuvieron los estadísticos descriptivos de la muestra para las puntuaciones en la DERS y se calculó el coeficiente Alfa de Cronbach para comprobar la confiabilidad de las puntuaciones por medio de la consistencia interna. Seguidamente, se realizó un análisis de diferencias entre grupos de acuerdo al

género mediante la prueba U de Mann Whitney. Finalmente, se construyeron dos tablas de baremos (una para la escala general y otra para cada una de las dimensiones) utilizando cálculo de percentiles y puntajes *T*. Se consideraron como puntos de corte los percentiles 25 y 75, propuestos en estudios similares a este (Dominguez-Lara, 2016; Morán et al., 2022) para clasificar los puntajes de la muestra en bajo, medio y alto. Con el fin de estimar la precisión de estos puntos de corte, se calculó el coeficiente *K2* (Livingston, 1972). Los análisis se realizaron con el software SPSS 23.

Resultados

El análisis de Kolmogorov-Smirnov arrojó valores p inferiores a .05 para todos los factores, por lo tanto se concluyó que estos no presentaban normalidad estadística. Por otra parte, la prueba de Levene mostró igualdad de varianza (homocedasticidad) solo en tres dimensiones: conciencia (p = .58), claridad (p = .51) e impulsos (p = .08). En base a estos resultados se decidió utilizar estadística no paramétrica.

En la Tabla 1 se muestran los resultados descriptivos de la escala general y de cada subescala para la muestra total. Como puede observarse, la consistencia interna fue satisfactoria en todos los casos con valores α entre .75 y .84 para las subescalas y .92 para la escala general.

Tabla 1
Estadísticos descriptivos de la muestra total para la DERS.

	N	M	DE	Mín	Máx	α
Conciencia	403	14,01	4,38	6	27	.76
Claridad	403	11,66	3,67	5	23	.75
Aceptación	403	13,81	5,57	6	30	.83
Estrategias	403	18,97	6,49	8	40	.83
Impulsos	403	12,83	5,06	6	30	.84
Objetivos	403	15,71	4,56	5	25	.80
Escala general	403	87,00	21,18	47	164	.92

Nota. M = Media; DE = Desviación estándar.

Se realizó una prueba U de Mann Whitney para explorar las diferencias en las dificultades en la regulación emocional de acuerdo con el género (se consideraron los grupos femenino y

Tabla 2
Prueba de U Mann-Whitney según género.

	U	*z*	*p*
Conciencia	13567	-1.76	.08
Claridad	14214	-1.12	.26
Aceptación	15240	-.11	.91
Estrategias	14430	-.91	.36
Impulsos	15253	-.10	.92
Objetivos	14144	-1.19	.23
Escala general	14956	-.39	.70

Tabla 3
Valores normativos de la DERS en población de estudiantes universitarios - Escala general.

PD	*Pc*	*T*	*K2*
59-61	5	34-36	1.00
62-64	10	37-39	0.97
65-66	15	40-41	0.97
67-69	20	42	0.96
70-73	25	43-44	0.95
74-76	30	45	0.95
77-78	35	46	0.94
79-81	40	47-48	0.93
82-83	45	49	0.93
84-86	50	50	0.92
87-89	55	51-52	0.92
90-91	60	53	0.92
92-95	65	54	0.92
96-99	70	55-56	0.92
100-103	75	57	0.93
104-107	80	58-59	0.94
108-114	85	60-62	0.95
115-127	90	63-65	0.96
128-163	95	66	0.97

Nota. *PD* = puntuación directa; *Pc* = puntuación percentilar; *T* = puntuación *T*; *K2* = coeficiente de confiabilidad de los puntos de corte.

masculino, dado que ningún participante señaló otra opción). Como se observa en la Tabla 2, los análisis no arrojaron diferencias estadísticamente significativas ($p > .05$), por lo tanto, se decidió no elaborar baremos distintos para cada género.

En las Tablas 3 y 4 se presentan los valores normativos para la escala total de dificultades en la regulación emocional y para cada una de sus dimensiones, respectivamente. Estos se construyeron utilizando el cálculo de percentiles y puntajes *T*. Asimismo, se obtuvieron los coeficientes *K2* (Livingston, 1972) para cada percentil, los cuales mostraron consistentemente puntuaciones muy elevadas (> .90).

Se utilizaron los puntos de corte percentilares propuestos por Dominguez-Lara (2016) y Morán et al. (2022) para delimitar los niveles alto (> *Pc* 75) y bajo (< *Pc* 25). Los elevados valores del coeficiente *K2* correspondientes a estos percentiles permitieron evidenciar la confiabilidad de los puntos de corte en cada dimensión evaluada y en la escala general. De manera práctica, obtener una puntuación directa de 100 a 103 en dicha escala indicaría que el nivel de dificultades en la regulación emocional del sujeto supera al del 75% de la población de referencia, lo cual sugeriría una puntuación elevada en dicha variable. Por el contrario, obtener una puntuación de 70 a 73 indicaría que el grado de dificultades en la regulación emocional evidenciado por la persona es superado por el del 75% de la población, porcentaje que representa un puntaje bajo. El mismo criterio interpretativo aplica a cada una de las subescalas.

Tabla 4
Valores normativos de la DERS en población de estudiantes universitarios.

PD						*Pc*	*T*	*K2*
Conciencia	**Claridad**	**Aceptación**	**Estrategias**	**Impulsos**	**Objetivos**			
7	6	6	10-11	6	9	5	34-36	1.00
8-9	7	7	12	7	10	10	37-39	0.97
10	8	8	13	8		15	40-41	0.97
		9			11	20	42	0.96
11	9		14	9	12	25	43-44	0.95
		10			13	30	45	0.95
12	10	11	15	10	14	35	46	0.94
			16	11		40	47-48	0.93
13	11	12	17		15	45	49	0.93
		13	18	12		50	50	0.92
14	12	14	19	13	16	55	51-52	0.92
15		15		14	17	60	53	0.92
	13	16-17	20-21		18	65	54	0.92
16			22	15		70	55-56	0.92
17	14	18-19	23	16	19	75	57	0.93
18	15		24-25	17	20	80	58-59	0.94
19	16	20-21	26-27	18	21	85	60-62	0.95
20-21	17-18	22-23	28-31	19-21	22-23	90	63-65	0.96
22	19	24	32	22	24	95	66	0.97

Nota. *PD* = puntuación directa; *Pc* = puntuación percentilar; *T* = puntuación *T*; *K2* = coeficiente de confiabilidad de los puntos de corte.

Discusión

La RE es un prerrequisito fundamental para el funcionamiento social que impacta tanto en el bienestar como en la salud mental de las personas (Kohn et al., 2014). La fase crítica que experimentan los jóvenes universitarios se caracteriza por enfrentar nuevas exigencias y tensiones, lo que incrementa la susceptibilidad a problemas de salud mental asociados al estrés. Los estudiantes se ven sometidos a diversos factores estresantes, como la carga académica, la constante presión por avanzar en sus estudios, la competencia entre compañeros y, de manera particular en el ámbito local, las implicaciones financieras que conlleva cursar una carrera universitaria (Moretti & Medrano, 2014). Ante este escenario, los estudiantes pueden experimentar consecuencias negativas en la regulación emocional debido al desbalance entre las exigencias del contexto académico y los recursos de la persona (Dominguez-Lara, 2016; Moretti & Medrano, 2014), lo que convierte a los universitarios en una población proclive a desarrollar trastornos psicológicos y malestar general (Micin & Bagladi, 2011).

En efecto, está demostrado que los estudiantes universitarios que presentan mayores dificultades para dirigir su comportamiento a ciertos objetivos y propósitos y mayor dificultad para aceptar sus emociones negativas, reconocerlas y diferenciarlas de otras, no sólo reportan menores niveles de bienestar psicológico, sino que también comienzan a tener mayor desinterés por los estudios y baja eficacia académica (Khalil et al., 2020).

Esto confirma la necesidad de continuar explorando la RE en estudiantes universitarios, considerando que la DERS está disponible como una herramienta que puede evaluar dificultades en la regulación emocional que son clínicamente significativas. Esto es importante debido al papel fundamental que juega este proceso en la comprensión de síntomas y comportamientos desadaptativos (Coutinho et al., 2010). Si bien la DERS se ha utilizado en diversas investigaciones argentinas (del Valle et al., 2020; Medrano et al., 2013; Medrano & Trógolo, 2014; Michelini & Godoy, 2022), aún no se encuentran baremos adecuados a la población, lo que impide el uso de la escala en contextos clínicos o de clasificación. Por esta razón, el objetivo de este trabajo fue obtener datos normativos para estudiantes universitarios argentinos.

Los análisis realizados proveen la transformación de los puntajes brutos de la escala y sus factores en puntajes normativos que permiten la interpretación de las mediciones brindadas por el instrumento. Teniendo en cuenta la distribución no normal de la variable, se utilizaron transformaciones no lineales que dan cuenta de la magnitud de las dificultades en la regulación emocional del individuo evaluado, respecto al grupo cultural y al ciclo evolutivo al cual pertenece (American Educational Research Association [AERA] et al., 2014). Este tipo de transformaciones, especialmente los percentiles, se caracterizan por su fácil interpretación para los profesionales que no cuentan con conocimientos avanzados en estadística y que desean emplear mediciones en sus tareas profesionales (Tornimbeni et al., 2008). De acuerdo a los resultados, estos puntajes obtuvieron elevados índices de confiabilidad e indicaron su precisión en el ordenamiento de las puntuaciones (Dominguez-Lara, 2016).

Para incrementar la utilidad práctica del presente estudio, siguiendo las recomendaciones de Tornimbeni et al. (2008), se sugirieron puntos de corte para identificar distintos niveles de dificultad en la regulación emocional. Las puntuaciones brutas que superan el percentil 75 señalan niveles elevados de dificultades. Por el contrario, los puntajes por debajo del percentil 75 no indican déficits y ubican al sujeto dentro de lo esperado para su grupo de referencia. En esta línea, puede

considerarse que quienes obtengan puntajes por debajo del percentil 25 pueden presentar una saludable gestión emocional.

Es importante mencionar que, previo a la obtención de estos datos normativos, se realizaron análisis para determinar si era pertinente presentar dichos datos diferenciados por género, tal como es usual en los estudios de construcción de baremos (AERA et al., 2014). Los análisis de diferencias de medias indicaron que no existían diferencias significativas entre hombres y mujeres en sus puntajes. Al respecto, existen múltiples estudios que reportan diferencias de género en la utilización de las diferentes estrategias de RE, y que las mujeres son más reactivas emocionalmente a los diferentes estresores (Nolen-Hoeksema, 2012). No obstante, la DERS específicamente indaga sobre la percepción de la persona sobre su capacidad para recurrir a las estrategias de regulación emocional en momentos de activación intensa y cuán efectivos son esos intentos (de la Rosa-Gómez et al., 2021). Por ello, los resultados obtenidos en este estudio coinciden con los publicados por Gratz y Roemer (2004) para la versión original estadounidense, así como también con los de la versión griega (Mitsopoulou et al., 2013), y parcialmente con las versiones turca, italiana, española y brasileña (Giromini et al., 2012; Hervás & Jódar, 2008; Koich-Miguel et al., 2017; Rugancı & Gençöz, 2010) en las cuales se encontraron diferencias significativas solo en algunos factores. En conclusión, y siguiendo a Giromini et al. (2017), estos antecedentes de la DERS y el presente estudio son consistentes en cuanto a que hombres y mujeres no presentan diferencias significativas en sus puntuaciones, por lo que se exime de la necesidad de contar con puntajes normativos diferenciados.

La principal importancia de este estudio radica en que presenta la posibilidad de contar con una herramienta que amplíe el campo de utilidad de la escala en contextos argentinos. El desarrollo de estudios psicométricos en Argentina se ha expandido en los últimos años (Becerra & Abal, 2021; de la Iglesia & Freiberg-Hoffmann, 2015), en los cuales se puede observar un incremento de publicaciones de estudios instrumentales cuyo objetivo es construir, adaptar y validar tests psicométricos de múltiples constructos. Sin embargo, la construcción de baremos no ha tenido el mismo avance, por lo que estos instrumentos solo están disponibles para su utilización en diseños específicos de investigación. Como resultado, quedan desprovistos de esta herramienta todos los estudios de prevalencia y aquellos en que la clasificación forma parte de su procedimiento. Sumado a ello, uno de los principales usos de las técnicas de evaluación psicológica es el diagnóstico y, en este contexto de aplicación, contar con datos normativos de interpretación resulta primordial. Específicamente en el ámbito clínico, la evaluación de las dificultades en la regulación emocional resulta de suma relevancia en cualquier proceso psicoterapéutico dado su demostrado efecto negativo en el bienestar psicológico, la adaptación social y la salud mental en general (Cludius et al., 2020; Hu et al., 2014; Saxena et al., 2011).

A pesar de la utilidad de los resultados, es importante indicar que el presente estudio tiene limitaciones. Por un lado, el método de muestreo fue no probabilístico y, por ende/por lo tanto, quedó privado de todas las bondades que ese tipo de muestreo presenta en términos de representatividad (Tornimbeni et al., 2008). Por otro lado, la diversidad de carreras universitarias abarcadas fue limitada, por lo que sería relevante indagar si existen antecedentes que den cuenta de una asociación entre el perfil vocacional y la regulación emocional que sustenten una ampliación de la muestra referencial. Otros estudios que completan el presente trabajo refieren al análisis sobre el poder discriminativo, la fiabilidad y la capacidad predictiva de los puntos de corte sugeridos (Dominguez-

Lara, 2016, 2018; Morán et al., 2023). A pesar de estas limitaciones, los datos normativos presentados permiten hacer uso de la escala para evaluaciones de tipo *screening*, en complemento con otras fuentes de información que contrasten los resultados obtenidos para cada individuo.

Referencias

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

Arnett, J. J., Žukauskienė, R., & Sugimura, K. (2014). The new life stage of emerging adulthood at ages 18-29 years: Implications for mental health. *The Lancet Psychiatry, 1*(7), 569-576. https://doi.org/10.1016/S2215-0366(14)00080-7

Becerra, L., & Abal, F. (2021, 24 de julio). *Análisis historiográfico de la medición en psicología en Argentina. Desarrollos e influencias en la psicometría en los fines del siglo XX [Presentación trabajo libre]*. XVII Jornadas Internacionales de Investigación en Psicología UCES 2021. XIX Jornadas Internacionales de Actualización del Algoritmo David Liberman. II Simposio de especialistas en Salud Mental en Emergencias y Desastres, Buenos Aires, Argentina. http://dspace.uces.edu.ar:8180/xmlui

Bolgeri-Escorza, P., Rojas-Duarte, F., Pastén, M. J., Vega-Vega, K., & San Martín-Henríquez, M. (2020). Regulación emocional, atención plena y bienestar psicológico en estudiantes universitarios de primer año. *Eureka, 17*(1), 22-37. https://ojs.psicoeureka.com.py/index.php/eureka

Cabanach, R. G., Fernández-Cervantes, R., Souto-Gestal, A., & Suárez-Quintanilla, J. A. (2018). Regulación emocional y estrés en estudiantes universitarios de fisioterapia. *Revista Mexicana de Psicología, 35*(2), 167-178. https://psycnet.apa.org/record/2018-39520-006

Cludius, B., Mennin, D., & Ehring, T. (2020). Emotion regulation as a transdiagnostic process. *Emotion, 20*(1), 37-42. https://doi.org/10.1037/emo0000646

Compas, B. E., Jaser, S. S., Bettis, A. H., Watson, K. H., Gruhn, M. A., Dunbar, J. P., Williams, E., & Thigpen, J. C. (2017). Coping, emotion regulation, and psychopathology in childhood and adolescence: A meta-analysis and narrative review. *Psychological Bulletin, 143*(9), 939-991. https://doi.org/10.1037/bul0000110

Coutinho, J., Ribeiro, E., Ferreirinha, R., & Dias, P. (2010). The Portuguese version of the Difficulties in Emotion Regulation Scale and its relationship with psychopathological symptoms. *Archives of Clinical Psychiatry, 37*(4), 145-151. https://doi.org/10.1590/S0101-60832010000400001

de la Iglesia, G., & Freiberg-Hoffmann, A. (2015, 25-28 de noviembre). *Psicometriar: Una propuesta de divulgación científica de la psicometría argentina [Presentación de póster]*. VII Congreso Internacional de Investigación y Práctica Profesional en Psicología, XXII Jornadas de Investigación, XI Encuentro de Investigadores en Psicología del MERCOSUR, Buenos Aires, Argentina. https://www.aacademica.org/000-015/111

de la Rosa-Gómez, A., Hernández-Posadas, A., Valencia, P. D., & Guajardo-Garcini, D. A. (2021). Análisis dimensional de la Escala de Dificultades en la Regulación Emocional (DERS-15) en universitarios mexicanos. *Revista Evaluar, 21*(2), 80-97. https://doi.org/10.35670/1667-4545.v21.n2.34401

del Valle, M., Zamora, E. V., Khalil, Y., & Altamirano, M. (2020). Rasgos de personalidad y dificultades de regulación emocional en estudiantes universitarios. *Psicodebate, 20*(1), 56-67. https://doi.org/10.18682/pd.v20i1.1877

Dominguez-Lara, S. A. (2016). Datos normativos de la Escala de Procrastinación Académica en estudiantes de psicología de Lima. *Revista Evaluar, 16*(1), 20-30. https://doi.org/10.35670/1667-4545.v16.n1.15715

Dominguez-Lara, S. (2018). Magnitud del efecto, una guía

rápida. *Educación Médica, 19*(4), 251-254. https://doi.org/10.1016/j.edumed.2017.07.002

Eichen, D. M., Chen, E., Boutelle, K. N., & McCloskey, M. S. (2017). Behavioral evidence of emotion dysregulation in binge eaters. *Appetite, 111*, 1-6. https://doi.org/10.1016/j.appet.2016.12.021

Giromini, L., Ales, F., de Campora, G., Zennaro, A., & Pignolo, C. (2017). Developing age and gender adjusted normative reference values for the Difficulties in Emotion Regulation Scale (DERS). *Journal of Psychopathology and Behavioral Assessment, 39*(4), 705-714. https://doi.org/10.1007/s10862-017-9611-0

Giromini, L., Velotti, P., de Campora, G., Bonalume, L., & Zavattini, G. C. (2012). Cultural adaptation of the Difficulties in Emotion Regulation Scale: Reliability and validity of an Italian version. *Journal of Clinical Psychology, 68*(9), 989-1007. https://doi.org/10.1002/jclp.21876

Gratz, K. L., & Roemer, L. (2004). Multidimensional assessment of emotion regulation and dysregulation: Development, factor structure, and initial validation of the Difficulties in Emotion Regulation Scale. *Journal of Psychopathology and Behavioral Assessment, 26*(1), 41-54. https://doi.org/10.1023/B:JOBA.0000007455.08539.94

Gross, J. J., & Jazaieri, H. (2014). Emotion, emotion regulation, and psychopathology: An affective science perspective. *Clinical Psychological Science, 2*(4), 387-401. https://doi.org/10.1177/2167702614536164

Guzmán-González, M., Mendoza-Llanos, R., Garrido-Rojas, L., Barrientos, J., & Urzúa, A. (2020). Propuesta de valores de referencia para la Escala de Dificultades de Regulación Emocional (DERS-E) en población adulta chilena. *Revista Médica de Chile, 148*(5), 644-652. http://doi.org/10.4067/S0034-98872020000500644

Guzmán-González, M., Trabucco, C., Urzúa, A. M., Garrido, L., & Leiva, J. (2014). Validez y confiabilidad de la versión adaptada al español de la Escala de Dificultades de Regulación Emocional (DERS-E) en población chilena. *Terapia Psicológica, 32*(1), 19-29. https://doi.org/10.4067/S0718-48082014000100002

Haynos, A. F., Wang, S. B., & Fruzzetti, A. E. (2018). Restrictive eating is associated with emotion regulation difficulties in a non-clinical sample. *Eating Disorders, 26*(1), 5-12. https://doi.org/10.1080/10640266.2018.1418264

Hervás, G., & Jódar, R. (2008). Adaptación al castellano de la Escala de Dificultades en la Regulación Emocional. *Clínica y Salud, 19*(2), 139-156. https://journals.copmadrid.org/clysa

Hu, T., Zhang, D., Wang, J., Mistry, R., Ran, G., & Wang, X. (2014). Relation between emotion regulation and mental health: A meta-analysis review. *Psychological Reports, 114*(2), 341-362. https://doi.org/10.2466/03.20.PR0.114k22w4

Khalil, Y., del Valle, M. V., Zamora, E. V., & Urquijo, S. (2020). Dificultades de regulación emocional y bienestar psicológico en estudiantes universitarios. *Subjetividad y Procesos Cognitivos, 24*(1), 69-83. http://dspace.uces.edu.ar:8180/xmlui/handle/123456789/41

Kohn, N., Eickhoff, S. B., Scheller, M., Laird, A. R., Fox, P. T., & Habel, U. (2014). Neural network of cognitive emotion regulation - An ALE meta-analysis and MACM analysis. *NeuroImage, 87*(1), 345-355. https://doi.org/10.1016/j.neuroimage.2013.11.001

Koich-Miguel, F., Giromini, L., Stivaleti-Colombarolli, M., Zuanazzi, A. C., & Zennaro, A. (2017). A Brazilian investigation of the 36- and 16-item Difficulties in Emotion Regulation Scales. *Journal of Clinical Psychology, 73*(9), 1146-1159. https://doi.org/10.1002/jclp.22404

Koole, S. L. (2009). The psychology of emotion regulation: An integrative review. *Cognition and Emotion, 23*(1), 4-41. https://doi.org/10.1080/02699930802619031

LeDoux, J. E., & Brown, R. (2017). A higher-order theory of emotional consciousness. *Proceedings of National Academy of Sciences of the United States of America, 114*(10), 2016-2025. https://doi.org/10.1073/pnas.1619316114

León, O. G., & Montero, I. (2002). *Métodos de investigación en Psicología y Educación* (3ra ed.). McGraw-Hill.

Ley 25.326. (2000). *Protección de los Datos Personales*. Argentina. https://servicios.infoleg.gob.ar/infolegInternet

Livingston, S. A. (1972). Criterion-referenced applications of classical test theory. *Journal of Educational Measurement, 9*(1), 13-26. https://doi.org/10.1111/j.1745-3984.1972.tb00756.x

Marín-Tejeda, M., Robles-García, R., González-Forteza, C., & Andrade-Palos, P. (2012). Propiedades psicométricas de la escala "Dificultades en la Regulación Emocional" en español (DERS-E) para adolescentes mexicanos. *Salud mental, 35*(6), 521-526. http://www.revistasaludmental.mx/index.php/salud_mental

Medrano, L. A., Moretti, L., Ortiz, A., & Pereno, G. (2013). Validación del Cuestionario de Regulación Emocional Cognitiva en universitarios de Córdoba, Argentina. *Psykhe, 22*(1), 83-96. https://doi.org/10.7764/psykhe.22.1.473

Medrano, L., & Trógolo, M. (2014). Validación de la Escala de Dificultades en la Regulación Emocional en la población universitaria de Córdoba, Argentina. *Universitas Psychologica, 13*(4), 15- 26. https://doi.org/10.11144/Javeriana.UPSY13-4.vedr

Michelini, Y., & Godoy, J. C. (2022). Adaptación argentina de dos Escalas de Dificultades en la Regulación Emocional en adultos emergentes universitarios. *Psykhe, 31*(1), 1-20. https://doi.org/10.7764/psykhe.2020.22585

Micin, S., & Bagladi, V. (2011). Salud mental en estudiantes universitarios: Incidencia de psicopatología y antecedentes de conducta suicida en población que acude a un servicio de salud estudiantil. *Terapia Psicológica, 29*(1),53-64. http://doi.org/10.4067/S0718-48082011000100006

Mitsopoulou, E., Kafetsios, K., Karademas, E., Papastefanakis, E., & Simos, P. G. (2013). The Greek version of the Difficulties in Emotion Regulation Scale: Testing the factor structure, reliability and validity in an adult community sample. *Journal of Psychopathology and Behavioral Assessment, 35*(1), 123-131. https://doi.org/10.1007/s10862-012-9321-6

Morán, V. E., Villarrubia, M. D., Fabbro, N., & Natera, M. Z. (2023). Datos normativos en población de adultos argentinos para la Escala Breve de Ansiedad y Fobia Social- SPAI-B. *Revista Argentina de Ciencias del Comportamiento, 15*(2), 50-59. https://revistas.unc.edu.ar/index.php/racc/article/view/33308

Morán, V., Villarrubia, M., & Curi, M. V. (2022). Datos normativos en población de adultos argentinos para la Escala de Conexión Social. *Revista de Investigación en Psicología, 25*(2), 89-104. https://doi.org/10.15381/rinvp.v25i2.23461

Moretti, L. S., & Medrano, L. A. (2014). Estructura factorial del Cuestionario de Estrés Percibido en la población universitaria. *Revista Evaluar, 14*(1), 67-83. https://doi.org/10.35670/1667-4545.v14.n1.11518

Nolen-Hoeksema, S. (2012). Emotion regulation and psychopathology: The role of gender. *Annual Review of Clinical Psychology, 8*(1), 161-187. https://doi.org/10.1146/annurev-clinpsy-032511-143109

Ochsner, K. N., & Gross, J. J. (2008). Cognitive emotion regulation: Insights from social cognitive and affective neuroscience. *Current Directions in Psychological Science, 17*(2), 153-158. https://doi.org/10.1111/j.1467-8721.2008.00566.x

Rugancı, R. N., & Gençöz, T. (2010). Psychometric properties of a Turkish version of the Difficulties in Emotion Regulation Scale. *Journal of Clinical Psychology, 66*(4), 442-455. https://doi.org/10.1002/jclp.20665

Saxena, P., Dubey, A., & Pandey, R. (2011). Role of emotion regulation difficulties in predicting mental health and well-being. *Journal of Projective Psychology & Mental Health, 18*(2), 147-155. https://2d-crowd.com/sisJournal/full_issues

Sierra-Bravo, R. (2007). *Técnicas de investigación social: Teoría y ejercicios (14ª Ed.)*. International Thomson.

Simons, R. M., Hahn, A. M., Simons, J. S., & Murase, H. (2017). Emotion dysregulation and peer drinking norms uniquely predict alcohol-related problems via motives. *Drug and Alcohol Dependence, 177*, 54-58. https://doi.org/10.1016/j.drugalcdep.2017.03.019

Tornimbeni, S., Pérez, E., & Olaz, F. (2008). *Introducción a*

la psicometría. Paidós.

Vargas-Gutiérrez, R. M., & Muñoz-Martínez, A. M. (2013). La regulación emocional: Precisiones y avances conceptuales desde la perspectiva conductual. *Psicologia USP, 24*(2), 225-240. https://doi.org/10.1590/S0103-65642013000200003

Vilhena-Churchill, N., & Goldstein, A. L. (2014). Child maltreatment and marijuana problems in young adults: Examining the role of motives and emotion dysregulation. *Child Abuse & Neglect, 38*(5), 962-972. https://doi.org/10.1016/j.chiabu.2013.10.009

Weiss, N. H., Bold, K. W., Sullivan, T. P., Armeli, S., & Tennen, H. (2017). Testing bidirectional associations among emotion regulation strategies and substance use: A daily diary study. *Addiction, 112*(4), 695-704. https://doi.org/10.1111/add.13698

World Medical Association. (2013). World Medical Association Declaration of Helsinki: Ethical principles for medical research involving human subjects. *Journal of the American Medical Association, 310*(20), 2191-2194. http://doi.org/10.1001/jama.2013.281053

EVALUAR

2024, Vol. 24, No. 1
ISSN 1667-4545
Recuperado de https://revistas.unc.edu.ar/index.php/revaluar

Laboratorio de Evaluación Psicológica y Educativa
Facultad de Psicología - Universidad Nacional de Córdoba

Psychosocial Maturity assessment on juvenile justice: A content validity analysis of a novel tool

Evaluación de la Madurez Psicosocial en Justicia Juvenil: Análisis de la validez de contenido de una herramienta novedosa

Elena Palacios-van Isschot [1], Karin Arbach * [2], Antonio Andrés-Pueyo [1]

1 - Departamento de Psicología Clínica y Psicobiología, Universidad de Barcelona. Barcelona, España.
2 - Instituto de Investigaciones Psicológicas, Consejo Nacional de Investigaciones Científicas y Tecnológicas y Universidad Nacional de Córdoba. Córdoba, Argentina.

Recibido: 27/11/2023 **Revisado:** 10/12/2023 **Aceptado:** 14/12/2023

Abstract

Psychosocial maturity (PM) is a potential factor influencing adolescent decision making and a wide range of social and interpersonal behaviors in adolescents. The current study represents an initial effort to design a new psychosocial maturity test suitable for assessing this construct in the forensic context and juvenile justice settings. Its aim is to establish content validity for a novel assessment tool. After a literature review, 38 items were selected from various existing specific tests and protocols, and 41 experts were instructed to conduct a content validity analysis on them. Content Validity Index results show that 92% of the items were classified by experts as representative to the construct, and Factorial Validity Index results show that experts associated 79% of the items with the correct component of the construct according to theoretical criteria. In conclusion, the majority of items were found to be representative of the construct and of their individual components, providing a valid foundation for the development of a new PM assessment tool. In this study the relevance and implications of the results for judicial tasks are discussed.

Keywords: *psychosocial maturity, juvenile justice, expert judges, content validity, psychological assessment*

Resumen

La madurez psicosocial (MP) es un factor potencial que influye en la toma de decisiones de los adolescentes y su comportamiento social e interpersonal. El estudio actual representa un primer paso para diseñar una nueva prueba de madurez psicosocial adecuada para evaluar este constructo en el contexto forense y en entornos de justicia juvenil. Su objetivo es establecer la validez de contenido para una nueva herramienta de evaluación. Después de una revisión de la literatura, se seleccionaron 38 ítems de diversas medidas y protocolos específicos existentes y se instruyó a 41 expertos para llevar a cabo un análisis de validez de contenido. Los resultados del Índice de Validez de Contenido muestran que el 92% de los ítems analizados fueron clasificados por los expertos como representativos del constructo, y los resultados del Índice de Validez Factorial muestran que los expertos asociaron el 79% de los ítems al constructo propuesto. En conclusión, la mayoría de los ítems se consideraron representativos del constructo y de sus componentes individuales, proporcionando una buena base para desarrollar una herramienta de evaluación de la MP. Se plantea la relevancia e implicaciones de los resultados en las tareas judiciales.

Palabras clave: *madurez psicosocial, justicia juvenil, juicio de expertos, validez de contenido, evaluación psicológica*

***Correspondence to:** Dra. Karin Arbach. Boulevard de la Reforma. Facultad de Psicología, Box B9. Universidad Nacional de Córdoba. CP 5000.
E-mail: k_arbach@unc.edu.ar. https://orcid.org/0000-0003-1753-4693

How to cite: Palacios-van Isschot, E., Arbach, A., & Andrés-Pueyo, A., (2024): Psychosocial Maturity assessment on juvenile justice: A content validity analysis of a novel tool. *Revista Evaluar, 24*(1), 14-27. Retrieved from https://revistas.unc.edu.ar/index.php/revaluar

Participaron en la edición de este artículo: Stefano Macri, Eugenia Barrionuevo, Florencia Ruiz, Jorge Bruera.

Introduction

Diminished capacity, resulting from immaturity, is relevant for criminal court decisions, in particular those related to the antisocial behavior of adolescents and other legal questions (Riggs-Romaine, 2018; Wakeling & Barnett, 2017). Despite the growing demand from forensic psychologists, the development of tools designed to assess maturity in the context of juvenile justice has received limited attention (Wakeling & Barnett, 2017). The existing research on maturity has predominantly focused on cognitive differences between adolescent and adult judgment, while the role of social environment, and its interaction with personal characteristics has been poorly developed (Steinberg & Cauffman, 1996). This issue requires attention because cognitive development results not only from the maturation process or independent interaction with the outside world but also from direct exposure to the world (Narváez-Burbano & Obando-Guerrero, 2023).

The scarcity of research on the role of the social environment is a critical concern, given that adolescents attain biological maturity before achieving social and emotional maturity, thus giving rise to a pronounced "maturity gap" (Belsky et al., 2020; Cavanagh, 2022; Moffitt, 1993, 2003; Ozkan & Worrall, 2017). This disparity is notably prevalent in life-course persistent offenders, who deal with neurological deficits that impede the maturation processes related to self-regulation (Moffitt, 2003; Ozkan & Worrall, 2017). Consequently, research in this area has not been sufficiently comprehensive.

A more constructive approach for the assessment of maturity in juvenile justice settings emerges with the concept of psychosocial maturity (PM) (Steinberg & Cauffman, 1996), which represents maturity of judgment. It refers to the complex process of individual decision-making influenced by cognitive, emotional and social factors (Steinberg & Cauffman, 1996). PM consists of three fundamental components: *temperance, perspective* and *responsibility*. *Temperance* denotes "the capacity to limit impulsivity, avoid extremes in decision-making, evaluate a situation thoroughly and to seek advice before acting" (p.745; Cauffman & Steinberg, 2000). Then, *perspective* involves recognizing the complexity of a situation and making decisions within a larger context (Steinberg & Cauffman, 1996). Finally, *responsibility* represents the ability to be in charge of one's behavior and to resist peer influences, and encompasses three subcomponents: *autonomy, work orientation* and *identity* (Cauffman & Steinberg, 2000).

Studies have demonstrated that PM offers a promising explanation for the process of avoiding criminal behavior during late adolescence and early adulthood, known as *desistance* (Monahan et al., 2009; Rocque et al., 2019). Differences in PM predict decision-making abilities, independently from age or gender (Riggs-Romaine, 2018) in potentially antisocial situations. This compelling empirical evidence has sparked interest in the development of a tool for the assessment of young offenders' PM through their transition into early adulthood (Steinberg et al., 2015). The operationalization of this construct into a measurable instrument carries significant implications for various facets within the domain of juvenile justice.

As of today, the juvenile justice system has yet to implement essential measures aimed at fostering the healthy development of adolescents, despite its primary mission of rehabilitating youth (Cavanagh, 2022). Research has demonstrated that, in legal proceedings involving young offenders, the legal aspects of their cases tend to carry more weight in influencing court decisions, compared to factors related to mental health or maturity (Cauffman et al., 2007; Lambie & Randell,

2013). Nonetheless, adolescents are inherently disadvantaged in comparison to adults when they are faced with choices in antisocial settings, primarily because they have a limited capacity to fully grasp the consequences of their actions. Studies indicate that the incarceration of adolescents can negatively impact their psychosocial development. This impact occurs as incarceration reduces their opportunities for typical social experiences, disrupts their contact with important social influences, and increases their interactions with peers engaged in antisocial behavior. Consequently, this increases the likelihood of adolescents engaging in further delinquent activities (Cavanagh, 2022).

The implementation of a tool for the assessment of PM holds significant potential for augmenting the effectiveness of forensic procedures and judicial decision-making. This innovative instrument would empower the legal system to tailor punitive measures and judgments precisely to the specific levels of maturity displayed by each offender, thereby heralding a new era in the pursuit of justice. Moreover, such a tool would play an essential role in mitigating the potential adverse effects of incarceration on the mental health and psychosocial development of young individuals (Cavanagh, 2022; Lambie & Randell, 2013).

At the professional practice level, the application of this instrument would markedly enhance the precision of diagnostic processes and facilitate the design of interventions targeted at adolescents at risk, thereby promoting effective prevention strategies. Notably, this endeavor aspires to create a novel assessment tool for PM tailored to the Spanish-speaking population, and aims to make it readily accessible to juvenile justice professionals in Latin America and Spain, where limited reviews are available regarding assessment tools for the juvenile justice context. In the long term, the availability of such an instrument would prove highly advantageous for the numerous adolescents navigating legal proceedings (Wenger & Andres-Pueyo, 2016).

In previous studies, the assessment of PM has commonly relied on pre-existing measures that have been validated for similar constructs. To illustrate, assessments of responsibility have frequently drawn from the Psychosocial Maturity Inventory (PSMI, Greenberger et al., 1975), while evaluations of *temperance* have leaned on the Weinberger Adjustment Inventory (WAI, Weinberger & Schwartz, 1990), whereas *perspective* has been explored using the Consideration of Future Consequences Scale (CFC, Strathman et al., 1994), as seen in consulted studies (Cauffman & Steinberg, 2000; Pailing & Reniers, 2018; Riggs-Romaine, 2018).

Nevertheless, the pursuit of a single valid and reliable assessment measure, specifically tailored for evaluating PM within the juvenile justice context, remains an ongoing challenge. In response to this gap in the literature and guided by the authors' definitions and a comprehensive literature review, a selection of 38 items was drawn from existing measures to aptly represent the PM construct. Given the frequent concern about potential misalignment between the data acquired from the instrument's application and the intricate reality it seeks to encapsulate (Expósito et al., 2023), several practicing forensic psychologists were engaged to participate in a content validity analysis of this newly devised assessment tool.

Conducting a content validity study is crucial, especially when no existing measure is available to operationalize the construct in question (Rubio et al., 2003). Commencing such a study is essential to prevent extensive revisions during testing and to ensure the representativeness of its content (Almanasreh et al., 2019). Consequently, the primary objective of this investigation was to conduct a thorough content validity assessment for the 38 selected items derived from prior re-

search. The study's specific objectives were delineated as follows:

O1. To assess, through the expert judgment of authorities in the field, the extent to which each item effectively represents the PM construct (Construct representativeness).

O2. To assess, via expert judgment, the degree to which each item represents a specific component of the PM construct (Component representativeness).

Method

Participants

The selection of experts for this study was based on rigorous criteria encompassing qualifications, substantial experience, clinical expertise and relevant training. Content validity analyses typically recommend a minimum of three experts for such assessments (Lynn, 1986), although some suggest involving as many as twenty experts for robust evaluations (Almanasreh et al., 2019). To ensure comprehensive evaluation, we assembled three distinct groups of experts, with a total of 41 participants. These experts were invited to assess and qualify the items comprising the scale by participating in an online survey administered through the Qualtrics platform. The expert groups were defined as follows:

Group 1: Comprising 14 researchers affiliated with the University of Barcelona, with specialized expertise in the field of Forensic Psychology.

Group 2: Comprising 19 Forensic Psychologists employed by the Spanish Ministry of Justice.

Group 3: Comprising 8 individuals who are members of the Association of Forensic Psychology (APF) and work within the Administration of Justice in Spain.

Recruitment process

The recruitment of experts for this study was meticulously organized, tailored to each expert group, and conducted with a focus on transparency and clarity.

Group 1. Experts from this group received personalized invitations via email. These invitations contained essential background information about the study, along with a URL link to access the survey. Participants in this group were provided with comprehensive definitions of the components of the construct under examination and detailed instructions for the content validity analysis.

Group 2. Experts in this group were invited to participate during an online course focused on Psychosocial Maturity (PM), which was led by the authors A.A. and E.P. The course included an in-depth presentation on the PM model, as proposed by Cauffman and Steinberg (2000). Subsequently, members of this group were also sent personalized email invitations, mirroring the information provided to Group 1, to access the survey.

Group 3. Experts in this group were invited via a representative from the Association of Forensic Psychology (APF), who directly provided them with the URL link to access the survey. Unlike Groups 1 and 2, experts from Group 3 did not receive any contextual information about the study. Their survey access was restricted to the instructions and definitions required for the analysis.

This approach was carefully designed to ensure that all expert groups could participate while considering their specific contexts and information needs.

Procedure and materials

To select items for our assessment, we conducted an extensive review of prior studies on PM. This review identified the most commonly utilized scales and instruments in the existing literature, including: (1) the Weinberger Adjustment Inventory (WAI) developed by Weinberger and Schwartz (1990), (2) the Consideration of Future Consequences Scale (CFC) by Strathman et al. (1994), (3) the Psychosocial Maturity Inventory (PSMI) by Greenberger et al. (1975) and (4) the Psychological Maturity Scale for Adolescents (PSYMAS), as introduced by Morales-Vives et al. (2013). This initial literature review served as the foundation for the creation of a preliminary questionnaire, titled the MAYAS (Wenger-Amengual, 2018), comprising 38 items.

Each of the 38 items underwent a complete semantic and syntactic revision, based on the original items extracted from the WAI, CFC, PSMI and PSYMAS scales mentioned above.

Item selection for preliminary version

In the development of the preliminary version of the questionnaire, we took a structured approach to represent the key components of PM.

Temperance. This component was represented by items drawn from both the Impulse Control and the Consideration of Others subscales of the Weinberger Adjustment Inventory (WAI) developed by Weinberger and Schwartz (1990).

Perspective. To capture the perspective component, we utilized items from the Consideration of Future Consequences Scale (CFC) by Strathman et al. (1994). Additionally, we incorporated items from the Consideration of Others scale of the WAI (Weinberger & Schwartz, 1990).

Responsibility. The responsibility component was represented through items sourced from the PSYMAS. This tool was selected due to its specific and contemporary nature compared to the traditionally employed PSMI. The PSYMAS model, based on the individual adequacy component of the PSMI, comprises the same three subcomponents that define responsibility in the Cauffman and Steinberg PM model (2000): *Autonomy (or self-reliance), Identity* and *Work orientation.*

This approach ensured that the preliminary version encompassed relevant items to accurately represent the fundamental components of PM, by the use of validated instruments from the literature.

Preliminary prevision

In the current study, a comprehensive renewal of the pilot protocol of the MAYAS (Wenger-Amengual, 2018) was undertaken, resulting in the creation of the "Psychological Maturity Test for Adolescents in a Forensic Context" (PMTAFC). This revision process involved two critical steps to ensure the comprehensive representation of the PM construct.

1. Literature Review and Item Categorization: Firstly, an additional review of specialized literature was conducted to facilitate a systematic categorization of the items. This categorization was instrumental in ensuring that every facet of the PM construct was thoroughly and appropriately represented. Each item was systematically associated with its original source, whether it originated from the Weinberger Adjustment Inventory (WAI) by Weinberger and Schwartz (1990), the Consideration of Future Consequences Scale (CFC) by Strathman et al. (1994) or the Psychological Maturity Assessment Scale (PSYMAS) by Morales-Vives et al. (2013).

2. Theoretical Assignment and Assessment: Subsequently, each item was theoretically linked to one of the components of the PM construct. This association allowed for a comprehensive evaluation of each item's comprehensibility, scope, and accurate translation from English to Spanish (through a back-translation process). To address potential gaps in the representation of the *perspective* component, several items from the Interpersonal Reactivity Index (IRI) developed by Davis (1980) were incorporated. Moreover, the vocabulary and syntactic structure of all items in this revised draft (PMTAFC) were thoughtfully adapted to optimize comprehension for Spanish-speaking adolescents.

In summation, the scale being examined in this study represents an enhanced iteration of the MAYAS, now referred to as PMTAFC. This revision process significantly improved the translation, comprehensibility, and the overall representativeness of each component.

Expert assessment protocol

With a panel of experts in place, a methodical process was implemented to solicit their evaluations of the PMTAFC items. This assessment involved the experts' completion of a specialized form, designed to guide their analysis through two distinct tasks.

Task 1: Construct Representativeness (O1) - In the initial task, experts were presented with a fundamental question: *Is the item representative of the psychosocial maturity construct?* Their responses were sought using a binary response system, allowing them to choose between *yes* or *no*. This task served as a crucial step to gauge the overall representativeness of the construct.

Task 2: Component Representativeness (O2) - Experts who provided an affirmative response in Task 1 (yes) were subsequently directed to Task 2. Here, their role was to determine the specific component of the PM construct that they believed each item was assessing. The available components for selection included (1) *autonomy*, (2) *identity*, (3) *work orientation*, (4) *temperance* or (5) *perspective*.

This structured evaluation process enabled the experts to provide their valuable assessments, contributing to a comprehensive appraisal of the PMTAFC items.

Data Analysis

To evaluate the representativeness of each item within the PM construct, we have employed the Item-Content Validity Index (I-CVI) and the Scale-Content Validity Index (S-CVI), as outlined by Lynn (1986). The Content Validity Index (CVI) stands as one of the most widely used tools for assessing content validity, determining whether individual items (I-CVI) and the instrument as a whole (S-CVI) accurately represent the construct (Yang & Chang, 2008). A minimum threshold value of (I-CVI $\geq$ .78) was established, signifying excellent construct representativeness based on expert input (Polit et al., 2007). It's important to note that the CVI can be affected by the number of reviewers, with an increase in reviewers potentially leading to a decrease in the CVI (Rubio et al., 2003).

The S-CVI plays a pivotal role in enhancing the construct validity of an instrument. In this study, we set a minimum threshold value of (S-CVI $\geq$ .80), denoting excellent scale content validity (Polit et al., 2007).

Factorial Validity Index (FVI)

The Factorial Validity Index (FVI) works as a valuable tool for the initial quantification of factorial validity, in line with the methodology described by Rubio et al. (2003). It is instrumental in determining the extent to which experts have appropriately assigned items to the correct component of the PM construct, guided by well-established theoretical criteria (Cauffman & Steinberg, 2000; Morales-Vives et al., 2013; Steinberg & Cauffman, 1996; Strathman et al., 1994; Weinberger & Schwartz, 1990). The calculation of this index involves dividing the number of experts who accurately associated each item with its designated component by the total number of expert respondents, following the approach outlined by Rubio et al. (2003).

One noteworthy aspect of the FVI is that, as a relatively new index, there is no existing criterion to determine the ideal level of attainment (Rubio et al., 2003). In this study, we established a minimum threshold value of (FVI ≥ .70). This decision was made with consideration for the impact of the number of experts on the likelihood of agreement among them, known that a higher number of experts may lead to greater variance in assessments. In alignment with the lack of established criteria for this index, this threshold was set to ensure a reasonable standard of agreement.

Results

Item Content Validity Index (I-CVI)

Our analysis of the Item Content Validity Index (I-CVI) yielded interesting insights based on expert assessments within each of the three groups. In Group 1, a substantial 89% of the items were deemed representative of the PM construct (I-CVI ≥ .78). Within this group, 50% of the items achieved a notably high level of representativeness (I-CVI ≥ .90), with a remarkable 26% of items obtaining complete consensus among the experts (I-CVI = 1.00).

Group 2 exhibited a similar trend, with 92% of the items meeting the criteria for construct representativeness (I-CVI ≥ .78). Among these items, 50% reached an exceptional level of relevance (I-CVI ≥ .90) and 32% achieved unanimous expert endorsement (I-CVI = 1.00).

In contrast, Group 3 presented a slightly lower percentage, with 61% of the items considered construct representative (I-CVI ≥ .78). Nevertheless, this group stood out with a remarkable 53% of items receiving unanimous approval from the experts (I-CVI = 1.00), marking it as the group with the most extreme scores.

When aggregating the assessments from all three expert groups, a substantial 92% of the items demonstrated construct representativeness (I-CVI ≥ .78). Among these, 45% achieved a very high level of relevance (I-CVI ≥ .90), with 21% of the items securing unanimous expert consensus (I-CVI = 1.00) (Table 1).

Scale Content Validity Index (S-CVI)

The combined assessment by all three expert groups resulted in an overall Scale Content Validity Index (S-CVI) of .89. This unified index reflects a shared agreement among experts that the questionnaire effectively covers the various aspects of the PM construct.

Remarkably, Group 2 demonstrated strong consensus, contributing to a robust S-CVI of .92. This underscores their collective belief in the questionnaire's ability to comprehensively represent the intricate facets of PM. In contrast, Group 1 and Group 3 displayed similar levels of agreement, achieving a commendable S-CVI of .89. This confirms the questionnaire's overall

Table 1
Percentage of items classified by range of I-CVI values for each group.

I-CVI	All groups	Group 1	Group 2	Group 3
= 1	21%	26%	32%	53%
.90 - .99	24%	24%	18%	0%
≥ .90	45%	50%	50%	53%
≥ .78	92%	89%	92%	61%
< .78	8%	11%	8%	39%

Note. I-CVI = Item-content validity index.

suitability for portraying the complexities of the PM construct.

These findings reiterate the robustness of the PM questionnaire, serving as a dependable and valid tool to evaluate PM from the vantage points of diverse expert cohorts. The collective S-CVI underscores the questionnaire's potential to encapsulate the multifaceted dimensions of the PM construct, which offers a valuable instrument for research and practical applications within the field.

Factorial Validity Index (FVI)

The Factorial Validity Index (FVI) plays a pivotal role in assessing the questionnaire's ability to correctly link items with the various components of the PM construct. The results, as displayed in Table 2, offer a comprehensive overview of these associations among the three expert groups.

In Group 1, an impressive 84% of items were correctly linked to their respective components of the PM construct (FVI ≥ .70). Group 2, while still proficient, associated 76% of items correctly (FVI ≥ .70). Group 3, though slightly lower in accuracy, successfully connected 58% of items with their correct components (FVI ≥ .70). When we consider the evaluations from all three groups, a substantial 79% of items were accurately associated with the PM construct's components (FVI ≥ .70). Additionally, about 26% of the items achieved an even higher level of consensus, with an FVI of ≥ .90, signifying a strong alignment among experts.

To gain a more detailed insight into these associations and their connection to specific PM components, Table 3 presents an in-depth breakdown of both the CVI and FVI scores at the item level. This categorization is organized in a descending order, providing a closer examination of how each item aligns with the various components of the PM construct.

Table 2
Percentage of items classified by range of FVI values for every group.

FVI	All groups	Group 1	Group 2	Group 3
= 1	5%	16%	23%	24%
.90 - .99	21%	24%	13%	0%
≥ .90	26%	26%	26%	18%
≥ .70	79%	84%	76%	58%
< .70	21%	16%	24%	42%

Note. FVI: Factorial validity index.

Table 3
PMTAFC structure and item- and scale-content validity indexes at the item-level.

Items	CVI	FVI
RESPONSIBILITY (Responsabilidad)		
Autonomy (autonomia)		
Me gusta tomar mis propias decisiones.[a]	100.00%	97.50%
Necesito consultar con mis amigos antes de tomar una decisión.[a]	95.12%	87.18%
Creo que mis decisiones son incorrectas cuando a mis amigos no les gustan.[a]	95.12%	84.62%
Antes de comprarme ropa u otras cosas para mí, consulto con mis amigos.[a]	90.00%	83.33%
Me siento incómodo cuando mi opinión es diferente a la de mis amigos.[a]	87.80%	80.56%
Identity (identidad)		
Creo que me conozco bastante bien.[a]	100.00%	100.00%
Tengo claro lo que me interesa.[a]	100.00%	60.98%
Siento que los demás me valora y me aceptan.[a]	87.50%	88.57%
Soy capaz de hacer muchas cosas bien.[a]	87.18%	76.47%
Muchas veces pretendo ser alguien que no soy.[a]	85.37%	97.14%
Siento que mi vida no tiene mucho sentido.[a]	73.32%	93.33%
Work Orientation (Orientación al trabajo)		
Siempre termino mis deberes y responsabilidades antes de dedicarme a las actividades que me gustan (videojuegos, ver amigos, hacer deporte, etc).[a]	100.00%	92.50%
Siempre hago lo que toca y cumplo con mis obligaciones.[a]	100.00%	90.24%
Casi nunca me retraso en cumplir con mis obligaciones.[a]	100.00%	87.80%
Me esfuerzo por conseguir buenos resultados, aunque sean a largo plazo.[c]	100.00%	78.05%
Cuando una tarea me requiere mucho esfuerzo o tiempo, me cuesta acabarla.[a]	85.37%	80.56%
Paso de una cosa a otra sin acabar ninguna. [a]	85.37%	74.29%
Si ahora no me esfuerzo lo suficiente, ya lo arreglare más adelante.[c]	80.49%	57.58%
TEMPERANCE (Templanza)		
Generalmente soy una persona controlada y no suelo perder los nervios.[b]	97.57%	97.50%
Aunque alguien me haga daño, no intento vengarme.[b]	90.00%	83.33%
Me porto bien, incluso con las personas que no me gusta.[b]	86.84%	57.58%
Cuando me enfado me dejo llevar, sin importarme nadie ni nada.[b]	85.37%	94.29%
Solo pienso en el resultado inmediato de mis acciones.[c]	85.37%	57.89%
Intento ser agradable, incluso con alguien que me ponga nervioso y me irrite.[b]	82.92%	82.35%
Trato bien a la gente, incluso la que no me cae bien.[b]	82.93%	53.00%
El que me haga enfadar debería tener cuidado conmigo.[b]	78.05%	96.88%
PERSPECTIVE (Perspectiva)		
Casi todas las cosas se pueden ver desde dos puntos de vista y siempre intento considerar ambos.[d]	100.00%	95.00%

Items	CVI	FVI
Cuando tomo una decisión pienso en cómo me podría afectar en el futuro.[c]	100.00%	92.86%
Intento que todo lo que hago ahora me sirva para el futuro.[c]	100.00%	80.49%
Solo atiendo a mis preocupaciones actuales, los problemas futuros ya se arreglarán.[c]	85.37%	80%
Para entender mejor a mis amigos me imagino cómo ven las cosas desde su perspectiva.[d]	92.68%	89.47%
Me preocupo más por los problemas de cada día que de los que puedan venir en el futuro.[c]	90.24%	72.97%
Pensar ahora en los problemas que me traerá el futuro no sirve de nada, es perder el tiempo.[c]	87.80%	83.33%
Muchas veces hago cosas que tardan mucho tiempo en dar resultados.[c]	87.80%	63.89%
Evito herir los sentimientos de los demás.[b]	87.80%	58.33%
Me cuesta ver las cosas desde el punto de vista de los demás.[d]	85.37%	88.57%
No soy de los que ayudan a los demás.[b]	67.50%	65.38%
No suelo hacer esfuerzos por los demás.[b]	62.50%	76.00%

Note. Item sources: a PSYMAS (Morales-Vives et al., 2013); b WAI (Weinberger & Schwartz, 1990); c CFC (Strathman et al., 1994); d IRI (Davis; 1980). The English version can be found here.

Discussion

In the domain of criminal justice, the process of desisting from criminal behavior during the transition from late adolescence to early adulthood has long piqued the curiosity of scholars and practitioners (Monahan et al., 2009; Rocque et al., 2019). This critical juncture can lead individuals to diverge from a life of crime or continue along a perilous path. It's in this context that the concept of PM has surfaced as a beacon of understanding and potential in the realm of juvenile justice.

Think of PM as a versatile tool, one that not only enhances diagnostic precision but also facilitates the creation of tailored interventions for adolescents navigating the complexities of maturity. It can serve as a guiding compass for judges, illuminating the path to fair and informed decisions. Moreover, it acts as a protective shield against the potential negative effects of incarceration, ensuring that it doesn't hinder psychosocial maturation but rather supports it (Cavanagh, 2022; Lambie & Randell, 2013).

This paper embarks on a journey of exploration, which delves into the content validity of a new measurement instrument comprising 38 items, each with the potential to assess PM. Our mission has two dimensions: O1). Construct representativeness: to assess the extent to which these items genuinely capture the essence of the PM construct, and O2). Component representativeness: to scrutinize their alignment with the specific components that define PM.

This novel study seeks to shed light on the transformative potential of the PM construct within the landscape of juvenile justice. Its implications go beyond the academic realm; they hold the power to make a real-world impact, shaping the future of justice for our youth.

Construct representativeness (I-CVI and S-CVI)

The findings regarding construct representativeness (I-CVI and S-CVI) are a resounding affirmation of the content validity of these items. The consensus among the three expert groups is clear: the majority of items and the scale as a whole are a robust representation of the PM construct. Items that received a content validity index (I-CVI) of

.78 or higher, are deemed as strong indicators of content validity (Polit et al., 2007). The interesting part is that when we compared the responses of the three groups, we noticed a remarkable degree of agreement between Groups 1 and 2. They were mostly on the same page when it came to the representativeness of the items. Group 3, however, provided more varied responses. A significant number of items received unanimous endorsement from experts (I-CVI = 1), while others scored lower.

The divergence in responses among the groups can be attributed to the variance in information provided to each. Groups 1 and 2, armed with a more profound comprehension of the PM construct, demonstrated heightened consensus on the relevance of items. Conversely, Group 3, operating with comparatively less contextual information, adopted a more cautious stance in assessing item relevance. This underscores the pivotal role of clear and comprehensive definitions, particularly for experts utilizing the assessment tool.

Component representativeness (FVI)

When considering the alignment of items with the specific components of the PM construct (FVI), a more nuanced narrative emerges compared to the robust CVI outcomes. However, given the substantial number of expert assessments, it is noteworthy that approximately three-quarters of the items, as evaluated across all groups, demonstrated accurate alignment with their respective components. This outcome is deemed satisfactory, taking into account the collective perspectives of experts.

An analysis of the results on a group-by-group basis unveils some intriguing patterns. Group 1, for instance, showcased a robust alignment, with over four-fifths of the items correctly associated with their respective components. Group 2 exhibited a commendable level of alignment, with roughly three-quarters of items accurately matched. Otherwise , Group 3 appeared to demonstrate a somewhat lower level of precision, with only around three-fifths of the items correctly linked to their components.

When aggregating the results across the groups, it is evident that more than three-quarters of the items harmoniously matched with their designated components. While this outcome may be viewed favorably, especially considering the substantial number of experts involved, several critical factors merit discussion, particularly regarding Group 3's performance.

To begin with, it should be noted that the definitions provided for each component, albeit consistent with the original author's terminology (Steinberg & Cauffman, 1996), may have lacked the depth necessary for experts to effectively discriminate between them in the context of Task 2 (O2. Represented component). It is evident that meticulously crafted definitions can substantially enhance the capacity of experts to discern between the various components. The superior FVI results observed in Groups 1 and 2 could be attributed to the additional contextual information and training pertaining to the PM construct that they received, emphasizing the pivotal role of specialized training for both experts and prospective users who will administer the measure.

Moreover, the inherent similarity shared among the components of the construct poses a formidable challenge when seeking items that exclusively represent each distinct facet. For instance, take into consideration the following item of our scale, "Often I engage in a particular behavior in order to achieve outcomes that may not result for many years" [*Me esfuerzo por conseguir buenos resultados, aunque sean a largo plazo*] (CFC, Strathman et al., 1994), which delves into

the attainment of long-term outcomes. This item encompasses elements of perspective by evaluating long-term consequences, ventures into the domain of *temperance* by probing the capacity to defer immediate gratification, and even extends into the realm of *work orientation*, which constitutes a sub-component of *responsibility*, by assessing the sense of pride derived from task accomplishment. Items that garnered relatively lower FVI scores are currently undergoing refinement to yield a more precise reflection of each individual component.

Lastly, it is imperative to acknowledge that the PM construct's interdependent components have not evolved in a uniform manner. Among these components, *responsibility* stands out as it features well-defined sub-components, such as autonomy, identity and work orientation. To facilitate a more equitable and comprehensive measurement, there may arise a necessity to establish specific sub-components for *temperance* and *perspective*. This could facilitate the realization of more uniformly structured measurement categories. The operative definitions of PM, as currently implemented, may not furnish the level of comprehensive assessment essential to holistically encapsulate the PM construct, echoing the earlier insights posited by Cauffman and Steinberg (2000).

In short, despite the overwhelming majority of the 38 items manifesting relevance to the PM construct, items that exhibited diminished performance in terms of CVI and FVI will undergo meticulous revision or potential elimination. In anticipation, a confirmatory factor analysis will be performed as a net step for this PMTAFC, promising insights into the organizational dynamics of these components, thereby contributing to the continued maturation of the PM construct's theoretical model.

Conclusions

The concept of PM is pivotal within the context of managing adolescents involved in legal proceedings and facilities. To be effective in the forensic and judicial context, PM needs to be accurately framed and assessed. The PM construct offers a promising framework, honing in on adolescents' decision-making capabilities and their influence on behavior. The validation of a PM assessment tool holds substantial value for adolescents navigating legal procedures within the juvenile justice system. It not only stands to enhance the precision of court decisions but also facilitates tailored clinical interventions.

Drawing from the findings of this study, it is evident that a substantial portion of the 38 items proposed for a valid self-report assessment of PM in adolescents successfully aligns with the core PM construct and its individual components. This achievement is a significant step toward the creation of a comprehensive measurement instrument. In the subsequent phase of our study, we will embark on a confirmatory factor analysis to delve into the intricate relationships between the sub-components of the construct, furthering our quest for a nuanced understanding of PM.

The robust content validation outcomes underscore the potential transformative influence of PM assessments on the juvenile justice landscape. This journey holds the promise of empowering professionals, streamlining interventions, and paving the way for more equitable and well-informed decisions within the realm of adolescent legal proceedings. As we move forward in this mission, the anticipation for the positive change that validated PM assessment tools can bring to the lives of young individuals within the justice system remains high.

References

Almanasreh, E., Moles, R., & Chen, T. F. (2019). Evaluation of methods used for estimating content validity. *Research in Social and Administrative Pharmacy, 15*(2), 214-221. https://doi.org/10.1016/j.sapharm.2018.03.066

Belsky, J., Caspi, A., Moffitt, T. E., & Poulton, R. (2020). *The origins of you: How childhood shapes later life*. Harvard University Press.

Cauffman, E., & Steinberg, L. (2000). (Im)maturity of judgment in adolescence: Why adolescents may be less culpable than adults. *Behavioral Sciences & the Law, 18*(6), 741-760. https://doi.org/10.1002/bsl.416

Cauffman, E., Piquero, A. R., Kimonis, E., Steinberg, L., Chassin, L., & Fagan, J. (2007). Legal, individual, and environmental predictors of court disposition in a sample of serious adolescent offenders. *Law and Human Behavior, 31*(6), 519-535. https://doi.org/10.1007/s10979-006-9076-2

Cavanagh, C. (2022). Healthy adolescent development and the juvenile justice system: Challenges and solutions. *Child Development Perspectives, 16*(3), 141-147. https://doi.org/10.1111/cdep.12461

Davis, M. H. (1980). *Interpersonal Reactivity Index (IRI)* [Database record]. APA PsycTests. https://doi.org/10.1037/t01093-000

Expósito, C. D., Marsollier, R. G., Difabio, H. E., & Castro-Santander, A. (2023). Construcción y validación del Cuestionario de Acoso y Ciberacoso Escolar (CACE) mediante juicio de expertos. *Revista Evaluar, 23*(1), 61-79. https://doi.org/10.35670/1667-4545.v23.n1.41014

Greenberger, E., Josselson, R., Knerr, C., & Knerr, B. (1975). The measurement and structure of psychosocial maturity. *Journal of Youth and Adolescence, 4*(2), 127-143. https://doi.org/10.1007/BF01537437

Lambie, I., & Randell, I. (2013). The impact of incarceration on juvenile offenders. *Clinical Psychology Review, 33*(3), 448-459. https://doi.org/10.1016/j.cpr.2013.01.007

Lynn, M. R. (1986). Determination and quantification of content validity. *Nursing Research, 35*(6), 382-386. https://doi.org/10.1097/00006199-198611000-00017

Moffitt, T. E. (1993). The neuropsychology of conduct disorder. Development and *Psychopathology, 5* (1-2), 135-151. https://doi:10.1017/S0954579400004302

Moffitt, T. E. (2003). Life-course persistent and adolescence-limited antisocial behavior. In B. B. Lahey, T. E. Moffitt, & A. Caspi (Eds.), *Causes of Conduct Disorder and Juvenile Delinquency* (pp. 49-75). The Guilford Press.

Monahan, K. C., Steinberg, L., Cauffman, E., & Mulvey, E. P. (2009). Trajectories of antisocial behavior and psychosocial maturity from adolescence to young adulthood. *Developmental Psychology, 45*(6), 1654-1668. https://doi.org/10.1037/a0015862

Morales-Vives, F., Camps, E., & Lorenzo-Seva, U. (2013). Development and validation of the Psychological Maturity Assessment Scale (PSYMAS). *European Journal of Psychological Assessment, 29*(1), 12-18. https://doi.org/10.1027/1015-5759/a000115

Narváez-Burbano, J. H., & Obando-Guerrero, L. M. (2023). Construcción y validación de la Escala de Exposición a Factores de Deprivación Sociocultural (EXFADESO). *Revista Evaluar, 23*(1), 27-39. https://doi.org/10.35670/1667-4545.v23.n1.41006

Ozkan, T., & Worrall, J. L. (2017). A psychosocial test of the Maturity Gap Thesis. *Criminal Justice and Behavior, 44*(6), 815-842. https://doi.org/10.1177/0093854817694924

Pailing, A. N., & Reniers, R. L. E. P. (2018). Depressive and socially anxious symptoms, psychosocial maturity, and risk perception: Associations with risk-taking behaviour. *PLOS ONE, 13*(8), e0202423. https://doi.org/10.1371/journal.pone.0202423

Polit, D. F., Beck, C. T., & Owen, S. V. (2007). Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Research in Nursing & Health, 30*(4), 459-467. https://doi.org/10.1002/nur.20199

Riggs-Romaine, C. L. (2018). Psychosocial maturi-

ty and risk-taking in emerging adults: Extending our understanding beyond delinquency. *Emerging Adulthood, 7(*4), 243-257. https://doi.org/10.1177/2167696818768013

Rocque, M., Beckley, A. L., & Piquero, A. R. (2019). Psychosocial maturation, race, and desistance from crime. *Journal of Youth and Adolescence, 48*(7), 1403-1417. https://doi.org/10.1007/s10964-019-01029-8

Rubio, D. M., Berg-Weger, M., Tebb, S. S., Lee, E. S., & Rauch, S. (2003). Objectifying content validity: Conducting a content validity study in social work research. *Social Work Research, 27*(2), 94-104. https://doi.org/10.1093/swr/27.2.94

Steinberg, L., Cauffman, E., & Monahan, K. C. (2015). Psychosocial maturity and desistance from crime in a sample of serious juvenile offenders. *Juvenile Justice Bulletin. US Department of Justice, Office of Justice Programs, Office of Juvenile Justice and Delinquency Prevention*. https://ojjdp.ojp.gov

Steinberg, L., & Cauffman, E. (1996). Maturity of judgment in adolescence: Psychosocial factors in adolescent decision making. *Law and Human Behavior, 20*(3), 249-272. https://doi.org/10.1007/BF01499023

Strathman, A., Gleicher, F., Boninger, D. S., & Edwards, C. S. (1994). The consideration of future consequences: Weighing immediate and distant outcomes of behavior. *Journal of Personality and Social Psychology, 66*(4), 742-752. https://doi.org/10.1037/0022-3514.66.4.742

Wakeling, H., & Barnett, G. (2017). *Development and validation of a screening assessment of psychosocial maturity for adult males convicted of crime* (Analytical summary). HM Prison & Probation Service. https://www.gov.uk

Weinberger, D. A., & Schwartz, G. E. (1990). Distress and restraint as superordinate dimensions of self-reported adjustment: A typological perspective. *Journal of Personality, 58*(2), 381-417. https://doi.org/10.1111/j.1467-6494.1990.tb00235.x

Wenger, L., & Andres-Pueyo, A. (2016). Personality and clinical tests in Spanish for assessing juvenile offenders. *Papeles del Psicólogo, 37*(2), 89-106. https://psycnet.apa.org/record/2016-31958-002

Wenger-Amengual, L. S. (2018). *Comportamiento antisocial, personalidad y madurez en adolescentes y jóvenes* (Doctoral dissertation). University of Barcelona. http://hdl.handle.net/2445/127229

Yang, Y.-T. C., & Chan, C.-Y. (2008). Comprehensive evaluation criteria for English learning websites using expert validity surveys. *Computers & Education, 51*(1), 403-422. https://doi.org/10.1016/j.compedu.2007.05.011

2024, Vol. 24, No. 1
ISSN 1667-4545
Recuperado de https://revistas.unc.edu.ar/index.php/revaluar

Laboratorio de Evaluación Psicológica y Educativa
Facultad de Psicología - Universidad Nacional de Córdoba

Structural Analysis of the Academic Motivation Scale (Spanish version) in Graduate Students

Análisis Estructural de la Escala de Motivación Educacional (versión en español) en Estudiantes Graduados

Juan Aníbal González-Rivera * [1], Ernesto Rosario-Hernández [2], Taimara Ortiz Santiago [3]

1 - Ponce Health Sciences University, Ponce, Puerto Rico. https://orcid.org/0000-0003-0622-8308
2 - Ponce Health Sciences University, Ponce, Puerto Rico. https://orcid.org/0000-0002-2523-6162
3 - Universidad Carlos Albizu, San Juan, Puerto Rico. https://orcid.org/0000-0002-3944-4835

Recibido: 24/08/2023 **Revisado:** 25/11/2023 **Aceptado:** 02/12/2023

Abstract

This research aimed to examine the factor structure of the Academic Motivation Scale (AMS) in master's and doctoral students from universities in Puerto Rico; 300 students between 21 to 40 years (M = 29.14; SD = 4.87) participated. Confirmatory factor analysis, internal consistency, correlation, and item analysis were performed. Results of the current study provide evidence that supports the internal structure of the AMS and the ancillary statistics use of the bifactor model presents some interesting information about the possible unidimensional or multidimensional uses of the AMS. The subscales of the AMS obtained good reliability coefficients, and the AMS appears to be invariant among gender and age, which permits comparison among these groups. The use of the AMS appears useful in the educational context with graduate students in Puerto Rico. The implications and limitations of the findings are discussed.

Keywords: *academic motivation, extrinsic motivation, intrinsic motivation, graduate students, psychometrics*

Resumen

Esta investigación se propuso examinar la estructura factorial de la Escala de Motivación Académica (EMA) en estudiantes de maestría y doctorado de universidades de Puerto Rico. Participaron 300 estudiantes con edades que fluctuaron entre 21 a 40 años (M = 29.14; SD = 4.87). Se realizaron análisis factoriales confirmatorios, análisis de consistencia interna, análisis de correlación y análisis de ítems. Los resultados aportan pruebas que respaldan la estructura interna de la EMA y el uso estadístico auxiliar del modelo bifactorial presenta alguna información interesante sobre los posibles usos unidimensionales o multidimensionales de la EMA. Las subescalas de la EMA obtuvieron buenos coeficientes de fiabilidad y la EMA parece ser invariante entre el género y la edad, lo que permite realizar comparaciones entre estos grupos. El uso de la EMA parece ser útil en el contexto educativo con estudiantes graduados en Puerto Rico. Se discuten las implicancias y limitaciones de los hallazgos.

Palabras clave: *motivación académica, motivación extrínseca, motivación intrínseca, estudiantes graduados, propiedades psicométricas*

***Correspondence to:** Dr. Juan Aníbal González-Rivera. E-mail: jagonzalez@psm.edu

How to cite: González-Rivera, A., J., Rosario-Hernández, E., & Ortiz Santiago. T., (2024): Structural Analysis of the Academic Motivation Scale (Spanish version) in Graduate Students. *Revista Evaluar, 24*(1), 28-43. Retrieved from https://revistas.unc.edu.ar/index.php/revaluar

Participaron en la edición de este artículo: Stefano Macri, Juan Balverdi Nieto, Abigail Pérez, Florencia Ruiz, Jorge Bruera.

Introduction

The guidance that university students receive from their professors has become a dynamic process that includes academic, personal, and professional areas; individually oriented to encourage students to create their way of working as independent professionals (Rodríguez et al., 2009). In the search to know the importance and the impact on the motivational variables in the academic performance, the literature reflects that self-efficacy strategies contribute to academic performance, organization, understanding and support strategies, and goals in learning (Becerra-González & Reidl-Martínez., 2015). This same author mentions that students with good academic performance recognize that the skills of the teacher and the didactic strategies for the development of the curriculum and institutional policies are not determining factors in academic performance. However, school performance, academic attributional style, motivation to school achievement, and academic self-efficacy have been related as determining factors (Becerra-González & Reidl-Martínez, 2015). Gutiérrez mentions that no statistically significant differences have been found between students' academic performance, intrinsic motivation, and self-efficacy. Nevertheless, positive correlations were found between cognitive strategies and subfactors of self-efficacy in general and motivation (Gutiérrez-Ruiz, 2015).

Motivation is one of the variables most frequently studied in the educational area and is also known as a multidimensional construct (Guzmán et al., 2006). It defines the internal and external components that promote the appearance of some behaviors (Candela et al., 2014). Moving to the educational area and considering the behavior of the human being, it is indisputable that the attitudes, perceptions, expectations, and ideas that the student has of himself, of the fulfillment of subjects, and of the objectives that he wishes to fulfill compose factors that offer direction. They lead student behavior in the school area. In any case, the external variables that come from the context in which the students operate must be considered, which is an aspect that influences them (Sánchez & Hernández-Pina, 2011).

Motivation has two variants: intrinsic motivation and extrinsic motivation; in teaching settings, emphasis is placed on differentiating one from the other on many occasions. Intrinsic motivation is related to behaviors for the interest of the activity itself, thinking only about its end and not as a way to achieve goals. Moreover, intrinsic motivation arises in people who are interested in learning, knowing and explaining phenomena. How obstacles are overcome is one of the most critical aspects of intrinsic motivation and is connected to the learning approach and competitive self-improvement of a high level of performance (Román-Pérez, 2013). On the other hand, extrinsic motivation is considered when the person performs an activity to satisfy other reasons unrelated to the activity itself (Ruiz, 2015). Motivation is an aspect that induces the way people think, and as a result, learning is affected (Alonso-Tapia, 1998). A student who performs activities out of interest, intrinsically motivated, is more receptive to make significant mental efforts while performing the activity, by using effective learning strategies (Lepper, 1988).

Unlike intrinsic motivation, extrinsic motivation leads to the accomplishment of the task with external aspects. The emotions that arise from the extrinsically motivated person occur through the expected results after performing the task (Sánchez & Hernández-Pina, 2011). In the educational area, Skinner justified using reinforcement and recognizing appropriate behavior; he understood that these strategies were more effective than applying punishment to modify the

behavior to be achieved. Over the years, studies on learning continued, and it was shown that applying reinforcement and recognition promotes superficial learning in which students perform tasks without showing genuine interest. The objective of performing them was to avoid punishment or receive some positive reinforcement (Sánchez & Hernández-Pina, 2011). This suggests that there is a possibility that the student, who is motivated in an extrinsic way, shows commitment in the academic area only when they have the opportunity of some reward. They are likely in search of more straightforward tasks to ensure success rewards (Lepper, 1988).

As found in the literature, in a study on the sense of self-efficacy in graduate students, it turned out that intrinsic motivation was the most influential personal factor in self-efficacy to complete master's studies (Reyes-Cruz & Gutiérrez-Arceo, 2015). On the other hand, in an investigation of self-regulation in university students: learning strategies, motivation, and emotions, which aimed to evaluate emotional incidents and motivational beliefs according to the results of the tasks performed, it turned out that significant differences were found depending on the emotion and subjective competence beliefs used by the student. Another component of motivation is amotivation. This is considered the lowest level of autonomy in the different types of motivation; it is identified when the person does not perceive causality between the behavior and the consequences and does not feel competent in obtaining the desired goal (Deci & Ryan, 2000).

Theoretical framework

The theory of self-determination tells us that in order to achieve a better understanding of human motivation; it is necessary to consider the innate psychological needs of individuals for competence, autonomy and relationships. These needs are necessary for continued psychological growth, integrity, and well-being (Zhang et al., 2005). These authors establish that human needs indicate the necessary conditions for psychological well-being, and therefore, their satisfaction could be associated with satisfactory levels of motivation. Satisfaction of the three needs is necessary since it has been shown that it is not enough to satisfy some of them. By satisfying these innate needs, intrinsic motivation is maintained or improved, while the internalization and integration of extrinsic motivation are facilitated. On the contrary, some of the frustrated satisfaction can be reflected in unfavorable functions towards persistence and performance. In addition, it is associated with less intrinsic motivation and more intense extrinsic aspirations, which impact experiences, performance, and decreased well-being (Zhang et al., 2005).

Academic Motivation Scale

The Academic Motivation Scale, validated by Núñez et al. (2010), was used to review and validate it, including the integrated regulation to measure motivation in initial teacher training (Burgueño et al., 2017). A study to discover the motivational profiles of studying medicine was accomplished. This study was completed in 3 universities in Ecuador (Chicaiza-Ayala & Cragno, 2018), while in Brazil, a study in which integrated regulation was added to the Educational Motivational Scale was performed, and the psychometric properties were observed (Silva et al., 2018). On the other hand, Núñez et al. in 2010 carried out an adaptation and validation of the version of the Academic Motivation Scale in post-compulsory secondary education students.

After identifying that the instruments for analyzing the types of motivation within the self-determination variable were limited, Vallerand et al. (1993) developed and validated *Echelle de Motivation en Education* in French. This instrument is designed with 28 items and, in turn, divided into seven subscales of four items. This scale is subdivided into three areas of motivation: extrinsic, intrinsic and amotivation. This validation study had satisfactory levels of internal consistency, with a Cronbach's alpha of .80. Later this scale was translated into English and resulted in the Academic Motivation Scale (AMS). This procedure was carried out with a sample of students from Canada. The validation study results yielded a Cronbach's alpha of .62 and .86, reflecting good internal consistency. Later, another study was performed and good levels of construct validity were obtained in the correlation analysis between the seven subscales (Vallerand et al., 1993).

However, the scale was translated into Spanish and subjected to psychometric analysis by Núñez et al. (2010). The AMS is made up of 28 self-report items that answer the question *Why are you going to university?* with Likert-type responses ranging from *Does not correspond at all* (1) to *totally corresponds* (7), with an intermediate score, *It corresponds moderately* (4). Some of the questions say: *Because for me it is a pleasure and satisfaction to learn new things; I don't know why I'm going to college and, frankly, I don't care*; and *Because I want to show myself that I am capable of succeeding in my studies*. The scale has internal consistency in its seven subscales (Amotivation, External Regulation, Introjected Regulation, Identified Regulation, Intrinsic Motivation to knowledge, Intrinsic Motivation to achievement and Intrinsic Motivation to stimulating experiences) with values between .76 and .84, except in the Regulation subscale identified with .67.

Research Purpose

This study aims to validate and adopt the Academic Motivation Scale (AMS) in Puerto Rico. At the moment, in Puerto Rico, no scale measures academic motivation. The AMS was initially validated in French (Vallerand, 1993) with a Cronbach alpha .80, then in English (Vallerand et al., 1992) with a Cronbach alpha .62 and .86, and finally in Spanish (Núñez et al., 2010) with a Cronbach alpha between. 67 and .84. In the studies that have been carried out previously, this instrument has proven to be adequate to evaluate motivation in the educational area; for this reason, carrying out validation for Puerto Rico is proposed. The construct validity and internal consistency of the AMS will be evaluated to benefit future researchers interested in the educational area in the Puerto Rican population.

Method

Research Design and Procedures

Following the classifications of Ato et al. (2013), this research was framed within the non-experimental model, under an instrumental model. The Institutional Review Board of Carlos Albizu University in San Juan, Puerto Rico, approved the research project. Data collection was made online (internet), and volunteers were recruited by propagating a paid ad on social media, directing them to informed consent and the survey. The consent included the purpose of the study, the inclusion criteria, the voluntary nature of the study, the possible risks, and benefits, as well as the right to withdraw from the study at any time.

Participants

A total of 300 students at the graduate level participated in the present investigation with ages ranging from 21 to 40 years, with a mean age equal to 29.14 and a standard deviation equal to 4.87. Table 1 shows that most of the students were female (77.3%), 65.0% were studying for a master's degree, 72.3% of the sample was between the first and third year of study, 86.7% reported having an academic average between 3.50 to 4.00, and 53.7% indicated that they studied daily between one to three hours.

Measurement

Academic Motivation Scale, Spanish version (AMS; Núñez et al., 2006). The scale contains seven subscales with four items each for a total of 28 items. Of these subscales, there are three that evaluate intrinsic motivation (IM): IM to knowledge, IM to achievement and IM to stimulating experiences. At the same time, three subscales measure extrinsic motivation (EM): external regulation, introjected regulation, and identified regulation. Finally, there is the subscale that measures amotivation. All the items answer the question, *Why are you going to college?* and the scale is anchored on a seven-point Likert scale from 1 (*Does not correspond at all*) to 7 (*It totally corresponds*), with an intermediate score of 4 (*It moderately corresponds*). The confirmatory factor analysis results using the modeling of structural equations found by Alonso et al. (2006) support the internal structure of seven factors and the correlations between the subscales with academic self-concept support the construct validity of the scale. The reliability of the subscales, according to Núñez et al. (2010), were acceptable since they fluctuated between .68 to .79 using Cronbach's alpha technique.

Data Analysis

The data were analyzed, first, with the IBM-SPSS version 28.0 program, and descriptive statistics, correlation, reliability, and item analysis were performed with it. In addition, we performed several confirmatory factors analyses

Table 1
Sociodemographic data of the study participants. (n = 300).

Variable	Frequency	Percent	Variable	Frequency	Percent
Gender			**Academic Year**		
Female	232	77.3	First	80	26.7
Male	67	22.3	Second	85	28.3
Program			Third	52	17.3
Master	195	65.0	Fourth	26	8.7
Doctorate	105	35.0	=>Fifth	57	19.0
GPA			**Hours Dedicated Daily to Studying**		
3.50 - 4.00	260	86.7	1-3	161	53.7
3.00 - 3.49	34	11.3	3-5	91	30.3
2.50 - 2.99	4	1.3	=>5	48	16.0
2.00 - 2.49	2	.7			

and invariance testing by gender and age of the Academic Motivation Scale using the weighted least squares-mean and variance adjusted (WLSMV) estimator with the "lavaan" package of the R3.6.3 program (Rosseel, 2012), which robustly deals with potentially non-normal data with items treated as ordinal (Li, 2016a, 2016b). To evaluate the results of the CFA, several fit indices of the structural equation models were used. Kline (2016) recommends the use of at least four fit indices, although more can be reported. One of the indices that is reported is Chi-Square (χ^2). This is a fundamental index of absolute adjustment and it is basically the same one that is used when the researcher wants to examine the association between nominal variables. However, the crucial difference when used as an index of fit in the structural equations model is that the researcher looks for no differences between the matrices to support that the tested model is representative of the data (Hair et al., 2019). Given that the χ^2 is sensitive to the sample size and therefore the probability of rejecting the hypothesized model increases when the sample size increases, it is recommended to take into account other indices (Marsh et al., 1996). In this way, the Root Mean Square Error of Approximation (RMSEA; Byrne, 2016; Hu & Bentler, 1999) was used, values ranging from .08 to .10 are considered as mediocres, less than .08 for the RMSEA indicate an acceptable fit, while values equal to .05 or less indicate a good fit of the model (Browne & Cudeck, 1989; MacCallum et al., 1996).

In addition, Standardized Square Root Mean Residual (SRMR; Hu & Bentler, 1999) was used, which examines the average difference between predicted and observed variances and covariances, based on the residual standard error. The lower the SRMR, the better the fit of the model; and to consider it an acceptable model it must be or close to .08, but it is preferred to be equal to or less than .05 (Hu & Bentler, 1999). On the other hand, the Bentler Comparative Fit Index (CFI) was used as an increased fit index to compare the theoretical model with the null model, which assumes that the latent variables of the model do not correlate with each other and values greater than .90 are considered acceptable (Hair et al., 2019). Another increased adjustment index is the Tucker-Lewis Index (TLI) that reflects the proportion in which the theoretical model improves the adjustment in relation to the null model (Littlewood-Zimmerman & Bernal-García, 2011; Tucker & Lewis, 1973). Values greater than .90 are considered acceptable.

Four competitive models were tested: a first model sought to examine the seven factors proposed and found by the authors of the scale (Núñez et al., 2010; Núñez et al., 2006); a second model was also evaluated in order to analyze a model of three factors, one for extrinsic motivation, another for intrinsic motivation, and a third intended to measure amotivation; a third model proposed an unifactorial internal structure in which all items loaded in just one factor. As it can be appreciated in Table 3, the best fitted model was the seven-factor model. Finally, a bifactor model was also tested which specifies that a general factor influences all items in addition to the seven specific factors. Furthermore, to examine whether there is a continuum in the general factor as a source of item variability, other indicators were applied to assess the general factor. The hierarchical omega (ω_H) refers to the amount of total variance that can be attributed to the overall factor (Flores-Kanter et al., 2018). The explained common variance (ECV) is also used, which can be interpreted as the amount of common variance of all items that is due to the overall factor. The percentage of uncontaminated correlations (PUC) are employed, which provides information on the percentage of

correlations not contaminated by multidimensionality (Rodriguez et al., 2016). Also, the percentage of uncontaminated correlations (ECV) is utilized, which is an indicator of unidimensionality; an ECV greater than .60 would indicate that there is little common variance between the factors other than the overall factor (Rodriguez et al., 2015). At the item level, the ECV-I was implemented, whose interpretation is similar to the ECV, indicating the percentage of the true variance of each item explained by the general factor. This coefficient requires that its magnitudes be greater than .60. Moreover, the average relative parameter bias (ARPB) is used, which is a measure for examining the difference between the factor loading of a unidimensional model and the general factor loading of the bifactor model (Rodriguez et al., 2016). A maximum difference of .12 to .15 may be acceptable. Finally, the factor loadings suggested by the literature (e.g., Reise et al, 2010) were examined by calculating the arithmetic means of the items. Ferrando and Lorenzo-Seva (2017) indicate that factor loading means lower than .30 in any specific factor can be considered as secondary evidence of unidimensionality.

Results

Table 2 shows the descriptive statistics of the 28 items of the AMS (mean, standard deviation, skewness and kurtosis). Item 3 had the highest mean, skewness, and kurtosis among all items, while item 21 had the highest standard deviation.

In terms of the results of the CFAs, the results can be seen in Table 3, in which it can be seen that the seven-factor model obtained the best fit indices, although the SRMR was above the threshold of .05. The bifactor model was the second best regarding fit indices and interestingly did not outperform the seven-factor model.

Table 2
Descriptive statistics of the 28 items of the Academic Motivation Scale (AMS).

# Item	Mean	SD	Skewness	Kurtosis	# Item	Mean	SD	Skewness	Kurtosis
1	4.88	2.27	-0.616	-1.115	15	5.47	1.91	-1.060	-.026
2	6.30	1.16	-1.913	3.543	16	5.67	1.63	-1.157	.520
3	6.50	1.07	-2.804	8.810	17	5.38	1.97	-1.091	-.004
4	4.28	2.05	-.145	-1.208	18	4.43	2.11	-.292	-1.250
5	1.91	1.67	1.952	2.777	19	1.58	1.50	2.748	6.473
6	6.19	1.25	-1.800	3.224	20	5.43	1.91	-1.060	-.019
7	4.96	2.33	-0.673	-1.159	21	4.39	2.36	-.282	-1.490
8	5.73	1.79	-1.458	1.155	22	5.83	1.67	-1.495	1.416
9	5.99	1.50	-1.691	2.387	23	6.13	1.41	-1858	2.954
10	6.25	1.30	-2.224	5.084	24	6.16	1.45	-2.046	3.786
11	4.14	2.05	.004	-1.245	25	4.99	1.94	-.609	-.785
12	2.58	2.13	1.080	-.327	26	1.70	1.64	2.429	4.635
13	6.15	1.40	-1.821	2.982	27	4.67	2.19	-.505	-1.137
14	4.78	2.22	-.544	-1.174	28	5.49	1.95	-1.108	.034

Note. n = 300; SD = Standard deviation.

Table 3
Fit indices obtained by the four competitive models of the Academic Motivation Scale.

Fit Index/Model	7-Factor	3-Factor	1-Factor	Bifactor
χ^2 (*df*)	707.007* (329)	2068.530* (347)	5905.536* (350)	1293.384* (322)
SRMR	.079	.126	.222	.108
RMSEA (CI)	.062 (.056 - .068)	.129 (.123 - .134)	.230 (.225 - .236)	.143 (.134 - .152)
CFI	.992	.962	.876	.979
TLI	.990	.958	.866	.975

Note. n = 300, *p < .05.

Table 4 shows the factor loadings (λ), discrimination indices (R_{bis}), reliability coefficients, and the average variance extracted (AVE) of the subscales of the AMS. All subscales obtained a reliability coefficient greater than .70 using Cronbach's alpha and McDonalds' omega with their respective confidence intervals; moreover, all subscales obtained an average variance extracted greater than .50 as recommended by some literature (Fornell & Larcker, 1981).

Table 4
Factor loadings (λ), discrimination index (r_{bis}), average variance extracted (AVE), and reliability of the items belonging to the AMS by subscale.

Subscale	Item	r_{bis}	λ	λ^2	AVE	Reliability	
						(CI)	(CI)
External Regulation	me1	.409	.438	.192	.615	.781 (.715 - .827)	.728 (.728 - .831)
	me8	.709	.863	.746			
	me15	.618	.872	.761			
	me22	.676	.872	.760			
Introjected Regulation	me7	.655	.775	.601	.719	.858 (.827 - .886)	.859 (.829 - .887)
	me14	.709	.841	.708			
	me21	.730	.836	.698			
	me28	.732	.932	.869			
Identified Regulation	me3	.462	.755	.570	.623	.750 (.660 - .810)	.764 (.677 - .824)
	me10	.611	.782	.611			
	me17	.550	.792	.627			
	me24	.650	.824	.679			
MI to Knowledge	me2	.622	.749	.561	.692	.830 (.772 - .869)	.836 (.787 - .878)
	me9	.712	.859	.738			
	me16	.571	.788	.621			
	me23	.763	.920	.846			
MI to Achievement	me6	.594	.812	.660	.690	.798 (.745 - .839)	.805 (.752 - .846)
	me13	.710	.889	.791			

Subscale	Item	r_{bis}	λ	$λ^2$	AVE	Reliability	
						(CI)	(CI)
MI Stimulating Experiences	me20	.668	.814	.663	.700	.848 (.810 - .878)	.859 (.828 - .885)
	me27	.583	.803	.644			
	me4	.490	.675	.456			
	me11	.754	.827	.685			
	me18	.800	.912	.833			
	me25	.719	.906	.820			
Amotivation	me5	.734	.886	.786	.831	.872 (.821 - .906)	.871 (.821 - .906)
	me12	.688	.850	.723			
	me19	.780	.965	.931			
	me26	.755	.923	.853			

Note. *n* = 300; ∝ = Cronbach's Alpha; ω = McDonald's Omega; CI = Confidence Interval.

Despite the fact that the bifactor model did not obtain the best fit indices, we believe it is necessary to present the results of this model, including the ancillary statistics that can help determine the unidimensionality or multidimensionality of the AMS (see Table 5). It is important to point out that 20 of the 28 items of the AMS obtained stronger factor loadings on the general factor in contrast to 8 items that obtained stronger factor loadings on its respective specific factor. For example, the factor loading mean of the general factor was $λ_{Mean}$ = .549 and the factor loading mean of the specific factors fluctuated between $λ_{Mean}$ = .154 and $λ_{Mean}$ = .879, being the lowest of the IMA subscale and the highest the amotivation subscale. Therefore, the average factor loadings of the subscales were $λ_{Mean}$ = .485. Regarding the $ω_H$ of the general factor, it obtained a value of .857, providing information on the amount of total variance that can be attributed to the general factor (Zinbarg et al., 2006), which is well beyond the threshold of .80 and probably it might be possible to consider the AMS as a unidimensional measure. Moreover, $ω_{HS}$ values obtained by the subscales fluctuated between $ω_{HS}$ =.032 and $ω_{HS}$ = .897 in which IMA subscale obtained the lowest value and the amotivation subscale the highest. There are authors (e.g., Smits et al., 2015), that consider values of $ω_{HS}$ between .20 and .30 is acceptable because they reflect a moderate proportion of the variance; however, there are other authors (e.g., Arias et al., 2018) who consider values less than .50 as an impediment to interpreting it as a factor. On the other hand, although the PUC was high (.889), the magnitude of the ECV (.580) was less than .70, which suggests that the data is not unidimensional enough (Quinn, 2014; Reise et al., 2013; Rodriguez et al., 2016). In terms of the ECV-I, 8 items have a significant influence on the overall factor: me3 and me17 (IndR), me16 and me23 (IMK), me13, me20 and me27 (IMA), and me12 (Amot). In other words, these are items that are essentially explained by the general factor and are better indicators of the general factor than of its specific factor as suggested by some literature (e.g., Montes & Sánchez, 2019). On the other hand, there are some items that seem to be good indicators for both the general factor and its respective specific factor, with corresponding ECV-I values around .50, such as items me15 and me22 of the ER subscale, and items me7 and me21 of the IntR subscale. Meanwhile, the

Table 5
Bifactor-CFA factor loadings of unifactor (Unif), general (GF), specific factors and ancillary statistics results.

Scale	Item	Unif	GF	Specific Factors							ECV-I	ARPB
				ER	IntR	IndR	IMK	IMA	IMSE	Amot		
ER	me1	.275	.244	.492							.197	.127
	me8	.659	.593	.675							.436	.111
	me15	.669	.625	.547							.566	.070
	me22	.674	.614	.608							.505	.098
IntR	me7	.652	.604		.522						.572	.079
	me14	.723	.674		.504						.641	.073
	me21	.711	.649		.572						.563	.096
	me28	.798	.766		.467						.729	.042
IndR	me3	.661	.675			.238					**.889**	.021
	me10	.690	.689			.445					.706	.001
	me17	.687	.712			.311					**.840**	.035
	me24	.725	.731			.416					.755	.008
IMK	me2	.676	.653				.484				.645	.035
	me9	.780	.775				.416				.776	.006
	me16	.713	.730				.206				**.926**	.023
	me23	.836	.830				.360				**.842**	.007
IMA	me6	.763	.783					.591			.637	.026
	me13	.836	.860					.178			**.959**	.028
	me20	.757	.806					.008			1.00	.061
	me27	.736	.804					-.163			**.961**	.085
IMSE	me4	.515	.533						.273		.792	.034
	me11	.701	.528						.690		.369	.328
	me18	.794	.625						.730		.423	.270
	me25	.770	.710						.480		.686	.085
Amot	me5	-.590	-.331							.809	.143	-.782
	me12	-.493	-.246							.811	**.840**	1.004
	me19	-.761	-.165							.955	.029	-3.612
	me26	-.702	-.112							.937	.014	-5.268
λ_{Mean}			.549	.581	.516	.353	.367	.154	.543	.878		
PUC			.889									
ECV			.580									
ARPB			-.318									
ω_H			.857									
ω_{HS}				.483	.338	.175	.175	.032	.404	.897		

ARPB value was equal to .318, which exceeds the threshold of .12 to .15 (Rodriguez et al., 2016), which is considered an acceptable criterion and therefore the ARPB value presents an inconsistency between the factor loadings of the unidimensional model and the general factor of the bifactor model. However, when we look at the ARPB of each item, we can see that only six items exceeded the ARPB criteria: me11 and me18 (IMSE) and the four items of the amotivation subscale (me5, me12, me19 and me26).

Since the seven-factor model was the best fitted, we examined the measurement invariance of the AMS by gender and age. Thus, measurement invariance was done with a bottom-up approach, from an unrestricted model to a model with strong restriction (Stark et al., 2006). Thus, we tested an unrestricted model of equality (configurational invariance) and continued with successive restrictions applied to factor loadings and thresholds (metric invariance), and intercepts (scalar invariance). Considering the sample size (> 300; Chen, 2007), the invariance criteria were: CFI < .010, SRMR < .030 and RMSEA < .015 (Chen, 2007). As such, measurement invariance in every group analyzed (i.e., gender and age) were good and complied with the established criteria. The differences between fit indices (Δ_{SRMR}, Δ_{RMSEA}, Δ_{CFI}, and Δ_{TLI}) were within limit, suggesting that the AMS was invariant among those groups (see Table 6).

Finally, the scores between the subscales of the AMS were correlated to demonstrate the presence of a continuum that goes from amotivation to IM. We should find high and positive correlations between the adjacent subscales and negative correlations among the scales opposite the construct on the continuum (Deci & Ryan, 1985). In Table 7, the correlations of the covariances (under the diagonal) that were high and strong between the latent variable IM-knowledge and IM-achievement, but IM-achievement with introjected regulation and identified regulation stand out. Also noteworthy are the correlations of the covariances between the latent variable of amotivation and the remaining latent variables close to zero and other negative ones. Similarly, correlations of the observed scores (above the diagonal) can be appreciated between IM at achievement and robust correlation introjected regulation.

Discussion

This study aimed to examine the internal structure and psychometric properties of the

Table 6
Measurement invariance of the Academic Motivation Scale by gender and age.

Model	χ^2 (*df*)	SRMR	RMSEA	CFI	TLI	Model of Reference	$\Delta\chi^2$	ΔSRMR	ΔRMSEA	ΔCFI	ΔTLI
Multigroup analysis by gender (male/female)											
1: Configural	916.06* (658)	.086	.051	.995	.994	-----	-----	-----	-----	-----	-----
2: Metric	1,027.03* (679)	.090	.059	.993	.992	1	+110.97	+.004	+.008	-.002	-.002
3: Scalar	993.54* (806)	.086	.040	.996	.996	2	-33.49	-.004	-.019	+.003	+.004
Multigroup analysis by age (21-30 /31-40)											
1: Configural	1,066.35* (658)	.091	.065	.992	.991	-----	-----	-----	-----	-----	-----
2: Metric	1,121.11* (679)	.092	.066	.991	.990	1	+54.76	+.001	+.001	-.001	-.001
3: Scalar	1,140.07* (806)	.091	.053	.993	.994	2	+18.96	-.001	-.013	+.002	+.004

Note. *p < .05; *df* = degree of freedom.

Table 7
Correlation matrix between latent variables (under the diagonal) and observed variables (above the diagonal) of subscales of the Academic Motivation Scale.

	Subscale	1	2	3	4	5	6	7
1.	External Regulation	1	.524**	.587**	.272**	.389**	.276**	.055
2.	Introjected Regulation	.717**	1	.495**	.429**	.682**	.423**	.081
3.	Identified Regulation	.791**	.682**	1	.569**	.583**	.416**	-.251**
4.	IM-Knowledge	.465**	.597**	.819**	1	.717**	.636**	-.249**
5.	IM-Achievement	.587**	.846**	.791**	.884**	1	.615**	-.176**
6.	IM-Stimulating Experiences	.376**	.519**	.560**	.797**	.732**	1	-.025
7.	Amotivation	.032	.067	-.427**	-.366**	-.318**	-.028	1

Note. $n = 300$; *$p < .05$; **$p < .01$.

AMS in a sample of graduate students in Puerto Rico. The results of the confirmatory factor analysis supported the internal structure of seven factors of the AMS, given that it was the model that obtained the best-fit indices and which is consistent with the theory used by the authors to construct it. Moreover, as reported in other studies (e.g., Alonso et al., 2006), these fit indices were achieved without correlating errors between the indicators. This is probably because a polychoric matrix and a more appropriate estimator (WLSMV) were used to perform the data analyses instead of the maximum likelihood estimator. On the contrary, even though the seven-factor model surpassed the bifactor model, some ancillary statistics suggest that the AMS could be more unidimensional than multidimensional due to the high value obtained from PUC. However, since the ECV is less than .70, this could also suggest multidimensionality, as proposed by the authors of the test (Núñez et al., 2010). These results of the ancillary statistics could probably suggest that both the seven factors and a general factor provide relevant information for the understanding of the data obtained from the AMS. It should be noted that the items of the amotivation subscale were the only ones with negative factor loadings both in the unidimensional model and in the general factor of the bifactor model, which could be affecting the values of the supplementary statistics of the bifactor model and making the unidimensionality/multidimensionality interpretation of the AMS more difficult.

The present study provides insight into measurement invariances of the AMS across gender and age. Since the seven-factor model obtained the best-fit indices and is consonant with the theory in which the AMS was developed, we tested the measurement invariance of this model. We tested the measurement invariances of AMS among students at different universities in Puerto Rico. Exploration on the first two levels revealed metric or factor loading invariance (i.e., weak measurement invariance) and scalar invariance (i.e., strong measurement invariance) of the seven-factor model across gender and age. Metric invariance is important to ensure the measure across multiple groups is on the same scale or that all groups' factors are measured similarly (Meredith & Teresi, 2006; Vandenberg & Lance, 2000; Wang & Wang, 2012). Scalar invariance refers to the item intercept being invariant across multiple groups in the present study. This indicates that none of the groups tends to respond systematically higher or lower to the items of scales than other groups (Meredith & Teresi, 2006; Vandenberg &

Lance, 2000; Wang & Wang, 2012). The present study met both invariance requirements. These results confirm that the compared groups had an equivalent understanding of each of the 28 items in the measure, an important prerequisite for making a meaningful comparison between groups on academic motivation. Researchers have argued that error variance invariance (i.e., strict measurement invariance) is not required for substantive analyses in many disciplines, and such invariance is considered unnecessary (Wang & Wang, 2012).

In terms of the correlations between the subscales of the AMS, it was possible to appreciate that, in general, it supports the conceptual framework of Deci and Ryan (1985) regarding the presence of a continuum that goes from amotivation to IM, where adjacent scales show higher correlations than opposite ones on the continuum. However, the IM to Achievement subscale presented higher covariance correlations with Introjected Regulation and Identified Regulation than with its adjacent dimensions of IM to Knowledge and MI to Stimulating Experiences, these results being similar to other studies (Cokley et al., 2001; Fairchild et al., 2005; Nuñez et al., 2010; Nuñez et al., 2006; Vallerand et al., 1993). Thus, we agree with Cokley et al. (2000) and Nuñez et al. (2006), who indicate that this could be because the difference between the EM and IM constructs is not as categorical as the self-determination theory proposes. Therefore, we echo Nuñez et al. (2006) that the items of the Introjected Regulation, Identified Regulation, and MI Achievement subscales should be reviewed for future research since they could be sharing a common factor given that these dimensions have some satisfaction, prove that they can achieve proposed goals and achieve a better future.

Limitations and Recommendations

The results, while satisfactory, should be interpreted with caution due to certain limitations of the study. Although the findings are consistent with the proposal by Deci and Ryan (1985) and exceed those of the bifactor model, the high correlations between the MI and EM subscales could be indicative that they do not differ as much as the theory supposes or that they come from a common factor as suggested by some of the ancillary statistics (e.g., PUC) of the bifactor model.

The bifactor model also presents some drawbacks that must be considered. As previously stated, good statistical results do not guarantee the existence of a general factor (Bonifay et al., 2017). Nonetheless, a CFA only allows the items of each factor to load on them, but it does not allow the items to load on other factors, which tends to be unrealistic with the psychological constructs, (Furr, 2022) since they tend to relate. In this way, for future research, it is recommended that the sample of students is expanded and that exploratory structural equation modeling (ESEM) and bifactor ESEM should be carried out to be clear about the unidimensionality/multidimensionality of the AMS and that the application ESEM and bifactor ESEM might help in this endeavor.

Conclusion

Results of the current study provide evidence that supports the internal structure of the AMS and the ancillary statistics use; additionally the bifactor model presents some interesting information about the possible unidimensional or multidimensional uses of the AMS. The subscales of the AMS obtained good reliability coefficients, and the AMS appears to be invariant among gender and age, which permits comparison

among these groups. Therefore, the AMS is useful in the educational context with graduate students in Puerto Rico.

References

Alonso, J. L. N., Lucas, J. M.-A., Izquierdo, J. G. N. (2006). Validación de la Escala de Motivación Educativa (AMS) en Paraguay. *Interamerican Journal of Psychology, 40*(2), 185-192. https://doi.org/10.1037/t85057-000

Alonso-Tapia, J. (1998). *Motivación y aprendizaje en el aula. Cómo enseñar a pensar*. Santillana.

Arias, V. B., Ponce, F. P., & Nuñez, D. E. (2018). Bifactor models of attention-deficit/hyperactivity disorder (ADHD): An evaluation of three necessary but underused psychometric indexes. *Assessment, 25*(7), 885-897. https://doi.org/10.1177/1073191116679260

Ato, M., López-García, J. J., & Benavente, A. (2013). Un sistema de clasificación de los diseños de investigación en psicología. *Anales de Psicología, 29*(3). https://doi.org/10.6018/analesps.29.3.178511

Becerra-González, C. E., & Reidl-Martínez, L. M. (2015). Motivación, autoeficacia, estilo atribucional y rendimiento escolar de estudiantes de bachillerato. *Revista Electrónica de Investigación Educativa, 17*(3). https://redie.uabc.mx/redie

Browne, M. W., & Cudeck, R. (1989). Single sample cross-validation indices for covariance structures. *Multivariate Behavioral Research, 24*(4), 445-455. https://doi.org/10.1207/s15327906mbr2404_4

Bonifay, W., Lane, S. P., & Reise, S. P. (2017). Three concerns with applying a bifactor model as a structure of psychopathology. *Clinical Psychological Science, 5*(1), 184–186. https://doi.org/10.1177/2167702616657069

Burgueño, R., Sicilia, Á., Medina-Casaubón, J., Alcaraz-Ibáñez, M., & Lirola, M. J. (2017). Academic motivation scale revised. Inclusion of integrated regulation to measure motivation in initial teacher education. *Anales de Psicología, 33*(3), 670-679. https://doi.org/10.6018/analesps.33.3.249601

Byrne, B. M. (2016). *Structural equation modelling with AMOS basic concepts, applications, and programming* (3rd ed.). Routledge.

Candela, F., Zucchetti, G., & Villosio, C. (2014). Preliminary validation of the Italian version of the original sport motivation scale. *Journal of Human Sport & Exercise, 1*(9), 136-147. http://dx.doi.org/10.4100/jhse.2014.91.14

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling, 14*(3), 464-504. https://psycnet.apa.org/doi/10.1080/10705510701301834

Chicaiza-Ayala, W., & Cragno, A. G. (2018). Motivation in three Medical Schools in Ecuador. *Educación Médica, 19*, 98-104. https://doi.org/10.1016/j.edumed.2017.03.031

Cokley, K. O. (2000). Examining the validity of the Academic Motivation Scale by comparing scale construction to self-determination theory. *Psychological Reports, 86*(2), 560–564. https://doi.org/10.2466/PR0.86.2.560-564

Cokley, K. O., Bernard, N., Cunningham, D., & Motoike, J. (2001). A psychometric investigation of the academic motivation scale using a United States sample. *Measurement and evaluation in Counseling and development, 34*(2), 109-119. https://doi.org/10.1080/07481756.2001.12069027

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic Motivation and Self-Determination in Human Behavior*. Springer. https://doi.org/10.1007/978-1-4899-2271-7

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry, 11*(4), 227-268. https://doi.org/10.1207/S15327965PLI1104_01

Fairchild, A. J., Horst, S. J., Finney, S. J., & Barron, K. E. (2005). Evaluating existing and new validity evidence for the Academic Motivation Scale. *Contemporary Educational Psychology, 30(*3), 331–358. https://doi.org/10.1016/j.cedpsych.2004.11.001

Ferrando, P. J., & Lorenzo-Seva, U. (2017). Assessing the quality and appropriateness of factor solutions and factor score estimates in exploratory item factor analysis. *Educational Psychological Measurement, 78*(5), 762-780. https://doi.org/10.1177/0013164417719308

Flores-Kanter, P. E., Dominguez-Lara, S., Trógolo, M. A., & Medrano, L. A. (2018). Best Practices in the Use of Bifactor Models: Conceptual Grounds, Fit Indices and Complementary Indicators. *Revista Evaluar, 18*(3). https://doi.org/10.35670/1667-4545.v18.n3.22221

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39-50. https://doi.org/10.2307/3151312

Furr, R. M. (2022). *Psychometrics: An introduction*. SAGE.

Gutiérrez-Ruiz, K. (2015). Perfil agentivo de estudiantes con bajo rendimiento académico: Estrategias cognitivas y de control de aprendizaje, autoeficacia académica y motivación. *Informes Psicológicos, 15*(1), 63-81. http://dx.doi.org/10.18566/infpsicv15n1a04

Guzmán, J. F., Carratalá, E., García-Ferriol, A., & Carratalá, V. (2006). Propiedades psicométricas de una escala de motivación deportiva. *European Journal of Human Movement, 16*, 85-98. https://www.eurjhm.com/index.php/eurjhm/article/view/161

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis*. Cengage.

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

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

Lepper, M. R. (1988). Motivational considerations in the study of instruction. *Cognition and Instruction, 5(*4), 289-309. https://doi.org/10.1207/s1532690xci0504_3

Li, C.-H. (2016a). Confirmatory factor analysis with ordinal data: Comparing robust maximum likelihood and diagonally weighted least squares. *Behavioral Research Methods, 48*(3), 936-949. https://doi.org/10.3758/s13428-015-0619-7

Li, C.-H. (2016b). 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

Littlewood-Zimmerman, H. F., & Bernal-Garcia, E. R. (2011). M*i primer modelamiento de ecuacion structural LISREL*. https://docplayer.es/50148837-Mi-primer-modelamiento-de-ecuacion-estructural-lisrel.html

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1*(2), 130–149. https://doi.org/10.1037/1082-989X.1.2.130

Marsh, H. W., Balla, J. R., & Hau, K. T. (1996). An Evaluation of Incremental Fit Indexes: A Clarification of Mathematical and Empirical Properties. In G. A. Marcoulides, & R. E. Schumacker (Eds.), *Advanced Structural Equation Modeling Techniques* (pp. 315-353). Lawrence Erlbaum.

Meredith, W., & Teresi, J. A. (2006). An Essay on Measurement and Factorial Invariance. *Medical Care, 44*(11, Suppl 3), S69-S77. https://doi.org/10.1097/01.mlr.0000245438.73837.89

Montes, S. A., & Sánchez, R. O. (2019). El factor p. ¿La estructura subyacente a la psicopatología? *Revista Evaluar, 19*(3), 20-41. https://revistas.unc.edu.ar/index.php/revaluar

Núñez, J. L., Lucas, J. M., Navarro Izquierdo, J. G., & Grijalvo Lobera, F. (2006). Validación de la Escala de Motivación Educativa (EME) en Paraguay. *Revista Interamericana de Psicología/Interamerican Journal of Psychology, 40*(3), 391-398. https://www.redalyc.org/pdf/284/28440314.pdf

Núñez, J. L., Martín-Albo, J., Navarro, J. G., & Suárez, Z. (2010). Adaptación y validación de la versión Española de la Escala de Motivación Educativa en estudiantes de educación secundaria postobligatoria.

*Estudios de Psicología, 31(*1), 89-100. https://doi.org/10.1174/021093910790744590

Quinn, H. (2014). *Bifactor models, explained common variance (ECV), and the usefulness of scores from unidimensional item response theory analyses*. University of North Carolina at Chapel Hill. https://doi.org/10.17615/t6ff-a088

Reise, S. P., Moore, T. M., & Haviland, M. G. (2010). Bifactor models and rotations: Exploring the extent to which multidimensional data yield univocal scale scores. *Journal of Personality Assessment, 92*(6), 544-559. https://doi.org/10.1080/00223891.2010.496477

Reise, S. P., Bonifay, W. E., & Haviland, M. G. (2013). Scoring and modeling psychological measures in the presence of multidimensionality. *Journal of Personality Assessment, 95*(2), 129-140. https://doi.org/10.1080/00223891.2012.725437

Reyes-Cruz, M. d. R., & Gutiérrez-Arceo, J. M. (2015). Sentido de autoeficacia en investigación en estudiantes de postgrado. *Sinéctica, 45*. https://sinectica.iteso.mx/index.php/SINECTICA/article/view/590

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2015). Applying bifactor statistical indices in the evaluation of psychological measures. *Journal of Personality Assessment, 98*(3), 223-237. https://www.tandfonline.com/doi/full/10.1080/00223891.2015.1089249

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Evaluating bifactor models: Calculating and interpreting statistical indices. *Psychological Methods, 21*(2), 137-150. http://dx.doi.org/10.1037/met0000045

Rodríguez, S., Núñez, J. C., Valle, A., Blas, R., & Rosario, P. (2009). Autoeficacia docente, motivación del profesor y estrategias de enseñanza. *Escritos de Psicología, 3*(1), 1-7. https://doi.org/10.24310/espsiescpsi.v3i1.13328

Román-Pérez, C. (2013). El buen rendimiento escolar en los estudiantes que ingresan a la universidad. *Calidad en la Educación, 38*, 147-179. https://doi.org/10.31619/caledu.n38.108

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

Ruiz, K. G. (2015). Perfil agentivo de estudiantes con bajo rendimiento académico: Estrategias cognitivas y de control del aprendizaje, autoeficacia académica y motivación. *Fundación Universitaria Tecnológico Comfenalco, 15*(1), 63-81. https://dialnet.unirioja.es/descarga/articulo/5229768.pdf

Sánchez, J. J. M., & Hernández-Pina, F. (2011). Influencia de la motivación en el rendimiento académico de los estudiantes de formación profesional. *Revista Electrónica Interuniversitaria de Formación de Profesorado, 14*(1), 81-100.

Silva, P. C. d. C., Sicilia, A., Burgueño, R., & Lirola, M. J. (2018). Academic motivation in the initial training of physical education teachers. *Revista Internacional de Medicina y Ciencias de La Actividad Física y Del Deporte, 18*(71), 537-554. https://doi.org/10.15366/rimcafd2018.71.009

Smits, I. A. M., Timmerman, M. E., Barelds, D. P. H., & Meijer, R. R. (2015). The Dutch symptom checklist-90-revised: Is the use of the subscales justified? *European Journal of Psychological Assessment, 31*(4), 263-271. https://psycnet.apa.org/doi/10.1027/1015-5759/a000233

Stark, S., Chernyshenko, O. S., & Drasgow, F. (2006). Detecting differential item functioning with confirmatory factor analysis and item response theory: Toward a unified strategy. *Journal of Applied Psychology, 91*(6), 1292–1306. https://doi.org/10.1037/0021-9010.91.6.1292

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. Psychometrika, 38, 1-10. https://doi.org/10.1007/BF02291170

Vallerand, R. J., Pelletier, L. G., Blais, M. R., Brière, N. M., Senécal, C., & Vallières, E. F. (1993). On the assessment of intrinsic, extrinsic, and amotivation in education: Evidence on the concurrent and construct validity of the Academic Motivation Scale. *Educational and Psychological Measurement, 53*, 159-172. https://journals.sagepub.com/toc/epma/53/1

Vallerand, R. J., Pelletier, L. G., Blais, M. R., Brière, N. M., Senecal, C., & Vallieres, E. F. (1992). The Academic Motivation Scale: A measure of intrinsic, extrinsic, and amotivation in education. *Educational and Psychological Measurement, 52*(4), 1003–1017. https://doi.org/10.1177/0013164492052004025

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational Research Methods, 3*(1), 4-69. https://doi.org/10.1177/109442810031002

Wang, J., & Wang, X. (2012). *Structural Equation Modeling Applications Using Mplus*. John Wiley & Sons Ltd. http://dx.doi.org/10.1002/9781118356258

Zhang, D., Wehmeyer, M. L., & Chen, L.-J. (2005). Parent and teacher engagement in fostering the self-determination of students with disabilities: A comparison between the United States and the Republic of China. *Remedial and Special Education, 26*(1), 55-64. https://doi.org/10.1177/07419325050260010701

Zinbarg, R. E., Yovel, I., Revelle, W., & McDonald, R. P. (2006). Estimating generalizability to a latent variable common to all of a scale's indicators: A comparison of estimators for ω_H. *Applied Psychological Measurement, 30*(2), 121-144. https://journals.sagepub.com/toc/apma/30/2

2024, Vol. 24, No. 1
ISSN 1667-4545
Recuperado de https://revistas.unc.edu.ar/index.php/revaluar

Laboratorio de Evaluación Psicológica y Educativa
Facultad de Psicología - Universidad Nacional de Córdoba

Development of a Scale to Assess Socioemotional Skills in Argentine Children Aged 9 to 12 Years

Desarrollo de una Escala para Evaluar Habilidades Socioemocionales en Niños Argentinos de 9 a 12 años

Eliana Vanesa Zamora * [1] , Macarena del-Valle [1] , Rosario Gelpi-Trudo, Cintia Olsen, Hernán López-Morales, Lorena Canet-Juric [1] , María Marta Richard's [1]

1 - Instituto de Psicología Básica, Aplicada y Tecnología (IPSIBAT), Universidad Nacional de Mar del Plata (UNMDP), Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET). Mar del Plata, Argentina.

Recibido: 30/08/2023 **Revisado:** 02/12/2023 **Aceptado:** 12/12/2023

Abstract

Socioemotional skills allow us to understand and manage emotions, to set and reach positive goals, to maintain positive relationships and to make responsible decisions. This study aimed to develop and validate a scale to assess these skills from 9 to 12 years old Argentinian children. 263 fourth, fifth and sixth graders (age M = 10 years, 4 months) from Mar del Plata, Argentina answered to the developed items, along with the scales CERQ-k and DERS. The final version contained 26 items, grouped, according to factorial analyses, in four dimensions in line with the background theory (social skills, emotional regulation, growth mindset and self-control). The scale showed adequate internal consistency (α=.87), test-retest stability (*r* = .67), evidence of content validity according to expert judgments, and of criterion validity through associations with adaptive strategies and difficulties in emotional regulation. This scale allows for easy and reliable assessment of socioemotional skills, in research, clinical and educational contexts.

Keywords: *social-emotional skills, growth mindset, emotional regulation, self-control, social skills, assessment, children*

Resumen

Las habilidades socioemocionales permiten comprender y gestionar emociones, establecer y alcanzar metas positivas, mantener relaciones positivas y tomar decisiones responsables. Este estudio tuvo como objetivo desarrollar y validar una escala para evaluar estas habilidades en niños argentinos de 9 a 12 años de edad. Así, 263 estudiantes de cuarto, quinto y sexto grado (promedio de edad M = 10 años, 4 meses) de Mar del Plata, Argentina, respondieron a los ítems desarrollados, junto con las escalas CERQ-k y DERS. La versión final de la escala está compuesta por 26 ítems, agrupados, según análisis factoriales, en cuatro dimensiones alineadas con la teoría de base (habilidades sociales, regulación emocional, mentalidad de crecimiento y autocontrol). La escala mostró una adecuada consistencia interna (α = .87), estabilidad test-retest (*r* = .67), evidencia de validez de contenido según juicios de expertos, y de validez de criterio a través de asociaciones con estrategias adaptativas y dificultades en la regulación emocional. Esta escala permite una evaluación sencilla y confiable de las habilidades socioemocionales, en contextos de investigación, clínicos y educativos.

Palabras clave: *habilidades socioemocionales, mentalidad de crecimiento, regulación emocional, autocontrol, habilidades sociales, evaluación, niños*

***Correspondence to:** Eliana V. Zamora. E-mail: eliana.zamora@conicet.gov.ar. UNMDP- Facultad de Psicología, Deán Funes 3280 - Cuerpo V - Nivel III (7600). Tel: +54 223 4752526. Mar del Plata, Argentina.

How to cite: Zamora, E., V., del-Valle, M., Gelpi-Trudo, R., Olsen, C., López-Morales, H., Canet-Juric, L., & Richatd´s, M., (2024): Development of a Scale to Assess Socioemotional Skills in Argentine Children Aged 9 to 12 Years. *Revista Evaluar, 24*(1) 44-59. Retrieved from: https://revistas.unc.edu.ar/index.php/revaluar

Participaron en la edición de este artículo: Gloria Nieve, Eugenia Barrionuevo, Florencia Ruiz, Benjamín Casanova, Jorge Bruera.

Author´s note: Este trabajo fue financiado con fondos de la Agencia de Promoción Científica y Tecnológica de Argentina, Ministerio de Ciencia y Tecnología de la Nación (MINCyT).

Introduction

School, beyond providing knowledge in areas, such as mathematics and language, provides an environment conducive to establishing friendships, fostering collaboration and developing personal responsibility skills (Zamora et al., 2019). These skills, known as 21st century skills (Schleicher, 2018), are essential for success in adult life. Internationally, social-emotional learning (SEL, *Social and Emotional Learning*) educational programs have emerged (Collaborative for Academic, Social, and Emotional Learning [CASEL], 2015) and categorized as *universal* interventions (since they are implemented in groups that do not present specific problems) that have shown an impact on academic performance, behavioral problems, and subsequent labor and social insertion (Durlak et al., 2022).

SEL has been defined as the process by which people understand and manage emotions, set and achieve positive goals, show empathy for others, establish and maintain positive relationships, and make responsible decisions (CASEL, 2015). The SEL skills model proposed by CASEL (CASEL, 2013) has influenced numerous interventions internationally (Cipriano et al., 2023) and nationally (Canet-Juric et al., 2020). From the effects observed, there is an agreement that programs should focus on both intrapersonal and interpersonal skills and attitudes (Durlak et al., 2022). Thus, different competencies are identified which should be addressed in children's environments (classroom, school, family and community) to favor and lead their comprehensive development (Durlak et al., 2015).

Within these competencies, we can highlight, primarily, Self-Control (SC), which is the process of regulating behavior, emotion and cognition to achieve important goals that conflict with more tempting goals in the present (Duckworth & Steinberg, 2015). Next, Emotional Regulation (ER) and Emotion Awareness, which involves the skills to recognize and identify emotions, and apply strategies and skills to monitor, evaluate, and modify the nature and course of an emotion to accomplish goals (Gross, 2014). Further, Growth Mindset (GM) refers to the belief that both intelligence and other human qualities can be developed through effort and learning (Dweck, 2006). Finally, Social Skills (SS), which are behaviors and abilities that are brought into play in interpersonal situations for an adequate adjustment to the social environment, such as listening, communication and cooperation skills, or the ability to establish and maintain positive relationships (Gresham, 2016).

Although the CASEL organization publishes SEL skills assessment guides (Crowe et al., 2011) and has at least a dozen specific SEL instruments (Humphrey et al., 2011; Gresham et al., 2018), the available instruments sometimes do not allow us to project lines of assessment for each SEL skill, cultural context, age range and/or participant or informant. In Table 1, there is a synthesis of the most used instruments. In general, the most used assessment format is the self-report, given that its ease of administration makes it especially attractive for the school context (del-Valle & Zamora, 2021).

The *Social Skills Improvement System-Rating Scales* (SSIS-RS; Gresham & Elliot, 2008) is a scale for students, teachers and families available in English and Spanish for ages 8 to 18, which assesses SEL skills, and considers teacher-student relationships, peer interactions and academic performance. On the contrary, the *Devereux Student Strengths Assessment Scale* (DESSA; LeBuffe et al., 2009) is a scale for teachers and families that assesses eight competencies that allow the construction of a strengths profile. Among the main advantages of the scale are the speed with

Table 1
Review of assessment instruments for SEL skills.

Authors	Scale name	Age	Variables evaluated	Who responds	Availability in Spanish	Paid/ Free
Gresham and Elliott, 2008	SISS-RS	8-18	Decision-making -Self-awareness -Self-management -SS -Teacher-student relationship -Academic performance	Teachers, parents and students	Yes	Pay
LeBuffe, Shapiro and Naglieri, 2009.	DESSA	4 - 13	-Personal responsibility -Optimistic thinking Goal-directed behavior -Social awareness Decision-making -SS -Self-awareness -Self-management	Teachers, after-school program staff and parents/ guardians	Yes	Pay
Districts Social Emotional Learning Surveys, Gehlbach and Hough, 2018.	CORE	9 - 18	-GM -Self-efficiency -Self-management -Social awareness	Teachers and students	Yes	Free access
Washoe County School District, 2018	WSCD Long	10 - 18	-SS -Responsible decision making Self-awareness of emotions Self-awareness of strengths -SC -Self-management of goals and school work -Social awareness	Students	Yes	Free access
Milicic, Alcalay, Berger and Torretti, 2014.	Social Emotional Learning Self-Report Scale for Children	8 - 12	Self-awareness -Awareness of others -Self-regulation -Peaceful conflict resolution -SS	Students	Yes	Free access
Washoe County School District, 2018	WCSD Student Social and Emotional Competency Assessment Short	10 - 18	-SS -Responsible decision making Self-awareness of emotions Self-awareness of strengths and weaknesses Self-management of emotions Self-management of goals Self-management of school work -Social awareness	Students	Yes	Free access

Note. ER = Emotional Regulation; GM = Growth Mindset; SS = Social Skills; SC = Self-Control.

which it can be administered and the age range it covers. Then, the CORE districts questionnaire (Gehlbach & Hough, 2018) assesses four competencies in children between 9 to 18 years: growth mindset, self-efficacy, self-management and social awareness. Advantages of this scale include its free access and the number of languages in which it is translated. The *WSCD scale* (Washoe County School District, 2018), in its 40-item version, assesses social and emotional competencies based on the strengths that students can report about themselves. Finally, the *Social-Emotional Learning Self-Report Scale for Children* (Milicic et al., 2014) aims to assess SEL competencies in children in both clinical and educational settings.

Although different scales have been developed in recent years and their availability has increased, they have some disadvantages for their application. First, the age range is limited (e.g., scales that cover 8 to 9 years) or, on the contrary, very wide (e.g., 8 to 18 years), without considering the variability associated with development. Likewise, the available instruments assess specific skills and not necessarily SEL skills (e.g., self-esteem, attention, motivation), have not been validated and adapted to the Argentine population (e.g., Socioemotional Learning Self-Report Scale, CORE scales), and finally, not all are available for free download and use. In addition, there are no adaptations that include the age range of 8 to 12 years, which in Argentine education corresponds to the last years of primary school. This is striking, since this period involves evolutionary and contextual changes that lay the foundations for a successful adjustment to early adolescence (Huston & Ripke, 2006).

Given the fundamental role of SEL skills in adapting to the school and social environment (Panayiotou & Humphrey, 2018), it is highly beneficial to possess valid and reliable assessment measures for their evaluation. Therefore, this paper aims to present the development and validation of a self-report questionnaire (ESH-A) designed to assess Socioemotional Skills in Argentine boys and girls between 9 to 12 years old. Three objectives have been outlined: (1) to analyze the development and content validity of the EHS-A through expert-judge analysis, (2) to examine the factor structure of the scale using confirmatory analysis (construct validity) and to verify its reliability and (3) to assess the criterion validity of the instrument by examining its association with theoretically related measures. The expectation is that the EHS-A will prove to be a valid and reliable scale, easily accessible, and adapted to the characteristics of the Argentine child population.

Methods

Design and participants

A correlational, non-experimental, cross-sectional design was used (Hernández-Sampieri et al., 2014). The sampling was non-probabilistic, purposive. Participants were 117 boys and 146 girls (N = 263) assisting 4th (n = 121), 5th (n = 46) and 6th (n = 96) grades at three private schools in Mar del Plata, Buenos Aires (Argentina). Ages ranged from 8 years, 9 months to 11 years, 9 months (M = 10 years, 4 months). Of the total, 91 children were re-evaluated after 9 months to analyze test-retest reliability. Inclusion criteria were as follows: absence of psychological or psychiatric treatment, normal or corrected vision, typical development, and absence of a history of learning or neurodevelopmental disorders.

Instruments

Socioemotional Skills: construction and selection of the ESH-A items. The main criterion for

the selection and formulation of the items was that they represented observable indicators at the cognitive (e.g., *I get distracted when I am studying*), behavioral (e.g., *I can stop doing something if I am told to*) and emotional (e.g., *I can control myself if I am angry*) domains. Moreover, these indicators were chosen based on their frequent manifestation in both school and home settings. Most items were selected from the available SEL scales, prioritizing those that presented a better fit to the constructs (see Table 1), and reformulated if necessary. For example, from the CORE scale, the items: *my intelligence is something I cannot change much* and *challenging myself will not make me smarter* were reformulated as *I think my intelligence is something I can change*, and for *every day I challenge myself to be smarter*, respectively. Consideration was given to crafting items with language that would be comprehensible to children. A preliminary version of the scale composed of 55 items was obtained.

Once the dimensions and items were defined, a content validity analysis was performed. For this purpose, the 55 items divided into the four SEL skills (SC, ER, GM and SS) were submitted by e-mail to 10 expert judges with knowledge in psychometrics and SEL skills. These items were assessed for their pertinence (suitability for the intended dimension), quality (clear wording and language appropriateness for the target population) and relevance (culturally applicable for the local population), expressed on a 5-point Likert scale. In addition, the judges were asked to indicate which SEL skill they considered each item referred to. In general, there was an agreement on which scale each item might belong to. The experts could suggest modifications and even the rejection of an item if needed. The results were analyzed using the validation coefficient (Aiken, 1985), and items considered relevant according to expert criteria were kept (see results). As well, some items were modified semantically or syntactically; afterwards, a pilot test was carried out. The 40 items best rated by the judges (10 per dimension) were administered to 8 children aged 9 to 11 years. Based on their responses, the comprehension of the items and instructions, the difficulty of the items, the length of the questionnaire, and the receptiveness of the sample to the instrument were assessed (Cohen et al., 2011). Once the pilot test was conducted, the final administration of the instrument was carried out.

Cognitive Emotional Regulation Questionnaire for Children (CERQ-k). It is a self-report Likert-type scale for children between 9 to 11 years, developed by Garnefski and Kraaij (2007) and adapted to the local context with good psychometric properties (e.g., Andrés, 2014). It is composed of 36 items that identifies the cognitive ER strategies that children use after experiencing negative events. It assesses nine factors, corresponding with nine ER strategies: five adaptive (acceptance, positive refocusing, putting into perspective, positive reappraisal, planning) and four maladaptive (self-blame, rumination, catastrophizing, other-blame). Composite reliability indices are above .70 in child population (Andrés, 2014).

Difficulties in Emotion Regulation Scale (DERS). Originally developed by Gratz and Roemer (2004) and adapted for the Argentinian population by Medrano and Trogolo (2014), this scale is composed of 36 items with a five-choice Likert-type response format. Its factor structure is composed of six dimensions: (1) Non-acceptance of emotional responses: tendency to experience secondary negative emotions as a response to a primary negative emotion; (2) Difficulty engaging in goal-directed behaviors: difficulties in concentrating or performing tasks when experiencing a negative emotion; (3) Impulse control difficulties: poor capacity to control one's behavior when experiencing a negative emotion; (4) Lack of emo-

tional awareness: extent to which one attends to and admit emotional states; (5) Lack of emotional clarity: extent to which a person knows and understands his or her emotions; (6) Limited access to emotion regulation strategies: belief that one will not be able to modify an unpleasurable emotional state. Previous studies support the suitability of the scale for the school-age population (e.g., Neumann et al., 2010). Internal consistency indices (Cronbach's α) range from .80 to .89 points in child population (Zamora et al., 2022). Specific dimension scores are often summed into a single overall ER difficulties score (e.g., del-Valle et al., 2020; Karatzias et al., 2016).

Procedure and ethical considerations

The present study was part of a larger research project approved by the Ethics Committee of the Interdisciplinary Thematic Program in Bioethics (PTIB) of the National University of Mar del Plata. The schools involved were informed of the goals and implications of the study, and informative meetings with teaching staff and families of the participants were held before data collection. An information sheet was handed out and families were invited to voluntarily participate in the study by signing an informed consent form. The children gave their informed assent to participate and could leave the study at any time if required. The guidelines of the National Council for Scientific and Technical Research of Argentina (CONICET) for ethical behavior in the Social Sciences and Humanities (Resolution No. 2857, 2006), the criteria for research recommended by the American Psychological Association (APA, 2010) and the Declaration of Helsinki (World Medical Association, 2013) were respected. The assessments were conducted by the authors of the study, at the educational institutions attended by the participants. Each item was read aloud to avoid possible difficulties in comprehension or interference of the reading processes.

Data analysis

To assess content validity (Objective 1), an analysis of agreement between experts was carried out using the validation coefficient (Aiken, 1985). This index allows the analysis of agreement between experts and establishes a criterion for making decisions about revising or eliminating items. Its magnitude ranges from 0.00 (*no agreement*) to 1.00 (*perfect agreement* among the judges regarding the highest validity score of the evaluated contents) (Soto & Segovia, 2009).

In relation to the analysis of the factorial structure of the scale and its reliability (Objective 2), all responses were coded and loaded into a general database. Reverse items were recorded. The applicability of the Exploratory Factor Analysis (EFA) was tested through Bartlett's sphericity test and the Kaiser-Meyer-Olkin (KMO) statistic, while the EFA was implemented through the FACTOR software (v. 12.03.02, Lorenzo-Seva & Ferrando, 2022). The number of factors to be extracted was estimated using parallel analysis with classical implementation (Horn, 1965), based on the polychoric correlation matrix (Ferrando & Anguiano-Carrasco, 2010). The factor extraction method was unweighted least squares (ULS), assumed robust when working with ordinal variables (Lloret-Segura et al., 2014). Oblique Promax was used for rotation, assuming interfactorial correlations (Lloret-Segura et al., 2014). With the retained items, a second EFA was performed. Then, a Confirmatory Factor Analysis (CFA) was accomplished with Lisrel (Scientific Software International, 2006), and the ULS was the estimation method used. Model fit

was evaluated through the following indices: χ^2, χ^2/df coefficient, GFI, AGFI, CFI, CFI, NFI and NNFI; while RMSEA was used as a measure of error (Hu & Bentler, 1998). The internal consistency of the retained factors was assessed using the omega coefficient. Interfactorial correlation was calculated using Pearson's *r* index. For the criterion validity analysis, correlations (Pearson's *r*) were performed between the factors and variables theoretically related to SEL skills: adaptive and maladaptive ER strategies (CERQ-k scale) and ER difficulties (DERS scale). Finally, the presence of differences according to gender and school year was analyzed for the extracted factors. For gender, Student's *t*-tests for independent samples were carried out (symmetry and kurtosis of the factors were between ± 2, suggesting normality; and Levene's tests suggested homoscedasticity). To evaluate whether there were differences according to school year, one-factor ANOVA tests were applied (after Levene's test suggested homoscedasticity).

Results

Aiken's *V* coefficient was calculated for content analysis. Results indicated that items 5, 9, 13 and 14 of the SC dimensions and items 1, 2 and 13 of the ER dimension had poor quality -with a score between .55 and .98-. However, items 5 and 9 referring to the GM dimension, items 4 and 6 referring to the SC dimension, and items 1, 2, 9, 11 and 13 referring to the SS dimension were low in terms of pertinence -with scores between .78 and .95-. In addition, these items showed lower values than the rest in terms of relevance (although above .50). Therefore, item 5 of the GM dimension, items 4, 5, 6, 9, 13 and 14 of the SC dimensions, items 1, 2, 9, 11 and 13 of the SS dimensions, and items 1, 2 and 13 of the ER dimension were eliminated for presenting low agreement in their content validity (see detail in Appendix 1).

Then, the applicability of the EFA was confirmed through Bartlett's sphericity test (2853.3; $gl = 780$; $p < .01$) and the KMO statistic (.77). The initial EFA suggested the retention of 4 factors that explained 41.8% of the total variance. In general, items developed to assess ER, GM and SC tended to cluster together, whereas items developed to assess SS tended to have low and duplicated factor loadings on different factors. Items 1, 37 (developed to assess GM), 10, and 38 (developed to assess ER) were eliminated because they had loadings below .30 on all factors. Only items with higher loadings in the expected factors were kept. Thus, items 11, 31 (developed to evaluate SS), 14, 22 (developed to evaluate ER), 29, 33 (developed to evaluate GM), 12, 16, 20 and 36 (developed to evaluate SC) were eliminated.

A second EFA was performed with the remaining items, which suggested the retention of 4 factors that explained 47.7% of the total variance. One more time, satisfactory loadings were observed for all factors, except for the items developed to assess SS, which tended to present duplicated or low loadings.

Besides, a CFA was carried out with the remaining items. The model fit indices were good ($\chi^2 = 470.54$, $p < .01$; $\chi^2 / df_{(293)} = 1.60$; GFI = .94; AGFI = .93; CFI = .97; NFI = .92; NNFI = .97; RMSEA = .048) and the factor loadings of the final items were also adequate. The results of the EHS-A in its final 26-item version are shown in Figure 1. The internal consistency of the retained factors was adequate (HHSS: .61; GM: .76; ER: .75; SC: .75; total scale: .85). The interfactorial correlations are presented in Table 2 (along with descriptive statistics) and were also adequate. Test-retest reliability was also good (GM: $r = .41$, $p < .01$; ER: $r = .52$, $p < .01$; SS: $r = .60$, $p < .01$; SC: $r = .70$, $p < .01$; total scale: $r = .67$, $p < .01$).

Table 2
Interfactorial correlations and descriptive statistics of the EHS-A dimensions.

	1	2	3	4	5
1. Social skills.	-	.51**	.52**	.28**	.74**
2. Emotional regulation.	-	-	.55**	.47**	.83**
3. Growth mindset.	-	-	-	.40**	.79**
4. Self-control.	-	-	-	-	.72**
5. Total scale.	-	-	-	-	-
ME	3.99	3.68	3.93	3.81	3.86
DE	0.50	0.71	0.64	0.75	0.49

Note. ** $p < .01$.

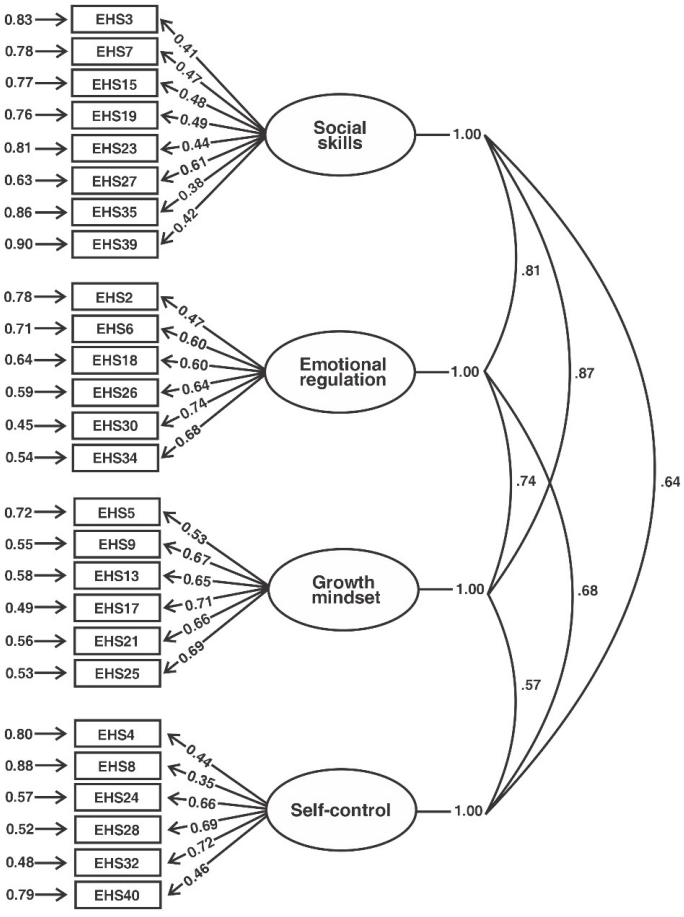


Figure 1
Factorial solution (CFA) for the socioemotional skills model.

Table 3 shows the relationships between the EHS-A dimensions and theoretically related constructs. Low and moderate correlations were observed suggesting that higher SEL skills scores relate to a more frequent use of adaptive ER strategies (CERQ-K) and less difficulties in emotional regulation (DERS). The frequency of use of maladaptive ER strategies (CERQ-K) was not associated with SEL skills.

Finally, the presence of gender and school grade differences was analyzed for the four factors of the scale. Results showed that there were no differences between boys and girls, and no dissimilarities between grades, in any of the four EHS-A factors.

Discussion

Thanks to the contributions of organizations, such as CASEL, and the demonstrated importance of assessing SEL skills in various domains, empirical and theoretical research in this field has significantly increased. However, no instruments for assessing these skills have been developed in Argentina. For this reason, the main objective of this study was to develop and validate a Social-Emotional Skills Scale for Argentine Children

Table 3
Correlations between the EHS-A dimensions, adaptive strategies, and total emotional regulation difficulties.

	1	2	3	4	5	6	7	8
1. Social skills.	-	.51**	.52**	.28**	.74**	.41**	.06	-.25*
2. Emotional regulation.	-	-	.55**	.47**	.83**	.59**	.09	-.41**
3. Growth mindset.	-	-	-	.40**	.79**	.43**	.18	-.27*
4. Self-control.	-	-	-	-	.72**	-.18	.10	-.42**
5. EHS-A total scale .	-	-	-	-	-	-.01	.49**	-.48**
6. Adaptive strategies (CERQ-k).	-	-	-	-	-	-	.45**	-.04
7. Maladaptive Strategies (CERQ-k).	-	-	-	-	-	-	-	.46**
8. Difficulties in emotional regulation (DERS).	-	-	-	-	-	-	-	-

Note. ** $p < .01$; * $p < .05$.

(EHS-A) between 9 to 12 years and to provide evidence of its reliability and validity.

To design and create the scale, other SEL skills scales were used as references, and the operational definitions of the constructs were followed. Consequently, a 55-item scale was initially developed, which, after undergoing expert analysis, was refined to 40 items (10 for each skill, i.e., SS, SC, GM, and emotions and ER). The scale was tested preliminarily with a pilot sample and subsequently with the total sample.

The final version of the EHS-A comprises 26 items, demonstrating satisfactory psychometric properties and enabling the identification of four principal factors that account for 47.7% of the variance. The logical structure of these factors aligns well with the theoretical propositions found in the literature (CASEL, 2015), and the items reflect the essential characteristics of each SEL skill.

The first factor, designated as SS, encompasses items related to the willingness to listen to peers, volunteer to assist others, and seek help when needed skills that contribute to more fulfilling interpersonal relationships (Gresham, 2016). These findings are in line with the principles of social reciprocity, suggesting that the readiness to offer, express gratitude, and receive help fosters connections among children (Bono et al., 2022; Froh et al., 2010).

The second factor, labeled ER, involves the application of strategies and skills to monitor, evaluate, and modify emotions in order to achieve goals and respond appropriately to environmental demands (Garnefski & Kraaij, 2007; Gross, 2014). The items assess understanding of emotion regulation, particularly for unpleasant emotions like anger and sadness, control over emotional reactions, and the use of strategies to modify emotions, such as shifting attention or positive refocusing.

The third factor, named GM, encompasses items reflecting the belief that children possess the ability to develop skills, overcome challenges and learn through perseverance. This factor collectively represents the notion that abilities are not fixed but can expand and improve with time and effort, emphasizing the idea that perseverance and learning from mistakes contribute to overcoming obstacles (Dweck, 2017).

Finally, the last factor, labeled SC, primarily pertains to the process by which children balance long-term goals or norms with more immediately rewarding short-term desires or impulses. The items illustrate the conflict between emotion/desire and the expected behaviors from others, exemplifying the process of regulating behavior, emotion, or cognition to achieve meaningful goals (Duckworth & Steinberg, 2015).

Concerning the internal consistency of the retained factors, it was satisfactory for ER, SCand GM. The factor related to SS exhibited a lower internal consistency. Nevertheless, test-retest reliability was robust for all factors, indicating the reliability of the EHS-A scale. In summary, regarding the analysis of construct validity (objective 2 of the study), the results affirm the relative independence of the various SEL skills, consistent with existing literature (Gresham et al., 2018). It is noteworthy that no differences were observed based on gender or school grade for any dimensions of the EHS-A, aligning with reviews that do not identify gender as a moderator of SEL skills (Durlak et al., 2022; Zamora et al., 2020).

In terms of criterion validity, both GM, SS, and ER exhibited positive associations with adaptive ER strategies (Garnefski et al., 2001). Among the adaptive ER skills, also known as functional or positive-focused, acceptance, distraction, planning, cognitive reappraisal, and perspective-taking stand out (Garnefski et al., 2001). In other words, children with higher GM, good SS, and effective emotion regulation are more inclined to use adaptive strategies frequently when dealing with negative emotions or experiencing a negative mood. It is noteworthy that no relationships were identified between adaptive strategies and the SC dimension, nor between SEL skills and maladaptive ER strategies.

Conversely, it was observed that all the dimensions proposed by EHS-A exhibited negative associations with ER difficulties (DERS). In this context, emotional dysregulation refers to challenges in the ability to regulate or modulate emotions in response to negative situations or events (Gratz & Roemer, 2004). It encompasses non-acceptance of emotional responses, a lack of emotional awareness and clarity, difficulties in engaging in goal-directed behavior, challenges in impulse control, and limited access to effective ER strategies. Consequently, children with higher SEL skill scores demonstrated lower ER difficulties, implying that these skills may play a role in the ability to control and manage emotions (Domitrovich et al., 2017; Eisenberg et al., 2010).

However, it is important to note some limitations of the present study. First, while the analyzed sample size is adequate for the number of items in the EHS-A, a larger number of participants could enhance the empirical results and facilitate the generalization of findings. Additionally, gathering reports from both families and teachers could enrich the assessment and provide a complementary perspective on children's skills.

Finally, it is relevant to comment on the use of self-report tools in children. The literature suggests that children may tend to overestimate their behavioral assessments due to their relative mastery of knowledge about themselves (Molina et al., 2013) and their inclination to respond according to normative patterns of what they consider appropriate or inappropriate (Lemos, 2006). Despite this, the development and use of self-report instruments for SEL skills from the age of 9 present a current area of interest with potential for the future development of measurement instruments. It is recommended that the items should be read one by one to children to enhance their comprehension.

In general terms, the outcomes of the current study signify progress in the availability of instruments for evaluating SEL skills in Argentina.

Additionally, unlike other instruments that concentrate on problematic behaviors, the EHS-A appraises positive or strengths-focused aspects, aligning with the models proposed by CASEL. While future studies are warranted, this work serves as an initial impetus for evaluating SEL skills in the everyday life settings of school-aged children.

Appendix 1

Aiken V Results.

Item	Description	Quality	Relevance	Relevance
MC 1	My intelligence is something I can change.	.95	.93	.95
MC 2	Every day I challenge myself to be smarter.	.85	.9	.93
MC 3	I can learn any subject if I put my mind to it.	.93	.93	.95
MC 4	I can try harder to make things work out for me.	.93	.85	.9
MC 5	I am able to overcome challenges or problems, even if they are difficult.	.85	.78	.88
MC 6	Even if I make a mistake, I know I can start again.	1	1	1
MC 7	I am able to work hard, even when things are difficult.	.93	.93	.95
MC 8	I am confident that I can achieve anything I set my mind to.	.95	.88	.88
MC 9	I like challenges.	.88	.78	.95
MC 10	When I try to do something, I think I'm going to fail.	.93	.95	.95
MC 11	I know I can learn more to be smarter.	.88	.88	.93
AU 1	I do things even though I know they are wrong.	.9	.93	.95
AU 2	I wait until the last minute to do my homework.	1	.98	.98
AU 3	When others speak, I wait my turn.	.95	1	1
AU 4	I can calm down when I am nervous or worried.	.93	.85	.93
AU 5	I can do my homework, even if I don't like it.	.85	.93	.93
AU 6	I complete tasks, even if they seem difficult to me.	.98	.95	.95
AU 7	I do my homework, even if I don't feel like it.	1	.98	.98
AU 8	I can concentrate in class, even if there are things that distract me.	.93	.93	.98
AU 9	At home or at school, I lose my temper.	.78	.83	.85
AU 10	When I am very angry, I talk back to adults.	.98	.98	.98
AU 11	If a colleague insults me, I go and do the same to him/her.	.9	.98	.98
AU 12	Even if I want to play, I do my homework first.	.98	.9	.95
AU 13	I think I am impulsive.	.55	.83	.9
AU 14	I say everything that comes to my mind.	.8	.9	.9
AU 15	If I want something, I find it hard to wait.	.98	.98	1
AU 16	I do things without thinking about the consequences.	.95	.95	.95

Item	Description	Quality	Relevance	Relevance
E1	I can describe my emotions.	.73	.88	.93
E2	I notice when my emotions distract me.	.78	.9	.93
E3	I know what to do to feel better when I am sad.	.98	1	.98
E4	Even if I am angry, I try to treat others well.	1	1	1
E5	When I see someone's face, I realize how they feel.	.95	.93	.93
E6	I talk about my emotions with others.	.98	.95	.98
E7	I can calm down when I am angry.	.95	.95	.95
E8	I get angry when things go wrong.	.9	1	1
E9	When something bad happens to me, I try to think about nicer things.	.95	.95	.95
E10	When something bad happens to me, I try to look on the bright side.	.93	.95	.98
E11	When something bad happens to me, I think about how I can fix it.	.98	1	.93
E12	I know the difference between being sad, scared or angry.	.98	1	.95
E13	I notice when I am tense or nervous.	.85	.95	.95
HHSS1	I care about the feelings of others.	.95	.93	.95
HHSS2	I accept my colleagues, even if they think differently from me.	.93	.88	.83
HHSS3	I like to listen to what my colleagues have to say.	.98	.9	.95
HHSS4	I am alone during recess.	1	.9	.88
HHSS5	I fight with my peers and colleagues.	.9	.9	.95
HHSS6	When I need help, I ask for it.	1	1	1
HHSS7	I am grateful when people do something for me.	1	.98	1
HHSS8	I take care of my colleagues' things as if they were my own.	.93	.83	.83
HHSS9	I call or write to my friends.	.98	.88	.9
HHSS10	When someone needs help, I offer to help.	.98	.98	.98
HHSS11	I get together with classmates (outside of school).	.98	.9	.85
HHSS12	I like to participate in group games.	1	.98	.95
HHSS13	When people are good to me, I am good to them.	.85	.8	.85
HHSS14	I find it easy to make friends.	.98	.98	.98
HHSS15	My friends trust me.	.98	.88	.93

References

Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement, 45*(1), 131-142. https://doi.org/10.1177/0013164485451012

Andrés, M. L. (2014). *Mediating effect of cognitive strategies of emotional regulation in the relationship between personality traits and anxiety, depression and happiness in children aged 9 to 12 years old* (Doctoral dissertation). National University of Mar del Plata, Mar del Plata. http://rpsico.mdp.edu.ar/handle/123456789/538

American Psychological Association. (2010). Ethical prin-

ciples of psychologists and code of conduct. https://www.apa.org

Bono, G., Duffy, T., & Moreno, S. (2022). Gratitude in school: Benefits to students and schools. In K. A. Allen, M. J. Furlong, D. Vella-Brodrick, & S. M. Suldo (Eds.), *Handbook of Positive Psychology in Schools (3rd Ed.)* (pp. 118-134). Routledge.

Canet-Juric, L., García-Coni, A., Andrés, M. L., Vernucci, S., Aydmune, Y., Stelzer, F., & Richard´s, M. M. (2020). Intervención sobre autorregulación cognitiva, conductual y emocional en niños: Una revisión de enfoques basados en procesos y en el currículo escolar, en Argentina [Intervention on cognitive, behavioral and emotional self-regulation in children: A review of process-based and school curriculum approaches in Argentina]. *Revista Argentina de Ciencias del Comportamiento/Argentinean Journal of Behavioral Sciences, 12*(1), 1-25. https://doi.org/10.32348/1852.4206.v12.n1.24999

Cipriano, C., Strambler, M. J., Naples, L. H., Ha, C., Kirk, M., Wood, M., Sehgal, K., Zieher, A. K., Eveleigh, A., McCarthy, M., Funaro, M., Ponnock, A., Chow, J. C., & Durlak, J. (2023). The state of evidence for social and emotional learning: A contemporary meta-analysis of universal school-based SEL interventions. *Child Development, 94*(5), 1181-1204. https://doi.org/10.1111/cdev.13968

Cohen, L., Manion, L., & Morrison, K. (2011). *Research Methods in Education (7th Ed.)*. Routledge.

Collaborative for Academic, Social, and Emotional Learning. (2013). *2013 CASEL Guide. Effective social and emotional learning programs: Preschool and elementary school edition*. KSA-Plus Communications. https://files.eric.ed.gov

Collaborative for Academic, Social, and Emotional Learning. (2015). *2015 CASEL Guide. Effective social and emotional learning programs: Middle and high school edition*. KSA-Plus Communications.

Crowe, L. M., Beauchamp, M. H., Catroppa, C., & Anderson, V. (2011). Social function assessment tools for children and adolescents: A systematic review from 1988 to 2010. *Clinical Psychology Review, 31*(5), 767-785. https://doi.org/10.1016/j.cpr.2011.03.008

del-Valle, M. V., & Zamora, E. V. (2021). El uso de las medidas de auto-informe: Ventajas y limitaciones en la investigación en psicología [The use of self-report measures: Advantages and limitations in psychology research]. *Alternativas en Psicología/Alternatives in Psychology, 47*, 22-35. https://www.alternativas.me

del-Valle, M., Zamora, E. V., Khalil, Y., & Altamirano, M. (2020). Rasgos de personalidad y dificultades de regulación emocional en estudiantes universitarios [Personality traits and emotion dysregulation in university students]. *Psicodebate, 20*(1), 56-67. https://doi.org/10.18682/pd.v20i1.1877

Domitrovich, C. E., Durlak, J. A., Staley, K. C., & Weissberg, R. P. (2017). Social-emotional competence: An essential factor for promoting positive adjustment and reducing risk in school children. *Child Development, 88*(2), 408-416. https://doi.org/10.1111/cdev.12739

Duckworth, A. L., & Steinberg, L. (2015). Unpacking self-control. *Child Development Perspectives, 9*(1), 32-37. https://doi.org/10.1111/cdep.12107

Durlak, J. A., Domitrovich, C. E., Weissberg, R. P., & Gullotta, T. P. (Eds.). (2015). *Handbook of Social and Emotional Learning. Research and Practice*. The Guilford Press.

Durlak, J. A., Mahoney, J. L., & Boyle, A. E. (2022). What we know, and what we need to find out about universal, school-based social and emotional learning programs for children and adolescents: A review of meta-analyses and directions for future research. *Psychological Bulletin, 148*(11-12), 765-782. https://doi.org/10.1037/bul0000383

Dweck, C. S. (2006). *Mindset: The new psychology of success*. Random House.

Dweck, C. S. (2017). The journey to children's mindsets and beyond. *Child Development Perspectives, 11*(2), 139-144. https://doi.org/10.1111/cdep.12225

Eisenberg, N., Spinrad, T. L., & Eggum, N. D. (2010). Emotion-related self-regulation and its relation to

children's maladjustment. *Annual Review of Clinical Psychology, 6*, 495-525. https://doi.org/10.1146/annurev.clinpsy.121208.131208

Ferrando, P. J., & Anguiano-Carrasco, C. (2010). Factor analysis as a research technique in psychology. *Papeles del Psicólogo, 31*(1), 18-33. https://www.papelesdelpsicologo.es

Froh, J. J., Bono, G., & Emmons, R. (2010). Being grateful is beyond good manners: Gratitude and motivation to contribute to society among early adolescents. *Motivation and Emotion, 34*(2), 144-157. https://doi.org/10.1007/s11031-010-9163-z

Garnefski, N., Kraaij, V., & Spinhoven, P. (2001). Negative life events, cognitive emotion regulation and emotional problems. *Personality and Individual Differences, 30*(8), 1311-1327. https://doi.org/10.1016/S0191-8869(00)00113-6

Garnefski, N., & Kraaij, V. (2007). The Cognitive Emotion Regulation Questionnaire: Psychometric features and prospective relationships with depression and anxiety in adults. *European Journal of Psychological Assessment, 23*(3), 141-149. https://doi.org/10.1027/1015-5759.23.3.141

Gehlbach, H., & Hough, H. J. (2018). *Measuring social emotional learning through student surveys in the CORE districts: A pragmatic approach to validity and reliability*. Policy Analysis for California Education. https://files.eric.ed.gov

Gratz, K. L., & Roemer, L. (2004). Multidimensional assessment of emotion regulation and dysregulation: Development, factor structure, and initial validation of the difficulties in emotion regulation scale. *Journal of Psychopathology and Behavioral Assessment*, *26*(1), 41-54. https://doi.org/10.1023/B:JOBA.0000007455.08539.94

Gresham, F. M. (2016). Social skills assessment and intervention for children and youth. *Cambridge Journal of Education, 46*(3), 319-332. https://doi.org/10.1080/0305764X.2016.1195788

Gresham, F. M., & Elliott, S. N. (2008). *Social Skills Improvement System Rating Scales*. Pearson.

Gresham, F. M., Elliott, S. N., Metallo, S., Byrd, S., Wilson, E., & Cassidy, K. (2018). Cross-informant agreement of children's social-emotional skills: An investigation of ratings by teachers, parents, and students from a nationally representative sample. *Psychology in the Schools, 55*(2), 208-223. https://doi.org/10.1002/pits.22101

Gross, J. J. (2014). Emotion regulation: Conceptual and empirical foundations. In J. J. Gross (Ed.), *Handbook of Emotion Regulation (2nd Ed.)* (pp. 3-20). The Guilford Press.

Hernández-Sampieri, R., Fernández-Collado, C., & Baptista-Lucio, P. (2014). Metodología de la investigación *[Research Methodology] (6th Ed.)*. McGraw-Hill.

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika, 30*(2), 179-185. https://doi.org/10.1007/BF02289447

Hu, L.-T., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods, 3*(4), 424-453. https://doi.org/10.1037/1082-989X.3.4.424

Humphrey, N., Kalambouka, A., Wigelsworth, M., Lendrum, A., Deighton, J., & Wolpert, M. (2011). Measures of social and emotional skills for children and young people: A systematic review. *Educational and Psychological Measurement, 71*(4), 617-637. https://doi.org/10.1177/0013164410382896

Huston, A. C., & Ripke, M. N. (2006). Middle childhood: Contexts of development. In A. C. Huston & M. N. Ripke (Eds.). *Developmental contexts in middle childhood: Bridges to adolescence and adulthood* (pp. 1-22). Cambridge University Press. https://doi.org/10.1017/CBO9780511499760.002

Karatzias, T., Shevlin, M., Fyvie, C., Hyland, P., Efthymiadou, E., Wilson, D., Roberts, N., Bisson, J. I., Brewin, C. R., & Cloitre, M. (2016). An initial psychometric assessment of an ICD-11 based measure of PTSD and complex PTSD (ICD-TQ): Evidence of construct validity. *Journal of Anxiety Disorders, 44*,

73-79. https://doi.org/10.1016/j.janxdis.2016.10.009

LeBuffe, P. A., Shapiro, V. B., & Naglieri, J. A. (2009). *The Devereux Student Strengths Assessment (DESS). Assessment, technical manual and user's guide*. Apperson.

Lemos, V. N. (2006). La deseabilidad social en la evaluación de la personalidad infantil [Social desirability in the assessment of children's personality]. *Suma Psicológica, 13*(1), 7-14. https://ri.conicet.gov.ar

Lloret-Segura, S., Ferreres-Traver, A., Hernández-Baeza, A., & Tomás-Marco, I. (2014). The exploratory item factor analysis: A practical guide revised and updated. *Anales de Psicología/Annals of Psychology, 30*(3), 1151-1169. https://doi.org/10.6018/analesps.30.3.199361

Lorenzo-Seva, U., & Ferrando, P. J. (2022). *Factor Analysis* (Version 12.03.02). [Software de cómputo]. https://psico.fcep.urv.cat/utilitats/factor

Medrano, L. A., & Trógolo, M. (2014). Validación de la Escala de Dificultades en la Regulación Emocional en la población universitaria de Córdoba, Argentina [Validation of the Difficulties in Emotion Regulation Scale in the university population of Córdoba, Argentina]. *Universitas Psychologica, 13*(4), 1345-1356. https://doi.org/10.11144/Javeriana.UPSY13-4.vedr

Milicic, N., Alcalay, L., Berger, C., & Torretti, A. (2014). Aprendizaje socioemocional: Programa BASE (Bienestar y aprendizaje socioemocional) como estrategia de desarrollo en el contexto escolar [Socioemotional learning: The BASE program (socioemotional well-being and learning) as an intervention strategy in the school context]. Paidós.

Molina, M. F., Calero, A., & Raimundi, M. J. (2013). Escala de Evaluación del Comportamiento del Niño en su versión para Padres (PRS): Adaptación para su uso en Buenos Aires. *Revista Evaluar, 13*(1), 1-18. https://doi.org/10.35670/1667-4545.v13.n1.6797

Neumann, A., van Lier, P. A. C., Gratz, K. L., & Koot, H. M. (2010). Multidimensional assessment of emotion regulation difficulties in adolescents using the Difficulties in Emotion Regulation Scale. *Assessment, 17*(1), 138-149. https://doi.org/10.1177/1073191109349579

Panayiotou, M., & Humphrey, N. (2018). Mental health difficulties and academic attainment: Evidence for gender-specific developmental cascades in middle childhood. *Development and Psychopathology, 30*(2), 523-538. https://doi.org/10.1017/S095457941700102X

Resolution 2857. (2006). Lineamientos para el comportamiento ético en las ciencias sociales y humanidades [Guidelines for ethical behavior in the social sciences and humanities]. Argentina. National Council for Scientific and Technical Research. https://www.conicet.gov.ar

Schleicher, A. (2018). *World class: How to build a 21st-Century school system*. Organization for Economic Co-operation and Development. https://doi.org/10.1787/9789264300002-en

Scientific Software International. (2006). LISREL (V. 8). [Software de cómputo]. http://www.ssicentral.com

Soto, C. M., & Segovia, J. L. (2009). Intervalos de confianza asimétricos para el índice la validez de contenido: Un programa Visual Basic para la V de Aiken [Confidence intervals for the content validity: A Visual Basic computer program for the Aiken´s V]. *Anales de Psicología/Annals of Psychology, 25*(1), 169-171. https://doi.org/10.6018/analesps

Washoe County School District. (2018). Frequently asked questions about the WCSD social and emotional competency assessment. https://www.washoeschools.net

World Medical Association. (2013). *Declaration of Helsinki - Ethical principles for medical research involving human subjects*. https://www.wma.net

Zamora, E., del-Valle, M., & Andrés, M. L. (2019). Educar en emociones: Los programas de aprendizaje socioemocional (SEL). En S. Vernucci & E. Zamora (Comp.), La ciencia de enseñar: Aportes desde la psicología cognitiva a la educación (1st Ed. pp. 126-141) [Educating in emotions: social-emotional learn-

ing (SEL) programs. In S. Vernucci & E. Zamora (Comp.), *The science of teaching. Contributions from cognitive psychology to education* (1st Ed., pp. 126-141)]. Universidad Nacional de Mar del Plata.

Zamora, E. V., del-Valle, M. V., Galli, J. I., Vernucci, S., Canet-Juric, L., & Richard's, M. M. (2020). Evaluación de impulsividad por múltiples informantes: Escala de Impulsividad de Dominio Específico en niños [Assessment of Impulsivity by Multiple Informants: Domain-Specific Impulsivity Scale for children]. *Revista Evaluar, 20*(2), 35-50. https://doi.org/10.35670/1667-4545.v20.n2.30108

Zamora, E. V., del Valle, M. V., Bruna, O., Olsen, C., & Andrés, M. L. (2022, December 1-3). *Reliability indices in children: Internal consistency analysis in an Emotional Regulation Scale*. IX Congreso Marplatense Internacional de Psicología, Mar del Plata, Argentina.

2024, Vol. 24, No. 1
ISSN 1667-4545
Recuperado de https://revistas.unc.edu.ar/index.php/revaluar

Laboratorio de Evaluación Psicológica y Educativa
Facultad de Psicología - Universidad Nacional de Córdoba

Aplicación de un Modelo Exploratorio de Ecuaciones Estructurales para la Evaluación de la "Children and Adolescents Interpersonal Survey" (CAIS)

Application of an Exploratory Structural Equation Model for the Evaluation of the Children and Adolescents Interpersonal Survey (CAIS)

Malenka Areas * [1], Candela García-Domench [2], Agustin Freiberg-Hoffmann [1], Andrés Roussos [3], Juan Martín Gómez-Penedo [1]

1 - Universidad de Buenos Aires y CONICET, Buenos Aires, Argentina.
2 - Universidad Nacional de La Plata, Argentina.
3 - Universidad Nacional del Comahue y CONICET, Argentina.

Recibido: 30/11/2023 **Revisado:** 0101/2024 **Aceptado:** 21/01/2024

Resumen

La personalidad es un constructo complejo que puede entenderse desde diversos modelos. Uno de los modelos que ha ganado importancia para su interpretación es el modelo interpersonal circumplejo (IPC). La Children and Adolescents Interpersonal Survey (CAIS, por sus siglas en inglés) es un instrumento específicamente desarrollado para el estudio de dicho modelo en niños y adolescentes. Este estudio tiene como objetivo examinar las propiedades psicométricas de la CAIS en una muestra argentina de 353 participantes. Se analizó la confiabilidad y validez del instrumento. Los resultados de este estudio indican una adecuada consistencia interna y homogeneidad de los ítems, así como también evidencia de validez de constructo (mediante un modelo exploratorio de ecuaciones estructurales) y concurrente (mediante correlaciones r de Pearson , otra medida de rasgos interpersonales). De esta manera, se establece así a la versión en español de la CAIS como un instrumento confiable y válido para la evaluación de estilos interpersonales en niños y adolescentes de Argentina.

Palabras clave: *CAIS, propiedades psicométricas, modelo interpersonal circumplejo, personalidad*

Abstract

Personality is a complex construct that can be understood from various models. One of the models that has gained importance for its interpretation is the interpersonal circumplex model (IPC). The Children and Adolescents Interpersonal Survey (CAIS) is an instrument specifically developed for the study of this model in children and adolescents. This study aims to examine the psychometric properties of the CAIS in an Argentinean sample of 353 participants. The reliability and validity of the instrument were analysed. The results of this study indicate adequate internal consistency and item homogeneity, as well as evidence of construct validity (through an exploratory structural equation model) and concurrent validity (through Pearson's r correlations with another measure of interpersonal traits). Thus establishing the Spanish version of the CAIS as a reliable and valid instrument for the assessment of interpersonal styles in children and adolescents in Argentina.

Keywords: *CAIS, psychometric properties, interpersonal circumplex model, personality*

***Correspondencia a:** Malenka Areas. Universidad de Buenos Aires y CONICET. Lavalle 2353, C.P.: C1047AAG, Ciudad Autónoma de Buenos Aires, Argentina. malenka.areas@gmail.com.

Cómo citar este artículo: Areas, M., García-Domench, C., Freiberg-Hoffman, A., Roussos, A., & Gómez-Penedo, M., J., (2024): Aplicación de un Modelo Exploratorio de Ecuaciones Estructurales para la Evaluación de la "Children and Adolescents Interpersonal Survey" (CAIS). *Revista Evaluar, 24*(1), 60-75. Recuperado de https://revistas.unc.edu.ar/index.php/revaluar

Participaron en la edición de este artículo: Débora Camponetti, Florencia Ruiz, Jorge Bruera.

Los autores de este trabajo declaran que no existe conflicto de intereses.

Introducción

Desde la década de 1960, las teorías de la personalidad han sido el foco fundamental en estudios de la psicología. Sin embargo, recién a partir de la década del '90 es posible ver una sistematización no solo en las teorías que explican la personalidad, como por ejemplo, la teoría de los cinco factores, sino también un desarrollo asociado de técnicas que dan cuenta de esta forma de entender la personalidad (Lubinski, 2000). A pesar de estos avances, dichas teorías de la personalidad se centraron desde sus inicios en la posición del adulto.

En el caso de la infancia, han sido múltiples los debates que plantean si el niño ya tiene una personalidad conformada o si la misma está en vías de desarrollo hasta que se establezca en la adultez (Shiner & Caspi, 2003). Los estudios de la personalidad a lo largo de la infancia y la adolescencia fueron una consecuencia lógica de los estudios de la personalidad en el adulto. Concretamente estos trabajos brindan la posibilidad de hacer un seguimiento de vida en relación a cómo es el desarrollo de esa personalidad, cómo se relacionan con problemáticas de salud mental, cuáles son las cualidades en común y cuáles obedecen a fenómenos propios de etapas vitales (Shiner & Caspi, 2003).

Como puede suponerse, la personalidad es un constructo complejo, lo que lleva a que existan diversas formas y modelos para su interpretación y entendimiento. Entre los modelos más reconocidos se encuentra el *modelo de los cinco grandes* (John & Srivastava, 1999; McCrae & Costa, 1999), los modelos de 3 factores como el *sistema de tres factores* de Eysenck (1991), el *modelo de estructura de la personalidad* de Tellegen (1985), el *modelo de temperamento* (Cloninger et al., 1993),entre otros.

Otros modelos que han ganado importancia a lo largo de las últimas décadas son los interpersonales. El modelo de mayor predominancia en este enfoque es el del círculo *interpersonal* de Leary (1957), basado en el trabajo teórico de Sullivan (1953). Este enfoque ha evolucionado hasta convertirse en el modelo *interpersonal circumplejo* (IPC) (Wiggins, 1979, 1982, 1995). El modelo IPC comprende que las conductas interpersonales o estilos interpersonales se pueden ubicar en dos ejes de motivaciones interpersonales: los motivos de agencia y de comunión. Los primeros, se asocian con el poder y la necesidad de la persona de diferenciarse de otros y mostrar dominio en actividades valiosas; por lo tanto, los motivos de agencia se centran en el desempeño individual. Por su parte, los motivos de comunión responden a necesidades de conexión y establecimiento de relaciones íntimas con otros, como así también a la necesidad de formar parte de unidades sociales que trasciendan al individuo (Horowitz, 2004). A su vez, alrededor de estos dos ejes se distribuyen ocho sub-escalas de estilo interpersonal: dominador (PA); competitivo (BC); frío (DE); inhibido socialmente (FG); no asertivo (HI); explotable (JK); auto-sacrificado (LM); e intrusivo (NO) (ver Figura 1).

El modelo IPC ha sido respaldado para el estudio de la personalidad por un gran cuerpo de investigación en adultos, pero el bagaje de información acerca del mismo en niños y adolescentes es significativamente menor (Sodano & Tracey, 2006). Los modelos interpersonales en la niñez -y sobre todo en la adolescencia- pueden ser particularmente interesantes para el campo clínico si consideramos la importancia de los vínculos, el apego en dicha edad y su estrecha relación con los problemas emocionales que surgen en tal etapa (Mónaco et al., 2021). El contar con modelos comunes para el estudio de la personalidad de niños y adultos permite la integración de una gran cantidad de investigaciones sobre rasgos de personalidad (Jensen-Campbell et al., 2002; John & Srivastava, 1999). Al realizar una búsqueda fo-

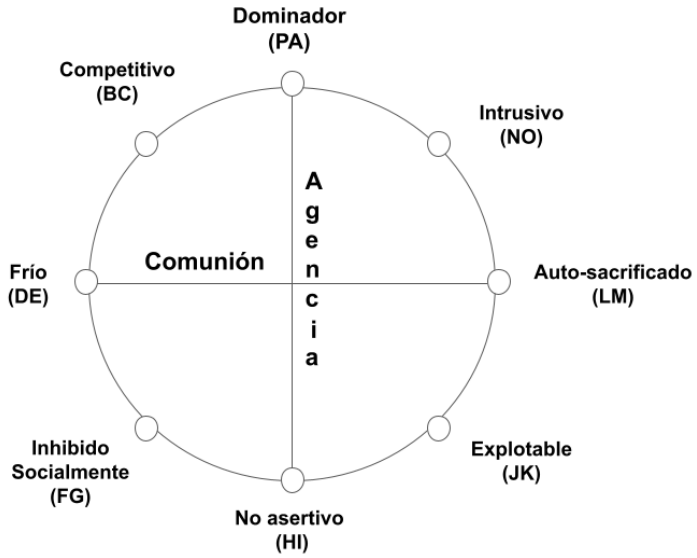


Figura 1. Subescalas y dimensiones del Modelo IPC (Horowitz et al., 2000).

calizada de instrumentos para el estudio de este modelo en niños y adolescentes se encontró que suelen utilizarse herramientas desarrolladas para adultos tales como el inventario de *problemas interpersonales* (Horowitz et al., 2000) o las escalas de *adjetivos interpersonales* (IAS) (Wiggins, 1995). Es por eso que resulta particularmente importante contar con instrumentos para la medida de estos rasgos en niños y adolescentes, ya que el vocabulario de los instrumentos en adultos suele ser más abstracto y complejo (Barbaranelli et al., 2003). Uno de los instrumentos que fue específicamente desarrollado para el estudio del modelo IPC en niños y adolescentes es la *Children and Adolescents Interpersonal Survey* (CAIS) (Sodano & Tracey, 2006).

Children and Adolescents Interpersonal Survey (CAIS)

La CAIS fue desarrollada en Estados Unidos por Sodano y Tracey (2006) con el objetivo de poder contar con un instrumento que explorase y comprendiera el estilo de interacción de niños y adolescentes con otras personas.. La misma se configuró en base a los dos ejes motivacionales del modelo IPC y sus ocho dimensiones [dominador (PA); competitivo (BC); frío (DE);inhibido socialmente (FG); no asertivo (HI); explotable (JK); auto-sacrificado (LM); e intrusivo (NO)], teniendo en cuenta que los ítems fueran acordes a la edad de la población a la cual se dirige. El instrumento cuenta con 37 ítems redactados en forma de afirmaciones donde los participantes indican qué tan a menudo dichas afirmaciones se aplican a ellos usando un formato de respuesta tipo Likert que comprende desde (1) *nunca* a (5) *siempre*. En su versión original en inglés, la validez de constructo para la CAIS fue apoyada por buenos ajustes al modo de orden circular (Sodano & Tracey, 2006). A su vez, se encontraron fuertes correlaciones entre las *escalas de agradabilidad* (LM), *extraversión* (NO) de la CAIS y las escalas del mismo nombre de una medida abreviada del modelo *de los cinco factores* en niños (Sodano & Tracey, 2006). Así como también se encontraron correlaciones fuertes entre algunas de las subescalas octantes correspondientes de la CAIS y el IAS (Wiggins, 1995) en adultos (Sodano & Tracey, 2006). En términos de confiabilidad, la mediana del α de Cronbach de las escalas de octantes fue de

.69 en la investigación original (Sodano & Tracey, 2006) y de .70 en un estudio secundario que utilizó dicho instrumento (Sodano, 2011). Estos resultados preliminares señalan de manera prometedora el desempeño del instrumento. Al momento de realizar el presente trabajo, no se encontraron versiones de dicha encuesta en español que posibilitaran su uso en población hispanohablante.

El presente estudio

Considerando la carencia de instrumentos en español disponibles para el estudio de la personalidad desde una perspectiva IPC, el presente estudio tuvo como objetivo estudiar las propiedades psicométricas de la CAIS en Argentina. Específicamente, se buscó explorar la consistencia interna, la validez del constructo y la validez concurrente de dicho instrumento. A partir de los resultados encontrados, se pretende brindar apoyo tanto a profesionales de la salud mental, como a investigadores de Argentina ofreciendo un instrumento válido y fiable para el análisis del estilo interpersonal desde una perspectiva circumpleja en adolescentes.

Métodos

Participantes

Análisis de evidencia de validez del constructo: participaron 353 adolescentes de Argentina. Del total de participantes, 211 eran mujeres (59.8%), 136 hombres (38.5%) y 6 se identificaron con otros (1.7%). La edad promedio de la muestra fue de 16.4 años (DE = 2.12). Al agrupar los participantes de acuerdo con su edad, se observó que el 19.8% tenía entre 12 y 14 años, el 27.2% tenía entre 15 y 16 años y el 53% tenía entre 17 y 19 años. Todos los jóvenes eran estudiantes. El 81.3% de la muestra pertenecía a la Provincia de Buenos Aires. El 18.7% restante se distribuyó en un 3.6% de adolescentes del norte del país, 2.2% un del centro del país y un 12.9% de adolescentes del sur del país. En el momento de la recolección de los datos, el 27.8% de los participantes asistía a un tratamiento psicoterapéutico, el 44.1% nunca había realizado psicoterapia y el 28.1% había realizado una psicoterapia en el último año.

Análisis de evidencia de validez concurrente: participaron 352 adolescentes. Del total de participantes, 275 eran mujeres (78.1%), 74 hombres (21%) y 3 se identificaron con otros (0.9%). Al igual que en la primera muestra, la mayoría de los participantes pertenecían a la Provincia de Buenos Aires (95.2%). En el momento de la recolección de datos, el 27.6% de los participantes asistía a un tratamiento psicoterapéutico, el 42% nunca había realizado psicoterapia y el 30.4% había realizado una psicoterapia en el último año.

Instrumentos

Children and Adolescents Interpersonal Survey (Sodano & Tracey, 2006). La CAIS es un instrumento basado en el modelo IPC y explora las características interpersonales de niños y adolescentes . Los participantes indican qué tan a menudo se aplican a ellos las afirmaciones a través de un formato de respuesta tipo Likert que comprende desde (1) *nunca* a (5) *siempre*. Los ítems conforman ocho subescalas de estilos interpersonales que están basadas en los octantes del modelo IPC: dominador (o controlador) (PA); competitivo (o auto-centrado) (BC); frío (o distante) (DE); inhibido socialmente (o evitativo) (FG); no asertivo (o sumiso) (HI); explotable (o acomodaticio) (JK); auto-sacrificado (o sobre protector) (LM); e intrusivo (o demandante) (NO). Algunos de los ejemplos de los ítems de la CAIS para cada subescala

son: *soy una persona dificil* (PA), *invento apodos molestos* (BC), *lastimo a la gente* (DE), *paso mucho tiempo solo* (FG), *soy tímido* (HI), *soy tranquilo* (JK), *soy amable con los otros* (LM), *estar conmigo es divertido* (NO). Adicionalmente, es posible calcular las dos dimensiones interpersonales de dominancia y afiliación con base en las distintas combinaciones de las ocho subescalas. Las propiedades psicométricas de la versión original han sido presentadas en la introducción del presente trabajo.

Inventario de Problemas Interpersonales (IIP-32; Horowitz et al., 2000). La versión del IIP de 32 ítems (Horowitz et al.,1988) presenta 32 afirmaciones con respecto a problemas que la gente podría experimentar al interactuar con otros. Los primeros 20 ítems empiezan con *es difícil para mí* -por ejemplo- *decir que no a otras personas*. Los restantes 12 ítems empiezan con *cosas que hago mucho* -por ejemplo- *me abro demasiado a la gente*. Los sujetos indican qué tan de acuerdo se sienten con las afirmaciones en una escala de tipo Likert que va desde 0 (*Nada*) a 4 (*Mucho*). Los 32 ítems conforman ocho subescalas de problemas interpersonales que están basadas en los octantes del modelo interpersonal circumplejo: dominador (o controlador) (PA); competitivo (o auto-centrado) (BC); frío (o distante) (DE); inhibido socialmente (o evitativo) (FG); no asertivo (o sumiso) (HI); explotable (o acomodaticio) (JK); auto-sacrificado (o sobre protector) (LM); e intrusivo (o demandante) (NO). Adicionalmente, las dos dimensiones interpersonales de dominancia y afiliación pueden ser calculadas teniendo en cuenta distintas combinaciones de las ocho subescalas. El IIP-64 tiene una adaptación argentina (Maristany, 2005), la cual fue adaptada a la versión de 32 ítems (Gómez Penedo et al., 2022). Esta versión ha reportado buena fiabilidad con coeficientes alfa en el rango de .67 - .87 y es la utilizada en este estudio (Areas et al., 2023).

Cuestionario Sociodemográfico ad hoc. Un cuestionario ad hoc fue incluido para medir las variables sociodemográficas en este estudio. Este incluyó: la edad del participante, el género, la provincia de residencia, y una pregunta acerca de si estaban asistiendo a psicoterapia al momento de realizar el inventario, si habían asistido a psicoterapia en el último año o si nunca habían asistido a psicoterapia.

Procedimiento

Considerando que el idioma original de la encuesta es el inglés, la primera etapa del proceso de adaptación correspondió a la traducción del instrumento. Se buscó generar una versión que se correspondiera con el idioma local y los estilos lingüísticos propios del contexto en el que se aplicaría. Se realizó una evaluación lingüística de los ítems adecuando los mismos a los modismos locales de la población que sería objeto de estudio. Se utilizó el método de traducción inversa o *backward translation*, la cual implicó una traducción inicial del idioma original al español, y luego una traducción al idioma inglés realizada por psicólogos bilingües, comparándose posteriormente las dos versiones para determinar su grado de equivalencia. Dicho trabajo fue realizado por dos estudiantes avanzados de la carrera de psicología y supervisado por la primera autora del presente trabajo. Una vez determinada la versión final, se realizó una prueba piloto cualitativa donde se le pidió a adolescentes entre 12 y 18 años que respondieran la encuesta y a su vez explicaran qué creían que significaban los ítems. En dicho paso no se encontraron dificultades para la comprensión de los ítems traducidos.

Una vez determinada la versión final de la encuesta, se convocó a participantes mediante un diseño de *bola de nieve*. Como primer paso,

se contactaron centros de psicoterapia, maestros de escuelas y sujetos conocidos en todo el país y fueron invitados a participar. Si estas personas adultas estaban interesadas en colaborar con el estudio, se les enviaba un enlace de la plataforma de *SurveyMonkey*® para compartirlo con adolescentes interesados en participar. Los adolescentes accedían al link donde se encontraban, en primer lugar, con un formulario de consentimiento para firmar, los participantes más jóvenes (menores de 18 años) necesitaban previamente el consentimiento de sus padres. Este consentimiento era requerido por los profesionales involucrados en la investigación antes de compartir el link. El consentimiento también aclaraba que la participación era voluntaria y anónima. Una vez aceptada la participación, el adolescente completaba la CAIS y la información sociodemográfica. Los participantes podían cambiar sus respuestas y volver atrás si así fuera necesario hasta presionar el botón de finalización. En un segundo paso se recolectaron otros 352 casos, se hizo una nueva administración de las técnicas de la CAIS junto al IIP-32 y se correlacionaron para analizar validez concurrente. Asimismo, se aclaraba que la información obtenida no podría identificar al participante y que la plataforma cumplía con los estándares de seguridad internacional para cuidar los datos y el anonimato del participante. Finalmente, es importante informar que este estudio es parte de uno más grande que se enfoca en las características interpersonales en adolescentes. Este estudio más amplio recibió la aprobación ética del Comité de Conductas Responsables en Investigación de la Universidad de Buenos Aires.

Análisis estadístico

Todos los análisis de este trabajo se realizaron con el software libre R (RStudio Team, 2020) y Mplus 8 (Muthén & Muthén, 2017). Las puntuaciones medias de los ítems se utilizaron para calcular las puntuaciones totales y de subescala.

Para estudiar la fiabilidad del instrumento se utilizaron medidas de consistencia interna y homogeneidad de los ítems. Por un lado, se evaluó la consistencia interna de la CAIS calculando el α de Cronbach, el alfa ordinal y el coeficiente omega de McDonald para cada una de las subescalas. Si bien el α de Cronbach es la medida de consistencia interna más utilizada en la literatura, diversos estudios sugieren que no es la mejor estrategia cuando los ítems tienen una escala Likert con menos de siete opciones de respuesta (Freiberg-Hoffmann et al., 2013; Gadermann et al., 2012). En dichos casos se recomienda la utilización del alfa ordinal, una medida basada en las correlaciones policóricas (Elosua & Zumbo, 2008; Zumbo et al., 2007). El estado del arte sugiere que un valor aceptable de α de Cronbach y ordinal oscila entre .70 y .90 (Tavakol & Dennick, 2011). Valores más bajos representan una consistencia interna cuestionable; mientras que, valores más altos implicarían que algunos ítems sean redundantes representando la misma pregunta con diferente formato (Tavakol & Dennick, 2011). Por otro lado, para evaluar la homogeneidad de los ítems, se analizaron las correlaciones ítem-total corregidas, es decir, la correlación promedio de cada ítem con el resto de los ítems de la escala, sin incluir dicho ítem. En este caso, se sugieren valores de correlación ítem-total en el rango de .30 a .80 (Rattray & Jones, 2007). Correlaciones más débiles de .30 implican una escasa homogeneidad del ítem con el resto de los reactivos; mientras que, correlaciones superiores a .80 señalan que el ítem sería repetitivo e innecesario.

Para el estudio de la evidencia de validez de constructo del instrumento, se procedió a examinar su estructura interna mediante un modelo exploratorio de ecuaciones estructurales (ESEM

-por sus siglas en inglés). Un enfoque ESEM estima libremente todas las cargas cruzadas rotadas entre indicadores y factores latentes. Además, el ESEM ofrece las mismas ventajas que el *análisis factorial confirmatorio* (AFC) en términos de índices de ajuste, errores estándar y pruebas de significación. Por lo tanto, la flexibilidad del marco ESEM proporciona una sinergia entre AFC y el *análisis factorial exploratorio* (AFE) (Guay et al., 2015). El uso del ESEM en la evaluación de modelos de medidas de personalidad es común en diversos estudios. Esto se debe a que, en general, el ESEM tiende a exhibir índices de ajuste superiores y un mayor grado de diferenciación entre los factores, entre otras ventajas (Dominguez-Lara & Merino-Soto, 2018; Trógolo et al., 2022).

Para la evaluación del ajuste del modelo en el ESEM, se utilizaron diversas medidas de bondad de ajuste tales como el *Comparative Fit Index* (CFI), *Tucker-Lewis Index* (TLI) y *Root Mean Square Error of Approximation* (RMSEA). Como indicadores de buen ajuste del modelo, la literatura suele recomendar valores de CFI y TLI superiores a .90 como adecuados, mayores a .95 como óptimos, y valores de RMSEA inferiores a .08 (Schumacker & Lomax, 2016). Por otra parte, para los modelos de AFC, se consideraron cargas factoriales aceptables aquellas iguales o por encima de .40. Se utilizó el estimador de mínimos cuadrados ponderados robustos (WLSMV, por sus siglás en inglés) que ha presentado superioridad al estudiar ítems que presentan un nivel de medición ordinal (Li, 2016; Mindrila, 2010). Los análisis se realizaron sobre la totalidad de los casos, ya que no faltaban datos y no hubo que aplicar métodos de imputación de datos.

Para el estudio de la evidencia de validez concurrente se analizaron las correlaciones de Pearson entre el IIP-32 (Horowitz et al., 2000) y la CAIS. Las correlaciones deberían oscilar entre .50 y .85, indicando asociación, pero, a su vez, independencia de los constructos (Rial-Boubeta et al., 2006). El pre-registro del plan de análisis fue cargado en la plataforma osf.io (https://osf.io/hq76r/).

Resultados

Análisis descriptivos

En la Tabla 1 se presentan los análisis descriptivos de los ítems de la CAIS. Asimismo, en la Tabla 2 se reportan las medias y desvío estándar de los puntajes de las subescalas del CAIS.

Confiabilidad

Consistencia interna. Los α de Cronbach de las distintas subescalas oscilaron entre .34 y .74. Considerando que los ítems del instrumento tienen una naturaleza ordinal y únicamente siete categorías de respuestas, además del α de Cronbach se calculó también el alfa ordinal, una medida de consistencia interna menos difundida, pero considerada más adecuada para este tipo de reactivos (Elosua-Oliden & Zumbo, 2008; Zumbo et al., 2007). Los alfa ordinales de las subescalas de la CAIS oscilaron entre .33 y .79. Además, como medida alternativa se analizaron los omega de Mcdonald de las distintas subescalas que oscilaron entre .48 y .75. Los valores detallados para cada subescala pueden observarse en la Tabla 3.

Homogeneidad de ítems. Las correlaciones ítem-total corregidas de los reactivos de la CAIS se presentan en la Tabla 1. Como es posible observar, los ítems 1 (i.e., *soy difícil*, subescala PA) y 22 (i.e., *sé muy poco*, subescala JK) presentaron una correlación ítem-total corregida por debajo de .30. Por otra parte, ningún ítem obtuvo correlaciones ítem-total corregidas por encima de .85.

Tabla 1
Análisis descriptivos y correlaciones ítem-escala corregidas de los ítems de la CAIS.

Ítem	M	DE	Escala	*r*
Soy una persona difícil	2.81	.85	PA	-.04
Invento apodos molestos	1.77	.96	BC	.36
Lastimo a la gente	1.81	.84	DE	.47
Paso mucho tiempo solo	3.03	1.02	FG	.55
Soy tímido	2.80	1.11	HI	.32
Soy tranquilo	3.36	1.02	JK	.29
Soy amable con los otros	4.22	.69	LM	.51
Estar conmigo es divertido	3.73	.86	NO	.36
Sé mucho	3.17	.86	PA	.32
Me gusta hacer lío	2.09	1.06	BC	.40
Hago llorar a la gente	1.36	.65	DE	.47
Estoy solo	2.57	1.14	FG	.55
Estoy triste	2.89	.97	HI	.34
Soy callado	2.54	1.09	JK	.22
Intento ayudar a los demás a que se sientan mejor	4.19	.82	LM	.54
Estoy feliz	3.43	.86	NO	.40
Creo que puedo hacer mucho	3.47	1.03	PA	.31
Engaño a la gente	1.72	.89	BC	.38
Soy malo con los demás	1.48	.66	DE	.52
Soy difícil de conocer	2.61	1.17	FG	.31
Sé muy poco	2.27	.94	HI	.30
Engañar a la gente es malo	4.08	1.02	JK	.03
Soy amigable	4.06	.80	LM	.39
Soy generoso	3.96	.80	NO	.28
Me hago valer	3.39	1.08	PA	.29
Le digo a la gente lo que tiene que hacer	2.53	1.01	BC	.27
Me gusta cuando otros se sienten mal	1.17	.49	DE	.29
Juego solo	2.48	1.16	FG	.37
Me rindo rápido	2.45	1.07	HI	.39
Ayudo a la gente	3.96	.81	LM	.60
Juego con otros	3.33	1.12	NO	.41
Creo que tengo razón	3.42	.91	PA	.31
Soy tramposo	1.94	.99	BC	.41
Soy gruñón	2.77	1.14	DE	.33
Soy miedoso	3.03	1.14	HI	.35
Comparto	3.94	.89	LM	.42
Tengo muchos amigos	3.09	1.17	NO	.47

Nota. M = Media; DE = Desvío Estándar.

Tabla 2
Análisis Descriptivos de las Escalas del CAIS.

Escalas	M	DE	Rango
PA	3.25	.53	[1.8-3.8]
BC	2.01	.61	[1-4.8]
DE	1.72	.50	[1-4]
FG	2.67	.79	[1-4.75]
HI	2.69	.64	[1-4.8]
JK	3.33	.68	[1-5]
LM	4.07	.56	[1.6-5]
NO	3.51	.61	[1.8-5]

Nota. M = Media; DE = Desvío Estándar.

Tabla 3
Alfa de Cronbach, alfa Ordinal y omega de McDonald para las escalas de la CAIS.

Escalas	alfa de Cronbach	alfa Ordinal	omega de McDonald
PA	.44	.47	.48
BC	.64	.67	.65
DE	.66	.77	.70
FG	.65	.70	.68
HI	.62	.62	.62
JK	.34	.33	.51
LM	.74	.79	.75
NO	.63	.67	.64

Validez

Análisis de la estructura interna. *El* ESEM verificó un ajuste adecuado a los datos empíricos con índices CFI de .96, TLI de .93 y RMSEA [CI 90%] de .037 [.031, .043]. La Tabla 4 muestra los parámetros del ESEM y el coeficiente de determinación para cada ítem.

Validez Concurrente. Las correlaciones de Pearson mostraron asociaciones directas significativas de la CAIS con el IIP-32. En casi todos los casos, las correlaciones demostraron un pequeño tamaño del efecto en las asociaciones [(r = .10 - .44), p < .001]. La Tabla 5 presenta los detalles de estas correlaciones.

Discusión

El presente estudio tuvo como objetivo estudiar las propiedades psicométricas de la CAIS en una muestra argentina. Los análisis incluyeron el estudio de la confiabilidad mediante el estudio de la consistencia interna de las escalas, y de la estructura del instrumento a través de un ESEM. Los resultados de este estudio arrojaron evidencias a favor de la confiabilidad y validez del instrumento en población argentina.

Tabla 4

Parámetros y Coeficientes de Determinación del ESEM.

		BC			DE			FG			HI			JK			LM			NO			PA		
		λ	S.E.	R^2	λ	S.E.	R^2	λ	S.E.	R^2	λ	S.E.	R^2	λ	S.E.	R^2	λ	S.E.	R^2	λ	S.E.	R^2	λ	S.E.	R^2
BC	CAIS2	.52	.08	.36	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	CAIS10	.43	.12	.45	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	CAIS18	.62	.08	.56	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	CAIS26	.16	.15	.46	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	CAIS33	.47	.09	.40	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
DE	CAIS3	.01	.06	-	.87	.05	.69	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	CAIS11	.31	.08	-	.73	.09	.65	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	CAIS19	.22	.17	-	.52	.11	.67	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	CAIS27	.37	.11	-	.20	.11	.53	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	CAIS34	.06	.09	-	.25	.08	.46	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
FG	CAIS4	-.06	.09	-	-.03	.05	-	.62	.06	.54	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	CAIS12	-.03	.06	-	.17	.06	-	.57	.07	.60	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	CAIS20	.02	.08	-	.30	.06	-	.09	.08	.33	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
	CAIS28	.18	.06	-	-.13	.06	-	.62	.07	.41	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HI	CAIS5	.03	.07	-	-.13	.08	-	.23	.06	-	.35	.18	.70	-	-	-	-	-	-	-	-	-	-	-	-
	CAIS13	-.19	.06	-	.22	.05	-	.30	.08	-	.45	.10	.57	-	-	-	-	-	-	-	-	-	-	-	-
	CAIS21	.12	.16	-	.08	.14	-	.12	.07	-	.31	.39	.53	-	-	-	-	-	-	-	-	-	-	-	-
	CAIS29	.13	.06	-	-.09	.06	-	.21	.07	-	.56	.07	.41	-	-	-	-	-	-	-	-	-	-	-	-
	CAIS35	-.00	.09	-	.09	.07	-	.05	.06	-	.37	.07	.25	-	-	-	-	-	-	-	-	-	-	-	-
JK	CAIS6	-.09	.05	-	-.13	.07	-	.31	.08	-	-.35	.10	-	.35	.13	.42	-	-	-	-	-	-	-	-	-
	CAIS14	-.14	.05	-	-.08	.05	-	.19	.05	-	.19	.12	-	.71	.05	.75	-	-	-	-	-	-	-	-	-
	CAIS22	-.17	.11	-	-.02	.10	-	-.05	.08	-	.14	.15	-	-.05	.08	.22	-	-	-	-	-	-	-	-	-
LM	CAIS7	.02	.10	-	-.02	.10	-	-.03	.07	-	-.06	.14	-	.06	.06	-	.65	.09	.56	-	-	-	-	-	-
	CAIS15	-.12	.07	-	.05	.06	-	.05	.07	-	.11	.10	-	-.12	.05	-	.73	.05	.52	-	-	-	-	-	-
	CAIS23	.16	.06	-	-.14	.05	-	-.10	.06	-	-.09	.10	-	-.28	.05	-	.50	.06	.57	-	-	-	-	-	-
	CAIS30	-.12	.05	-	.05	.05	-	.01	.05	-	.04	.06	-	-.01	.05	-	.81	.05	.64	-	-	-	-	-	-
	CAIS36	-.06	.07	-	-.02	.07	-	-.10	.06	-	.11	.11	-	.27	.05	-	.67	.05	.51	-	-	-	-	-	-
NO	CAIS8	.22	.06	-	-.24	.07	-	-.12	.07	-	.07	.09	-	-.31	.06	-	.16	.06	-	-.06	.16	.35	-	-	-
	CAIS16	.33	.07	-	-.20	.06	-	-.25	.07	-	-.04	.09	-	.15	.08	-	.11	.06	-	.01	.38	.57	-	-	-
	CAIS24	.17	.09	-	-.21	.11	-	-.06	.07	-	.11	.16	-	-.00	.05	-	.65	.10	-	-.29	.21	.70	-	-	-
	CAIS31	.17	.09	-	-.12	.07	-	-.18	.06	-	-.09	.10	-	-.01	.05	-	.27	.06	-	.14	.20	.29	-	-	-
	CAIS37	.19	.15	-	-.24	.12	-	-.19	.08	-	-.04	.25	-	-.12	.05	-	.26	.08	-	.38	.23	.51	-	-	-
PA	CAIS1	-.08	.08	-	.54	.06	-	-.09	.07	-	.27	.11	-	.18	.10	-	.09	.05	-	-.08	.30	-	.22	.11	.46
	CAIS9	.00	.09	-	-.09	.08	-	.14	.05	-	-.28	.27	-	-.11	.05	-	.00	.07	-	-.17	.09	-	.65	.12	.57
	CAIS17	.00	.08	-	.00	.07	-	.06	.07	-	-.46	.07	-	.02	.09	-	.22	.06	-	.13	.24	-	.21	.17	.35

	BC			DE			FG			HI			JK			LM			NO			PA		
	λ	S.E.	R²	λ	S.E.	R²	λ	S.E.	R²	λ	S.E.	R²	λ	S.E.	R²	λ	S.E.	R²	λ	S.E.	R²	λ	S.E.	R²
CAIS25	.06	.06	-	-.08	.06	-	-.19	.06	-	-.27	.10	-	.10	.05	-	.13	.06	-	.02	.17	-	.36	.08	.38
CAIS32	.23	.09	-	.04	.07	-	.11	.06	-	.04	.07	-	-.03	.06	-	.00	.06	-	.15	.06	-	.54	.06	.50

Tabla 5
Correlaciones entre las subescalas de la CAIS y el IIP-32.

	IIP_PA	IIP_BC	IIP_DE	IIP_FG	IIP_HI	IIP_JK	IIP_LM	IIP_NO
CAIS_PA	.25*							
CAIS_BC		.15*						
CAIS_DE			.17*					
CAIS_FG				.44*				
CAIS_HI					.36*			
CAIS_JK						.19*		
CAIS_LM							.29*	
CAIS_NO								.10*

Nota. N = 352. Correlaciones * = $p < .001$.

En relación con la confiabilidad, se puede observar que algunas de las subescalas mostraron niveles de consistencia interna aceptables y otras cuestionables, esto podría deberse a que dentro de cada subescala el contenido de los ítems podría no ser homogéneo. Los α de Cronbach y ordinales oscilaron entre .33 - .79, y los omega de McDonald entre .48 - .75, evidenciando una consistencia aceptable en la mayoría de las subescalas y moderada en otras (Tavakol & Dennick, 2011). Particularmente, la subescala JK obtuvo los niveles de consistencia interna más bajos. La misma contiene tan sólo 3 ítems, por lo cual es esperable que al disminuir la cantidad de ítems dentro de la subescala también bajen los índices de consistencia interna. A su vez, el contenido de los ítems representativos de esta dimensión evalúa aspectos diferentes. Esas diferencias pueden observarse en estos índices, ya que los ítems serían heterogéneos y poco redundantes. Además, en concordancia con la bibliografía previa, esta escala fue también la de menor consistencia interna en el trabajo original realizado por Sodano y Tracey (2006), donde obtuvo un α de Cronbach de .32. Al mismo tiempo, esta subescala ha mostrado tener consistencia baja en otros instrumentos de estructura similar, como el IIP desarrollado por Horowitz et al. (2000) (Alden et al., 1990; Areas et al., 2023; Bailey et al., 2018; Gómez-Penedo et al., 2022).

Adicionalmente, al analizar la homogeneidad de los ítems como parte de la evaluación de la consistencia interna, todos los ítems, a excepción del 1 (*soy difícil*, subescala PA) y el 22 (*engañara la gente es malo*, subescala JK), presentaron correlaciones ítem-total en el rango sugerido y por debajo de .85, lo cual indica que los ítems no eran repetitivos ni redundantes (Rattray & Jones, 2007). Respecto al ítem 1, si se compara su contenido conceptual, puede observarse que el mismo tiene una orientación más negativa que los otros ítems incluidos en la subescala, lo cual podría explicar dicho resultado. En relación con el ítem 2, esta subescala cuenta solo con 3 ítems (ítem 6,

ítem 14 e ítem 22), mientras que, los dos ítems restantes son descripciones de la personalidad y no una creencia, por lo que también esto podría explicar el resultado. Asimismo, este resultado podría explicar la baja consistencia interna identificada en la subescala JK, así como la observación de que los coeficientes omega de McDonald para esta subescala mostraron un ajuste más elevado en comparación con la medida α de Cronbach.

En relación con el estudio de la estructura del instrumento, los resultados del ESEM mostraron un ajuste adecuado de la estructura interna. Todas las medidas de bondad de ajuste (CFI, TLI, RMSEA) indicaron un buen ajuste dentro de los parámetros sugeridos en la literatura (CFI y TLI superiores a .90 como adecuados, y mayores a .95 como óptimos, y valores de RMSEA inferiores a .08) (Schumacker & Lomax, 2016).

Por último, las correlaciones de Pearson fueron significativas entre las distintas subescalas de la CAIS y la IIP-32 con una relación baja en algunas de las escalas. Esto podría explicarse porque el IIP-32 mide los problemas interpersonales, mientras que, la CAIS se centra en los rasgos interpersonales desde una perspectiva no conflictiva. Por lo tanto, la naturaleza contrastada de estas conceptualizaciones contribuye a las diferencias esperadas en los resultados. Además, puede haber otras variables como la variabilidad de los datos, la falta de linealidad y las características de la muestra que pueden afectar al tamaño de una correlación de Pearson (Goodwin & Leech, 2006).

Este estudio es el primero en analizar el uso de una escala basada en el modelo IPC,creado especialmente para niños y adolescentes en Latinoamérica. Poder contar con instrumentos válidos para su uso en esta población brinda herramientas al campo de la psicología, la cual considera la importancia de la personalidad desde una perspectiva relacional en el campo de la salud mental. Esto puede ser particularmente útil en el trabajo psicológico si se considera el rol crítico de las relaciones en dicha etapa de la vida (DiRico et al., 2016; Inglés et al., 2005) y el impacto de tales variables relacionales en la salud mental y el trabajo en psicoterapia (Auerbach et al., 2014; Moreno-Peral et al., 2020; Taubner et al., 2023).

Resulta importante mencionar que el presente trabajo cuenta con numerosas limitaciones que deberán considerarse en futuras investigaciones. En primer lugar, el presente trabajo se enfocó particularmente en el estudio de la confiabilidad y validez del constructo, es decir que no se realizaron otros análisis de relevancia para el estudio del instrumento. Los posibles futuros trabajos podrían beneficiarse al incorporar otros tipos de análisis como la validez de criterio, concurrente o discriminante, entre otros. Estos resultados podrían proveer información valiosa sobre el uso del instrumento. En segundo lugar, la muestra se recolectó de forma no probabilística, lo cual afecta la posibilidad de generalizar los resultados. Además, dicha metodología para la recolección de datos no permite conocer información sobre las características del paciente que podrían ser de relevancia para el ámbito de la salud mental, por ejemplo: posibles diagnósticos psiquiátricos, situación socio-económica, núcleo familiar, entre otras cuestiones. Contar con dicha información podría resultar útil para conocer si niños y adolescentes con ciertos estilos interpersonales condicen con presencia de patologías o cuestiones que puedan resultar problemáticas para el desarrollo, o inclusive, identificar si los estilos interpersonales varían a lo largo de la niñez/adolescencia. En esta línea, es importante destacar que la gran mayoría de la muestra fue recolectada en Buenos Aires, lo cual limita también la generalización de dichos resultados a otras provincias del país. En tercer lugar, como se señaló anteriormente, algunas de las subescalas del instrumento revelaron un nivel de consistencia interna bajo.

Investigaciones futuras que se centren en el comportamiento de estas subescalas podrían proporcionar información valiosa para perfeccionar los instrumentos de evaluación. Acciones tales como explorar nuevos ítems para ser incorporados a nuevas versiones de la encuesta podrían contribuir a la mejora de la consistencia interna y aportar valor al instrumento. Además, resulta relevante destacar que la ausencia de estudios previos sobre el tema plantea un desafío, ya que dificulta la capacidad de comparar y contextualizar los resultados con la información existente.

Finalmente, es importante destacar que no se aplicaron medidas repetidas del instrumento, que podrían ayudar con el estudio de la sensibilidad del instrumento al cambio, por lo que se desconoce dicha sensibilidad.

A pesar de dichas limitaciones, los resultados de este estudio establecen a la CAIS como una herramienta confiable y válida para su implementación con niños y adolescentes en Argentina. Tanto los coeficientes de consistencia interna, como la homogeneidad de ítems y el ESEM, evidenciaron parámetros adecuados de acuerdo a la literatura. En consecuencia, se puede afirmar que la misma es un recurso importante para estudios clínicos e investigaciones en psicología y que, resulta necesario continuar estudiando su uso en esta población.

Referencias

Alden, L. E., Wiggins, J. S., & Pincus, A. L. (1990). Construction of circumplex scales for the Inventory of Interpersonal Problems. *Journal of personality assessment, 55*(3-4), 521–536. https://doi.org/10.1080/00223891.1990.9674088

Areas, M., García-Domench, C., Babl, A., Freiberg-Hoffmann, A., Piazzese Spratte, S., Roussos, A., & Gómez-Penedo, J.M. (2023). Inventory of Interpersonal Problems-32 (IIP-32): Psychometric Properties in Adolescents from Argentina. *Cogent Mental Health, 2*(1), 2267590. https://doi.org/10.1080/28324765.2023.2267590

Auerbach, R. P., Ho, M. H. R., & Kim, J. C. (2014). Identifying cognitive and interpersonal predictors of adolescent depression. *Journal of Abnormal Child Psychology, 42*(6), 913–924. https://doi.org/10.1007/s10802-013-9845-6

Bailey, C., Abate, A., Sharp, C., & Venta, A. (2018). Psychometric evaluation of the Inventory of Interpersonal Problems 32. *Bulletin of the Menninger Clinic, 82*(2), 93–113. https://doi.org/10.1521/bumc.2018.82.2.93

Barbaranelli, C., Caprara, G. V., Rabasca, A., & Pastorelli, C. (2003). A questionnaire for measuring the big five in late childhood. *Personality and Individual Differences, 34*, 645–664. https://doi.org/10.1016/S0191-8869(02)00051-X

Cloninger, C.R., Svrakic, D.M., & Przybeck, T.R. (1993). A psychobiological model of temperament and character. *Archives of General Psychiatry, 39*, 1242–1247.

DiRico, E., Paternain, N., Portillo, N., & Galarza, A. L. (2016). Análisis de la relación entre factores interpersonales y riesgo suicida en adolescentes de la ciudad de Necochea. *Perspectivas en Psicología: Revista de Psicología y Ciencias Afines, 13*(2), 95–106. Análisis de la relación entre factores interpersonales y riesgo suicida en adolescentes de la ciudad de Necochea (redalyc.org)

Dominguez-Lara, S., & Merino-Soto, C. (2018). Estructura interna del BFI-10P y BFI-15P: Un estudio complementario con enfoque CFA y ESEM. *Revista Argentina de Ciencias del Comportamiento, 10*(3), 22-34. Recuperado de https://revistas.unc.edu.ar/index.php/racc/article/view/21037

Elosua, O. P., & Zumbo, B. D. (2008). Coeficientes de fiabilidad para escalas de respuesta categórica ordenada. *Psicothema, 20*(4), 896-901. Redalyc.Coeficientes de fiabilidad para escalas de respuesta categórica ordenada

Eysenck, H.J. (1991). Dimensions of personality: 16, 5 or 3? Criteria for a taxonomic paradigm. *Personality and Individual Differences, 12*, 773–790.

Freiberg-Hoffmann, A., Stover, J. B., de la Iglesia, G., & Fernández Liporace, M. (2013). Correlaciones policóricas y tetracóricas en estudios factoriales exploratorios y confirmatorios. *Ciencias Psicológicas, 7*(2), 151-164. https://doi.org/10.22235/cp.v7i1.1057

Gadermann, A. M., Guhn, M., & Zumbo, B. D. (2012). Estimating ordinal reliability for Likert-type and ordinal item response data: A conceptual, empirical, and practical guide. *Practical Assessment, Research and Evaluation, 17*(3), 1-13. https://doi.org/10.7275/n560-j767

Gómez-Penedo, J. M., Fernández-Álvarez, J., Maristany, M., & Freiberg-Hoffmann, A. (2022). Inventory of Interpersonal Problems-32 (IIP-32): Psychometric properties and normative data in a clinical sample from Argentina. *Revista Evaluar, 22*(2), 64-75. http://dx.doi.org/10.35670/1667-4545.v22.n2.38688

Goodwin, L.D., & Leech, N.L. (2006). Understanding Correlation: Factors That Affect the Size of r. *The Journal of Experimental Education, 74*(3), 249-266, https://doi.org/10.3200/JEXE.74.3.249-266

Grosse-Holtforth, M., Altenstein, D., Krieger, T., Flückiger, C., Wright, A. G., & Caspar, F. (2014). Interpersonal differentiation within depression diagnosis: Relating interpersonal subgroups to symptom load and the quality of the early therapeutic alliance. *Psychotherapy Research, 24*(4), 429-441. https://doi.org/10.1080/10503307.2013.829253

Guay, F., Morin, A. J. S., Litalien, D., Valois, P., & Vallerand, R. J. (2015). Application of Exploratory Structural Equation Modeling to Evaluate the Academic Motivation Scale. The *Journal of Experimental Education, 83*(1), 51-82. https://doi.org/10.1080/00220973.2013.876231

Horowitz, L. M. (2004). *Interpersonal Foundations of Psychopathology*. Washington, DC: American Psychological Association. https://doi.org/10.1037/10727-000

Horowitz, L. M., Alden, L. E., Wiggins, J. S., & Pincus, A.L. (2000). *IIP, Inventory of Interpersonal Problems Manual*. San Antonio: The Psychological Corporation.

Horowitz, L. M., Rosenberg, S. E., Baer, B. A., Ureño, G., & Villaseñor, V. S. (1988). Inventory of Interpersonal Problems: Psychometric properties and clinical applications. *Journal of Consulting and Clinical Psychology, 56*(6), 885-892. https://doi.org/10.1037/0022-006X.56.6.885

Inglés, C. J., Hidalgo, M. D., & Méndez, F. X. (2005). Interpersonal difficulties in adolescence: A New self-report measure. *European Journal of Psychological Assessment, 21*(1), 11–22. https://doi.org/10.1027/1015-5759.21.1.11

Jensen-Campbell, L. A., Adams, R., Perry, D. G., Workman, K. A., Furdella, J. Q., & Egan, S. (2002). Agreeableness, extraversion, and peer relations in early adolescence: Winning friends and deflecting aggression. *Journal of Research in Personality, 36*, 224–251.

John, O.P., & Srivastava, S. (1999). The Big Five trait taxonomy: History, measurement, and theoretical perspectives. In L.A. Pervin & O.P. John (Eds.), *Handbook of personality: Theory and research* (2nd Ed., pp. 102–138). New York: Guilford.

Jöreskog, K., & Sörbom, D. (1993). LISREL 8 [Computer software]. Chicago, IL: Scientific Software International, Inc. Recuperado de https://ssicentral.com

Leary, T. (1957). *Interpersonal diagnosis of personality*. New York: Ronald.

Li, C.H. (2016). The performance of ML, DWLS, and ULS estimation with robust corrections in structural equation models with ordinal variables. *Psychological Methods, 21*, 369-87. http://doi.org/10.1037/met0000093

Lubinski, D. (2000). Scientific and social significance of assessing individual differences: 'Sinking shafts at a few critical points'. *Annual Review of Psychology, 51*, 405–444.

Maristany, M. (2005). Problemas en las relaciones interpersonales y trastornos de la personalidad. Universidad de Belgrano. Recuperado de http://repositorio.ub.edu.ar

McCrae, R.R., & Costa, P.T., Jr. (1999). A five-factor theory of personality. In L.A. Pervin & O.P. John (Eds.), *Handbook of personality theory and research* (pp. 139–153). New York: Guilford.

Mindrila, D. (2010). Maximum likelihood (ML) and diagonally weighted least squares (DWLS) estimation procedures: A comparison of estimation bias with ordinal and multivariate non-normal data. *International Journal of Digital Society, 1(*1), 60-66. https://doi.org/10.20533/ijds.2040.2570.2010.0010

Mónaco, E., de la Barrera, U., & Montoya-Castilla, I. (2021). Parents and peer attachment and their relationship with emotional problems in adolescence is stress mediating? *Revista de Psicología Clínica con Niños y Adolescentes, 8*(3), 67-73. http://dx.doi.org/10.21134/rpcna.2021.08.3.8

Moreno-Peral, P., Bellón, J. A., Huibers, M. J. H., Mestre, J. M., García-López, L. J., Taubner, S., Rodríguez-Morejón, A., Bolinski, F., Sales, C. M. D., & Conejo-Cerón, S. (2020). Mediators in psychological and psychoeducational interventions for the prevention of depression and anxiety. A systematic review. *Clinical Psychology Review, 76*, 101813. https://doi.org/10.1016/j.cpr.2020.101813

Muthén, L. K., & Muthén, B. O. (2017). *Mplus: Statistical Analysis with Latent Variables: User's Guide* (Version 8). Los Angeles, CA: Authors.

Rattray, J., & Jones, M. C. (2007). Essential elements of questionnaire design and development. *Journal of Clinical Nursing, 16*, 234–243. https://doi.org/10.1111/j.1365-2702.2006.01573.x

Rial-Boubeta, A., Varela-Mallou, J., Abalo-Piñeiro, J., & Lévy-Mangin, J. P. (2006). *El análisis factorial confirmatorio*. In J. P. L. M. En y J. Varela (Eds.), Modelización con estructuras de covarianzas en ciencias sociales (pp. 119–143). Recuperado de BERITA 2024 – Berita Terupdate 2024 (netbiblo.com)

RStudio Team (2020). RStudio: Integrated Development for R. RStudio, PBC, Boston, MS. Retrieved July, 7th, 2023, from http://www.rstudio.com

Schumacker, R. E., & Lomax, R. G. (2016). *A beginner's guide to structural equation modeling* (4th Ed.). Psychology press. https://doi.org/10.4324/9781410610904

Shiner, R., & Caspi, A. (2003). Personality differences in childhood and adolescence: measurement, development, and consequences. *Journal of Children Psychology and Psychiatry, 44*(1), 2-32. https://doi.org/10.1111/1469-7610.00101

Sodano, S.M., & Tracey, T.J.G. (2006). Interpersonal Traits in Childhood: Development of the Child and Adolescent Interpersonal Survey. *Journal of Personality Assessment, 87*, 3, 317-329. https://doi.org/10.1207/s15327752jpa8703_12

Sodano, S.M. (2011). Integrating vocational interests, competencies, and interpersonal dispositions in middle school children. *Journal of Vocational Behavior, 79*(1), 110-120. https://doi.org/10.1016/j.jvb.2010.12.013

Sullivan, H. S. (1953). *The interpersonal theory of psychiatry*. New York: Norton.

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach alpha. *International Journal of Medical Education, 2*, 53-55. https://doi.org/10.5116/ijme.4dfb.8dfd

Taubner, S., Ioannou, Y., Saliba, A., Sales, C., Volkert, J., Protić, S., Adler, A., Barkauskiene, R., Conejo-Cerón, S., DiGiacomo, D., Mestre, J. M., Moreno-Peral, P., Mucha-Vieira, F., Pinheiro-Mota, C., Rangel-Santos-Henriques, M. I., Røssberg, J. I., Perdih, T. S., Schmidt, S. J., & Heinonen, E. (2023). Mediators of outcome in adolescent psychotherapy and their implications for theories and mechanisms of change: A systematic review. *European Children and Adolescent Psychiatry*. https://doi.org/10.1007/s00787-023-02186-9

Tellegen, A. (1985). Structure of mood and personality and their relevance to assessing anxiety, with an emphasis

on self-report. In A.H. Tuma & J.D. Maser (Eds.), *Anxiety and the anxiety disorders* (pp. 681–706). Hillsdale, NJ: Erlbaum.

Trógolo, M. A., Montes, S. A., & Ledesma, R. D. (2022). Adaptación argentina del Personality Inventory for DSM-Brief Form (PID-5-BF): Un análisis ESEM. *Revista Evaluar, 22*(3), 20-34. Recuperado de https://revistas.unc.edu.ar/index.php/revaluar

Wiggins, J. S. (1979). A psychological taxonomy of trait-descriptive terms: The interpersonal domain. *Journal of Personality and Social Psychology, 37*(3), 395–412. https://doi.org/10.1037/0022-3514.37.3.395

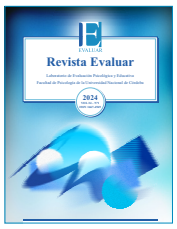
Wiggins, J. S. (1982). Circumplex models of interpersonal behavior in clinical psychology. In P. C. Kendall & J. N. Butcher (Eds.), *Handbook of research methods in clinical psychology* (pp. 183–221). New York: Wiley.

Wiggins, J. S. (1995). *IAS: Professional manual. Tampa*, FL: Psychological Assessment Resources.

Zumbo, B. D., Gadermann, A. M., & Zeisser, C. (2007). Ordinal versions of coefficients alpha and theta for Likert rating scales. *Journal of Modern Applied Statistical Methods, 6*(1), 21-29. https://doi.org/10.22237/jmasm/1177992180

2024, Vol. 24, No. 1
ISSN 1667-4545
Recuperado de https://revistas.unc.edu.ar/index.php/revaluar

Laboratorio de Evaluación Psicológica y Educativa
Facultad de Psicología - Universidad Nacional de Córdoba



Validation of the Thought-Action Fusion Scale (TAFS) for the Mexican population

Validación de la Escala Fusión Pensamiento-Acción para población Mexicana

Gabriela Sánchez-Jáuregui [1] , Julia Gallegos-Guajardo * [1] ,
Vania Luna-Huguenin [1] , Viviana Sánchez-Salinas [1] , Brian Fisak [2]

1 - School of Psychology, Universidad de Monterrey.
2 - Departament of Psychology, University of Central Florida.

Recibido: 04/08/2023 **Revisado:** 20/11/2023 **Aceptado:** 01/12/2023

Abstract

The purpose of this study was to validate the cultural adaptation of the Thought Action Fusion Scale (TAFS), which is likely the most common measure of thought action fusion, with a sample of Mexican undergraduate students. 430 undergraduate students, between 18-28 years, from a Mexican university participated in this study (M = 20.59, DE = 1.92). Based on a factor analysis, support was found for the original three-factor structure with factors measuring TAF-likelihood-for-self (TAF-LS), TAF-likelihood-for-others (TAF-LO) and TAF-Moral (TAF-M), which was consistent with previous studies conducted with non-clinical samples. Further, support was found for the criterion validity of the current version of the TAFS-M. Overall, the findings of the present study indicate that the Mexican version of TAFS exhibits good psychometric properties.

Keywords: *thought action-fusion, thought action-fusion scale, Mexican population, intrusive thoughts, obsessive-compulsive disorder*

Resumen

Fusión pensamiento-acción se refiere a la creencia de que un pensamiento puede causar una acción indeseada o tener consecuencias morales. La Escala de Fusión Pensamiento-Acción (TAFS) evalúa la disfuncionalidad de los pensamientos en que el individuo vincula sus pensamientos con las acciones. La escala cuenta con tres subescalas, TAF-probabilidad -para-uno mismo (TAF-LS), TAF-probabilidad-para-otros (TAF-LO) y la última escala siendo la de TAF-moral (TAF-M). El objetivo de este estudio fue validar la adaptación del TAFS con una muestra de estudiantes universitarios mexicanos. 430 estudiantes universitarios entre 18 a 24 años de edad, de una universidad de México participaron en el estudio (M = 20.59, DE = 1.92). Los resultados mostraron que la estructura original de tres factores tiene un buen ajuste con los datos, lo cual fue consistente con estudios previos realizados con muestra no-clínica. Estos resultados indican una buena consistencia interna y una adaptación adecuada de la prueba original para ser utilizada con población universitaria mexicana.

Palabras clave: *pensamiento fusión-acción, scala de pensamiento fusión-acción, población mexicana, pensamientos intrusivos, trastorno obsesivo-compulsivo*

***Correspondence to:** Julia Gallegos-Guajardo, School of Psychology, University of Monterrey, Av. Ignacio Morones Prieto 4500, Zona Valle Poniente, 66238 San Pedro Garza García, N. L. Phone number: 52-1-8115316352. E-mail address: julia.gallegos@udem.edu

How to cite: Sánchez-Jáureguil, G., Gallegos-Guajardo, J., Luna-Hugenin, V., Sánchez-Salinas, V., & Fisak, B. (2024). Validation of the Thought-Action Scale (TAFS) for the Mexican population. *Revista Evaluar, 24*(1), 76-87. Retrieved from https://revistas.unc.edu.ar/index.php/revaluar

Participaron en la edición de este artículo: Abigail Pérez, Florencia Ruiz, Jorge Bruera, Eugenia Barrionuevo.

Introduction

Thought-action fusion (TAF) refers to the belief that thoughts and actions are linked, particularly that some thoughts, usually unwanted intrusive thoughts, can increase the likelihood of catastrophic events happening or that they may imply a certain immorality of the individual's character (Berle & Starcevic, 2005; Shafran et al., 1996). Having intrusive thoughts is common among the non-clinical population, with most people being able to dismiss them without having distressing emotions about them (Inozu et al., 2014). According to the cognitive model, it is the belief about the meaning of the thought that can change a benign intrusive thought into an obsession (Bailey et al., 2014; Barrera & Norton, 2011), which is consistent with what the thought-action fusion construct implies. This type of cognitive bias has been commonly associated with obsessive compulsive (OC) symptoms and has been observed in patients with anxiety disorders and obsessive-compulsive disorder (OCD). Furthermore, thought-action fusion may exacerbate and/or maintain OCD symptoms by providing unhealthy reassurance and/or by temporarily decreasing their anxiety symptoms. Other aspects involved in OCD have been found to be related to thought-action fusion, amongst these are thought suppression, neutralizing behaviors and inflated responsibility (Bailey et al., 2014; Berle & Starcevic, 2005).

The concept of thought-action fusion was originally conceptualized through the clinical work and analysis of patients with OCD. The relation between TAF and OCD symptoms has been analyzed in several studies with clinical and non-clinical populations. The following studies describe how these two variables are related and may be present in different groups of people. Aydin et al. (2012) conducted a study with a sample of 263 Turkish undergraduate students from ages 17 to 40. The authors hypothesized that thought-action fusion would be a detrimental factor in the worsening and severity of OCD symptoms. However, they found that this factor by itself was not a detrimental aspect, but rather it is its combination with other factors that was detrimental. Specifically, thought-action when it is accompanied by maladaptive interpersonal schemas and perfectionist attitudes showed to be a vulnerability agent in the development and severity of OCD symptoms.

Furthermore, in a subsequent study conducted by Amir (2017) with adolescents from Iran with an average age of 12.8, the positive and statistically significant association between thought-action fusion and OCD symptoms was also confirmed, particularly with the symptomatology related to being obsessive. Similarly, another study reported that intrusive and unwanted thoughts about oneself were classified as negative only in people who have reported a high thought-action fusion tendency and OCD symptoms (Hezel et al., 2017).

Nevertheless, there are studies that indicate that thought-action fusion is not exclusive to OCD and that it can be observed in other disorders, such as anxiety and depression (Abramowitz et al., 2003; O'Leary et al., 2009). In a study consisting of 37 patients with diverse anxiety disorders (excluding those with OCD), a positive and statistically significant correlation was found between Generalized Anxiety Disorder and thought action fusion within the subcategory of severity of the symptoms (Thompson-Hollands et al., 2013). A similar result was found in another study, where the authors have explored in more detail the nature of this relationship with OCD patients, notably investigating the link between the intensity of the reaction to the probability of a particular outcome within any situation (Odriozola-González et al., 2016).

Thought Action Fusion Scale

Thought-action fusion can be evaluated through the use of the *Thought-Action Fusion Scale* (TAFS) (Shafran et al., 1996), a self-report measure consisting of 19 items. Two forms of TAF are often described, moral TAF and likelihood TAF (Berman et al., 2011; Rassin et al., 2001; Shafran et al., 1996). Likelihood TAF is the belief that having a thought about a disturbing or unwanted event will increase the probability of the occurrence of that event (Shafran et al., 1996). This type of belief can be about oneself (e.g., "*If I think about being in a car accident, it makes it more likely that I could be in one"),* which is referred to as "Likelihood-self"; or it can be about an event that involves someone else (e.g., "*If I think about my sister being in a car accident, it will be more likely that she could be in one*"), which relates to as "likelihood-other" (Rachman & Shafran, 2004). Nevertheless, moral TAF, refers to the tendency to believe that having an unacceptable or unwanted thought is morally equivalent to the real action (e.g., believing that thinking about hitting someone is in itself as bad as the action of actually hitting them) (Cougle et al., 2013). It is possible that the misinterpretation in this type of belief revolves around the idea that this thought reveals the "true" nature of the person, being that they may be "wicked" or "bad" (Shafran et al., 1996).

Originally, the TAFS consisted of 34 items and was validated with a clinical and a non-clinical sample of undergraduate students, both from Canada. The internal consistency was found to be high for both samples (α = .95, .96, respectively). A two factor solution, moral TAF and likelihood TAF, was supported for the clinical sample and a three-factor solution for the student sample: moral TAF, likelihood-self TAF and likelihood-other TAF. The measure was later revised by the same authors and reduced to 19 items. It was validated again with a clinical and a non-clinical sample of undergraduate students from Canada. Just as with the first version, a two-factor solution was found for the clinical sample and a three-factor solution for the student sample and the internal consistency was found to be also high for both samples (α = .85 to .96) (Shafran et al., 1996).

The TAFS has also been revised and standardized for Spanish-speaking population from Spain (Jáuregui-Lobera et al., 2013) with a clinical sample of patients with eating disorders and a non-clinical sample of university students. In this study, the TAFS and its subscales showed a good internal consistency of α = .88 for patients and α = .90 for students. Results in this study demonstrated that a three-factor solution was best for both clinical and non-clinical samples, as opposed to earlier studies that found a two-factor structure for the clinical samples.

There was also an evaluation of an extended version of the TAFS conducted by Amir et al. (2001) in order to include a subscale that assessed the likelihood of events happening to others, and also subscales to rate the responsibility and cost for having these thoughts. In order to do this, they added eight subscales to the original version. The *TAFS-R* comprises 11 subscales with 30 items in total, and was validated with a sample of 424 undergraduate students from the United States. Results of this study showed that all subscales had adequate coefficient alphas and that the *TAFS-R* was well adapted to assess the role that thought action fusion has in OCD symptoms. The *TAFS-R* was later revised and standardized for the use of it with an Iranian sample of students (Pourfaraj et al., 2008). The reliability coefficients of the total scale were calculated by two methods: internal consistency and test-retest, which were .81 and .61, respectively. This revised version was later tested with an American clinical sample of adults with OCD symptoms receiving treatment at an outpa-

tient clinic. The reliability of this version was tested through a confirmatory factor analysis and was found to have high scale reliability (p = .97) (Myer & Brown, 2012).

Although previous studies about TAFS have been conducted with samples of undergraduate students, there is still a paucity of research focused on the evaluation of the psychometric properties TAFS in Spanish-speaking non-clinical populations. It is particularly noteworthy that the psychometric properties of the TAFS have not been examined in Latin-Speaking countries. Consequently, the purpose of the current study was to examine the psychometric properties of the first known Spanish version of the TAFS for the Mexican population (TAFS-MV).

The purpose of this study was to culturally adapt the TAFS through a back-to-back translation process and the selection of wording and phrases that could capture the cultural and linguistic nuances of the Spanish language spoken in Mexico. The psychometric properties of the TAFS-MV were evaluated in a sample of Mexican undergraduate students attending a university in the northern part of Mexico, using the adapted Spanish measures of the Yale-Brown Obsessive-Compulsive Scale – Self-Report Version (YBOCS-SR) and the White Bear Suppression Inventory (WBSI-MV) for Mexican population to examine convergent validity. It was hypothesized that the original three factor structure that has been found in non-clinical samples would be maintained in the TAFS-MV, along with a good reliability and convergent validity.

Method

Participants

430 students from a university in northern Mexico volunteered to participate in this online survey-based study in exchange for extra credit. Inclusion criteria included that the participants must be 18 years or older and had to be enrolled in an undergraduate program in order to participate in this study. As the data was collected in a community sample, data collection did not include gathering information about an existing psychiatric diagnosis, use of psychiatric medication or psychiatry history. The sample was 79.4% females, most of them Hispanic (91.2%), with an average participant age of 20.59 (SD = 1.92).

Measures

Thought-Action Fusion Scale – Mexican Version. The TAFS is a 19-item self-report measure designed to evaluate the construct of thought-action fusion (Shafran et al., 1996). This measure has been reported in its original version as an internal consistency of α = .93, with a Likert scale from 1 (*Strongly Disagree*) to 5 (*Strongly Agree*). The measure has three subscales, one of which is the Moral-TAF that assesses the TAF relation to morality: *"Having violent thoughts is almost as unacceptable to me as a violent act"*. Next, is the TAF-Likelihood-others which assesses negative outcomes for others: *"If I think of a relative/friend falling ill this increases the risk that he/she will fall ill"*. Lastly, the TAF-Likelihood-Self, which assesses the concerns of negative outcomes for themselves: *"If I think of myself falling ill, this increases the risk I will fall ill"*. The measure has shown good psychometric properties, including good internal consistency and good criterion validity, as the TAFS has been found to be associated with symptoms of OCD (Cougle et al., 2013; Rachman & Shafran, 2004).

This measure was back-to-back translated from the original English into Spanish. Within the translation process, language was adapted and cultural aspects were revised by two professors

from the School of Psychology at a Mexican university, in order to make the items more accurate and representative to the Mexican population. The back-translated version of the measure was evaluated by native English speakers (See Appendix B). Cronbach's alphas with the current sample are provided in the results section.

White Bear Suppression Inventory – Mexican Version (WBSI-MV). The WBSI is among the most commonly utilized measures to assess thought suppression (Wegner & Zanakos, 1994). This self-report measure consists of 15 items with response options on a 5-point Likert scale ranging from 1 (*Strongly disagree*) to 5 (*Strongly agree*). The WBSI exhibits good psychometric properties, including excellent internal consistency (α = 0.93), and validity, as the WBSI has been found to be positively and significantly associated with measures of anxiety and depression (Wegner & Zanakos, 1994). For the purpose of this study, this measure was also translated from English into Spanish and back-translated. Within the translation process, language was adapted and cultural aspects were revised by two professors from the School of Psychology at a Mexican university, in order to make the items more accurate and representative. The back-translated version of the measure was evaluated by native English speakers. The internal consistency of the WBSI-MV for the current sample was excellent, α = .90 (Gallegos-Guajardo et al., 2020). In the current study, this measure was used to assess concurrent validity, as thought-action fusion has been constantly associated with thought suppression.

Yale-Brown Obsessive Compulsive Scale Self Report Version (Y-BOCS-SR). The Symptom Severity Scale of the Y-BOCS-SR was administered in this study (Baer, 1991; Ólafsson et al., 2010). The Y-BOCS-SR assesses components of symptom severity together with the amount of distress, interference, time spent on obsessions or compulsions, and perceived control of obsessions and compulsions. The Y-BOCS-SR severity scale includes a total of seven items to measure obsessions and seven items to measure compulsions; each question has a Likert scale ranging from 0 to 4. For the purpose of this study, this measure was also translated from English into Spanish and back-translated. Within the translation process, language was adapted and cultural aspects were revised by two professors from the School of Psychology at a Mexican university, in order to make the items more accurate and representative. The back-translated version of the measure was evaluated by native English speakers. The internal consistency of the Y-BOCS-SR in Spanish for the current sample was excellent, α = .93. In the current study, this measure was used to assess concurrent validity, as thought-action fusion has consistently found to be associated with OCD symptoms.

Design and Procedures

The study was internet-based, in which participants completed a number of self-report measures. Before participating in the study, potential participants were required to complete and sign an informed consent form. As part of the informed consent, potential participants were reminded that their involvement was voluntary. Those who agreed were then asked to complete a demographic questionnaire and Spanish versions of the measures described below. The study was approved by the appropriate university Institutional Review Board.

Data Analysis Plan

Confirmatory factor analyses (CFAs) were planned by using Mplus version 8.1 to determine

the degree to which data obtained in the current sample fits with the previously established models. In particular, the purpose of the first CFA was to examine the degree to which the three-factor model obtained in the original study with non-clinical population (Shafran et al., 1996) fit with the current data. The cut-off to determine a good adjustment was .95 or greater. Items with loadings of .30 or greater on a single factor were to be retained.

Following establishment of the factor structure, an assessment of reliability was planned by examining Cronbach's alphas: .7 and above was considered acceptable, .8 and above good and .9

Appendix A

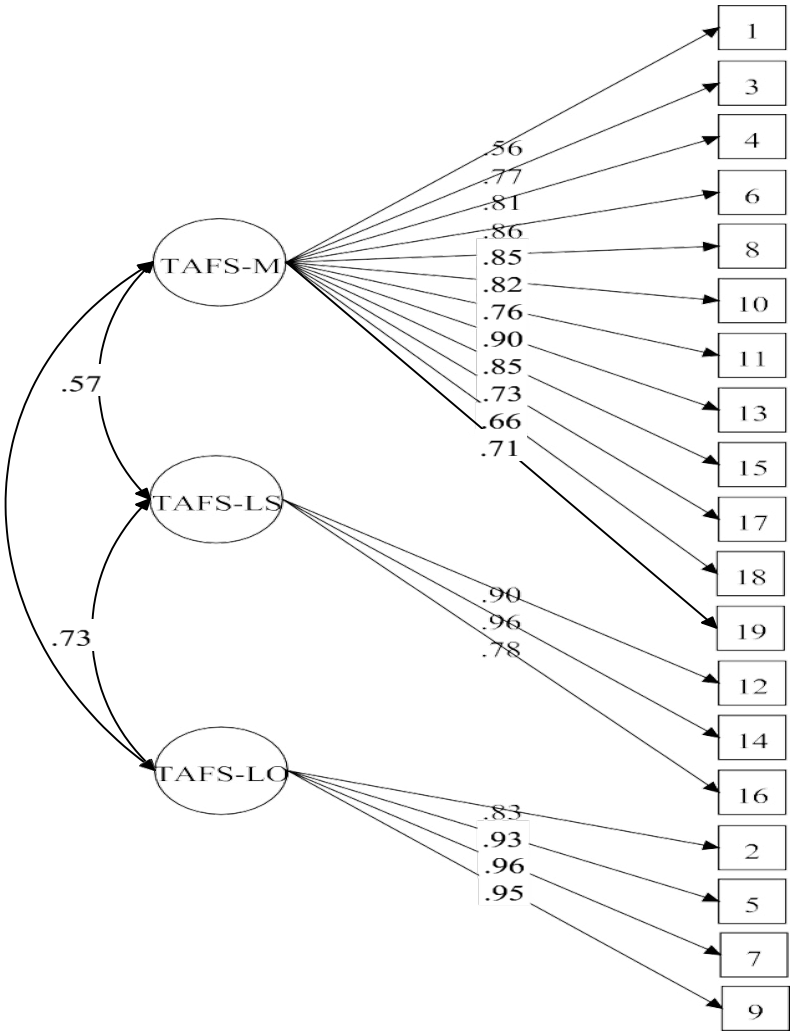


Figure 1.
Standardized Factor Loadings for the Thought Action Fusion Scale based on Confirmatory Factor Analysis.

and above excellent. An examination of the validity of the TAFS-MV was planned by assessing the magnitude of the association between the TAFS-MV and the designated validation measures (i.e., Y-BOCS-SR and WBSI-MV). Correlations were expected to be at least moderate in magnitude, .3 or greater. The SPSS version 26.0 was to be utilized for the bivariate correlations. Finally, qualitative analysis was conducted to determine if the translation process produced substantive differences and improved utility, for use in a Mexican sample, relative to the previously translated Spanish version of the TAFS with the Spanish sample.

Results

Results showed that the original three-factor structure was a good fit with the data, $\chi^2/df = 3.10$, CFI = .98, TLI = .98, RMSEA = .070 (see Figure 1 in Appendixes). Also, reliability analyses were performed to explore the internal consistency of the Spanish version of the TAFS-MV, reporting good internal consistency for the three subscales: TAFS- MV- Moral subscale (TAFS-MV-M), TAFS- MV- Likelihood- Other subscale (TAFS-MV-LO), and TAFS- MV-Likelihood-Self.

In order to assess the convergent validity, Pearson correlations were conducted with the TAFS-MV and Y-BOCS-SR and WBSI-MV. Results showed that all three TAFS-MV subscales were found to be significantly associated with the WBSI-MV and with the Spanish version of the Y-BOCS-SR (see Table 1).

Discussion

The purpose of the current study was to examine the psychometric properties of the Thought Action Fusion Scale for the Mexican population (TAFS-MV). Consistent with previous studies conducted with non-clinical samples, results from our current study supported a three-factor structure, being these factors: TAFS-MV Moral, TAFS-MV Likelihood to others and TAFS-MV- Likelihood to self (Jáuregui-Lobera et al., 2013; Pourfaraj et al., 2008; Shafran et al., 1996). As well, consistent with previous research, the TAFS-MV showed good internal consistency and support was found for the validity of the TAFS-MV as the three factors were found to be positively and significantly associated with thought suppression and OCD symptoms.

The present data indicates a good internal consistency for the use of TAFS-MV and a good adaptation of the original measure to be used in the Mexican population. This indicates a good fit with the original sample and that TAFS-MV is an effective tool to assess thought action fusion tendency in a non-clinical population. Particularly in Mexico, the importance of evaluating the presence of risk factors, such as thought action fusion, relays in the fact that the rates of anxiety disorders and

Table 1
Correlations between TAFS subscales and the WBSI-M and Y-BOCS-SR.

	TAFS-LS	TAFS-LO	TAFS-M
WBSI	.42**	.27**	.29**
Y-BOCS-SR	.21**	.25**	.16**

Note. ** = $p < .001$. TAFS-LS = Thought Action Fusion Scale- Likelihood-Self subscale; TAFS-LO = TAFS- Likelihood-Other subscale; TAFS-M = TAFS- Moral subscale. WBSI = White Bear Suppression Inventory and Y-BOCS-SR= Yale Brown Obsessive Compulsive Inventory.

OCD have been increasing very fast in the young and adult Mexican populations (Caraveo-Anduaga & Colmenares, 2004; Lozano-Vargas, 2017; Ulloa et al., 2011). By assessing this particular risk factor, mental health professionals could be better prepared to conduct an accurate diagnosis for the individual and provide a more effective treatment, as well as to track the patient's progress on risk factors that may be maintaining the disorder (Ulloa et al., 2011). Henceforth, it is encouraged that future research, in addition to measuring outcome variables, such as anxiety or OCD symptoms, also focuses on gathering valuable information of risk and protective factors for these disorders. For instance, factors such as emotional regulation (De la Rosa-Gómez et al., 2021) and mindfulness (Gustin-García & Alegre-Bravo, 2021), will be of great value to assess, as well as flourishing and wellbeing (González-Rivera, 2018).

In relation to the specific differences found between the Spanish European (Jáuregui-Lobera et al., 2013) and Mexican Spanish version of the TAFS, a detailed qualitative analysis that compares the items of both versions highlighted several differences. It was found that specific items meant different things in each version or used different words, one example can be observed in items 5 and 14 with the term for car *"carro"* which applies more to the Latin-American population as opposed to *"coche"* which is used specifically in the Spanish population. Moreover, in item 6 *"Tener un pensamiento desagradable sobre alguien, es casi igual de malo que llevar a cabo una acción desagradable"* the word *"desagradable"* is an accurate representation of an unpleasant thought to have about someone, which is more commonly used in Mexican population instead of the word *"repugnante"* which is used in the Spanish version and means abhorrence and complete hatred thought or action. Therefore, it was important to culturally adapt the measure in order to be more accurate and relevant for the Mexican population. Moreover, having the psychometric properties explored in a Mexican sample can ensure precise results in further investigations for this specific population, as opposed to using adaptations or norms from a different country.

In general, having the TAFS-MV is useful for researchers and clinicians, allowing them to study thought action fusion in both non-clinical and clinical samples in Mexico, by having a validated adaptation with accurate translation that applies specifically to the Mexican population. In addition, having this scale validated in Mexico serves as an opportunity to conduct more research regarding anxiety and OCD symptoms in clinical and non-clinical samples. For instance, research in Mexico about OCD is very scarce (i.e., Caraveo-Anduaga & Colmenares, 2004; Lozano-Vargas, 2017), and it has been shown that the symptoms reported by the clinical sample do not always match with those symptoms stated by the general population. Having a validated measure with the Mexican population may facilitate the study of thought action fusion and the specific patterns in which it predicts anxiety and/or OCD symptoms. It would also provide a more comprehensive view of how these symptoms manifest in this particular population.

Some limitations regarding this study need to be addressed further on and are related to the sample of this study. The sample is limited to a specific age range and by only including undergraduate students, therefore, narrowing down the generalizability of these results to the general population in Mexico. Another limitation is the demographics of the sample, as it is only composed of undergraduate students from northern Mexico, which is expected to display cultural differences to the southern and central regions of the country, again limiting its generalizability. Also, the current study did not include a clinical

sample, as it focuses only on the general population, some of them experiencing a wide range of OCD symptoms. Further research should evaluate this measure by including a clinical sample from Mexico, as previous studies have done this in different countries (Jáuregui-Lobera et al., 2013; Pourfaraj et al., 2008; Shafran et al., 1996). It is also of importance to continue exploring the TAFS-MV with samples of different age ranges, as well as different educational levels, SES levels and living in both, rural and urban areas from across the country.

In summary, current findings indicate that the TAFS-MV is a useful and valid tool to assess thought action fusion in the Mexican population as the measure showed good internal consistency and a three-factor structure supporting the factor structure reported in the original measure with non-clinical samples.

Appendix B

Escala de Fusión Pensamiento-Acción

Instrucciones: A continuación, se enumeran 24 afirmaciones. Por favor utiliza la siguiente escala para indicar si estás de acuerdo o en desacuerdo con cada una de ellas.

En completo desacuerdo 1	En desacuerdo 2	Neutral 3	De acuerdo 4	De acuerdo completamente 5

1. Pensar en hacer una observación extremadamente crítica a un amigo es casi igual de inaceptable como decirla.

2. Pensar que un pariente/amigo perderá su trabajo aumenta el riesgo de que suceda.

3. Para mí, tener un pensamiento blasfemo es casi igual de pecaminoso como la acción de la blasfemia en sí.

4. Para mí, pensar en maldecir a alguien es casi igual de inaceptable como en verdad maldecirlo.

5. Pensar que un pariente/amigo tendrá un accidente de auto aumenta el riesgo de que suceda.

6. Tener un pensamiento desagradable sobre alguien, es casi igual de malo que llevar a cabo una acción desagradable.

7. Pensar que un pariente/amigo se lastimará en una caída aumenta el riesgo de que la caída suceda y se lastime.

8. Para mí, tener pensamientos violentos es casi igual de inaceptable que llevar a cabo acciones violentas.

9. Pensar que un pariente/amigo se enfermará aumenta el riesgo de que la enfermedad suceda.

10. Cuando pienso en hacer una observación o gesto obsceno en un lugar de oración, es casi igual de pecaminoso que realmente hacerlo.

11. Desear que alguien se haga daño es casi igual de malo que causar el daño en sí.

12. Si yo pienso en que me haré daño en una caída, esto aumenta el riesgo de que me caiga y me lastime.

13. Pensar cruelmente sobre un amigo es casi igual de desleal que llevar a cabo un acto cruel.

14. Si yo pienso que tendré un accidente automovilístico, va a aumentar el riesgo de que tenga un accidente automovilístico.

15. Si yo pienso en hacerle un gesto obsceno a otra persona, es casi igual de malo como hacerlo.

16. Si yo pienso que me voy a enfermar, aumenta la posibilidad de que me vaya a enfermar.

17. Si tengo un pensamiento de celos, es casi igual de malo que hacer un comentario de celos.

18. Para mí, pensar en engañar a una persona con la que se tiene una relación amorosa, es casi igual de inmoral que en verdad engañarla.

19. Para mí, es inaceptable tener pensamientos blasfemos en un lugar de oración.

References

Abramowitz, J. S., Whiteside, S., Lynam, D., & Kalsy, S. (2003). Is thought-action fusion specific to obsessive-compulsive disorder? A mediating role of negative effect. *Behaviour Research and Therapy, 41*(9), 1069-1079. https://doi.org/10.1016/S0005-7967(02)00243-7

Altın, M., & Gençöz, T. (2010). How does thought-action fusion relate to responsibility attitudes and thought suppression to aggravate the obsessive-compulsive symptoms? *Behavioural and Cognitive Psychotherapy, 39*(1), 99-114. https://doi.org/10.1017/S1352465810000524

Amir, N., Freshman, M., Ramsey, B., Neary, E., & Bartholomew, B. (2001). Thought-action fusion in individuals with OCD symptoms. *Behavioural Research and Therapy, 39*(7), 765-776. https://doi.org/10.1016/S0005-7967(00)00056-5

Amir, F. (2017). The relationship between thought action fusion with obsessive compulsive dimensions and severity in adolescents. *Indian Journal of Positive Psychology*, *8*(4), 531-534.

Aydın, A., Boysan, M., Tutarel-Kışlak, S., Kalafat, T., Selvi, Y., & Beşiroğlu, L. (2012). The predictive value of interpersonal schemas, perfectionism and thought action fusion in obsessive compulsive disorder. *The Journal of Psychiatry and Neurological Sciences*, *25*, 108-118. https://doi.org/10.5350/dajpn2012250202

Baer, L. (1991). *Getting control: Overcoming obsessions and compulsions*. Little Brown.

Bailey, B. E., Wu, K. D., Valentiner, D. P., & McGrath, P. B. (2014). Thought–action fusion: Structure and specificity to OCD. *Journal of Obsessive-Compulsive and Related Disorders, 3*(1), 39-45. https://doi.org/10.1016/j.jocrd.2013.12.003

Barrera, T. L., & Norton, P. J. (2011). The appraisal of intrusive thoughts in relation to Obsessional–Compulsive symptoms. *Cognitive Behaviour Therapy, 40*(2), 98-110. https://doi.org/10.1080/16506073.2010.545072

Berman, N. C., Abramowitz, J. S., Wheaton, M. G., Pardue,

C., & Fabricant, L. (2011). Evaluation of an in-vivo measure of thought–action fusion. *Journal of Cognitive Psychotherapy, 25*(2), 155-164. https://doi.org/10.1891/0889-8391.25.2.155

Berle, D., & Starcevic, V. (2005). Thought–action fusion: Review of the literature and future directions. *Clinical Psychology Review, 25*(3), 263-284. https://doi.org/10.1016/j.cpr.2004.12.001

Caraveo-Anduaga, J., & Colmenares, E. (2004). The epidemiology of obsessive-compulsive disorder in Mexico City. *Salud Mental, 27*(2), 1-6.

Cougle, J. R., Purdon, C., Fitch, K. E., & Hawkins, K. A. (2013). Clarifying relations between thought-action fusion, religiosity, and obsessive-compulsive symptoms through consideration of intent. *Cognitive Therapy and Research*, *13*(2), 221-231. https://doi.org/10.1007/s10608-012-9461-8

De la Rosa-Gómez, A., Hernández- Posadas, A., Valencia, P. D., & Guajardo-Garcini, D. A. (2021). Análisis dimensional de la Escala de Dificultades en la Regulación Emocional (DERS-15) en universitarios mexicanos. *Revista Evaluar, 21*(2), 80-97. https://doi.org/10.35670/1667-4545.v21.n2.34401

Gallegos-Guajardo, J., Sanchez-Jauregui, G., Hidalgo, J., Alicea-Lozada, L., Luna-Huguenin, V., & Fisak, B. (2020). *Validation of the White Bear Suppression Inventory for the Mexican population*. Manuscript submitted for publication. https://doi.org/10.35670/1667-4545.v23.n2.42075

Ghamari-Kivi, H., Mohammadipour-Rik, N., & Sadeghi-Movahhed, F. (2012). Explanation of obsessive-compulsive disorder and major depressive disorder on the basis of thought-action fusion. *Iranian Journal of Psychiatry and Behavioral Sciences*, *7*(2), 44-50.

González-Rivera, J. A. (2018). Propiedades psicométricas de la Escala de Florecimiento en Puerto Rico. *Revista Evaluar, 18*(3), 30-43.

Gustin-García, M. G., & Alegre-Bravo, A. A. (2021). Validación de Child and Adolescent Mindfulness Measure en escolares de Lima, Perú. *Revista Evaluar, 21*(2), 63-79. https://doi.org/10.35670/1667-4545.v21.n2.34398

Hezel, D. M., Stewart, S. E., Riemann, B. C., & McNally, R. J. (2017). Clarifying the thought-action fusion bias in obsessive-compulsive disorder. *Journal of Obsessive-Compulsive and Related Disorders, 20*, 75-84. https://doi.org/10.1016/j.jocrd.2017.10.004

Inozu, M., Ulukut, F. O., Ergun, G., & Alcolado, G. M. (2014). The mediating role of disgust sensitivity and thought-action fusion between religiosity and obsessive compulsive symptoms. *International Journal of Psychology*, *49*(5), 334-341. https://doi.org/10.1002/ijop.12041

Jáuregui-Lobera, I., Santed-Germán, M., Bolaños-Ríos, P., & Garrido-Casals, O. (2013). Spanish version of the thought-action fusion questionnaire and its application in eating disorders. *Psychology Research and Behavior Management, 6*, 75-86. https://doi.org/10.2147/prbm.s51183

Lozano-Vargas, A. (2017). Aspectos clínicos del trastorno obsesivo compulsivo y trastornos relacionados. *Revista de Neuro-Psiquiátrica, 80(*1), 35. https://doi.org/10.20453/rnp.v80i1.3058

Marino-Carper, T., Negy, C., Burns, G., & Lunt, R. A. (2010). The effects of psychoeducation on thought-action fusion, thought suppression, and responsibility. *Journal of Behavior Therapy and Experimental Psychiatry, 41*(3), 289-296. https://doi.org/10.1016/j.jbtep.2010.02.007

Myer, J. P., & Brown, T. A. (2012). Psychometric evaluation of the Thought-action Fusion Scale in a large clinical sample. *Assessment, 20*(6), 764-775. https://doi.org/10.1177/1073191112436670

O'Leary, E. M., Rucklidge, J. J., & Blampied, N. (2009). Thought–action fusion and inflated responsibility beliefs in obsessive–compulsive disorder. *Clinical Psychologist, 13*(3), 94-101. https://doi.org/10.1080/13284200903369557

Ólafsson, R. P., Snorrason, I., & Smári, J. (2010). Yale-Brown Obsessive Compulsive Scale: Psychometric properties of the self-report version in a student sam-

ple. *Journal of Psychopathology and Behavioral Assessment, 32*, 226-235.

Odriozola-González, P., Pérez-Alvarez, M., García-Montes, J. M., Perona-Garcelán, S. Y., & Vallina-Fernández, O. (2016). El papel mediador de las variables metacognitivas en la relación entre la fusión pensamiento-acción y la sintomatología obsesivo-compulsiva. *Suma Psicológica, 23*(2), 80-89. https://doi.org/10.1016/j.sumpsi.2016.08.001

Pauls, D., Abramovitch, A., Rauch, S., & Geller, D. (2014). Obsessive-compulsive disorder: An integrative genetic and neurobiological perspective. *Natural Reviews Neuroscience, 15*, 410-424. https://doi.org/10.1038/nrn3746

Pourfaraj, M., Mohammadi, N., & Taghavi, M. (2008). Psychometric properties of revised Thought-Action Fusion Questionnaire (TAF-R) in an Iranian population. *Journal of Behavior Therapy and Experimental Psychiatry, 39*(4), 600-609. https://doi.org/10.1016/j.jbtep.2008.02.001

Rachman, S., & Shafran, R. (2004). Thought-action fusion: A Review. *Journal of Behavior Therapy and Experimental Psychiatry, 35*(2), 87-107. https://doi.org/10.1016/j.jbtep.2004.04.002

Rassin, E., Muris, P., Schmidt, H., & Merckelbach, H. (2000). Relationships between thought–action fusion, thought suppression and obsessive–compulsive symptoms: A structural equation modeling approach. *Behavior Research and Therapy, 38*(9), 889-897. https://doi.org/10.1016/s0005-7967(99)00104-7

Rassin, E., Merckelbach, H., Muris, P., & Schmidt, H. (2001). The Thought-action Fusion Scale: Further evidence for its reliability and validity. *Behaviour Research and Therapy, 39*(5), 537-544. https://doi.org/10.1016/s0005-7967(00)00031-0

Rintala, H., Chudal, R., Leppämäki, S., Leivonen, S., Hinkka-Yli, S., & Sourander, A. (2017). Register-based study of the incidence, comorbidities and demographics of obsessive-compulsive disorder in specialist healthcare. *BMC Psychiatry, 17*(64), 2-8. https://doi.org/10.1186/s12888-017-1224-3

Shafran, R., Thordarson, D. S., & Rachman, S. (1996). Thought-action fusion in obsessive-compulsive disorder. *Journal of Anxiety Disorders, 10*(5), 379-391. https://doi.org/10.1016/0887-6185(96)00018-7

Smári, J., & Hólmsteinsson, H. E. (2001). Intrusive thoughts, responsibility attitudes, thought-action fusion, and chronic thought suppression in relation to obsessive-compulsive symptoms. *Behavioural and Cognitive Psychotherapy, 29*(1), 13-20. https://doi.org/10.1017/s1352465801001035

Thompson-Hollands, J., Farchione, T. J., & Barlow, D. H. (2013). Thought-action fusion across anxiety disorder diagnoses. *The Journal of Nervous and Mental Disease*, *201*(5), 407-413. https://doi.org/10.1097/nmd.0b013e31828e102c

Ulloa, R., Palacios, L., & Sauer, T. (2011). Trastorno obsesivo-compulsivo en niños y adolescentes: Una revisión del tratamiento. *Salud Mental, 34*(5), 415-420.

Wegner, D. M., & Zanakos, S. (1994). Chronic thought suppression. *Journal of Personality, 62*(4), 615-640.

Williams, A. D., Lau, G., & Grisham, J. R. (2013). Thought-action fusion as a mediator of religiosity and obsessive-compulsive symptoms. *Journal of Behavior Therapy and Experimental Psychiatry, 44*(2), 207-212. https://doi.org/10.1016/j.jbtep.2012.09.004

EVALUAR

2024, Vol. 24, No. 1
ISSN 1667-4545
Recuperado de https://revistas.unc.edu.ar/index.php/revaluar

Laboratorio de Evaluación Psicológica y Educativa
Facultad de Psicología - Universidad Nacional de Córdoba

Initial assessment of attitudes towards reporting cyber-harassment in children and adolescents

Evaluación inicial de las actitudes hacia el reporte de casos de ciberacoso en niños/niñas y adolescentes

Nicolás Martínez-Ramos * [1], Camila, Arbeláez-Gómez [1], William Martínez [2], Consuelo Ramos [2], Judy Benavidez [3]

1 - Pontificia Universdad Javeriana, Bogotá.
2 - Universidad La Gran Colombia, Bogotá.
3 - Universidad Nacional de Colombia, Bogotá.

Recibido: 12/08/2023 **Revisado:** 10/10/2023 **Aceptado:** 12/10/2023

Abstract

This study aimed to validate the initial scale of attitudes toward reporting cyber-harassment in children-adolescents among people from the Colombian general community. An instrumental study was developed, which was reached in two non-probabilistic samples (n^1 = 220, n^2 = 204). With the first one, a Robust Exploratory Factor Analysis (EFA) was performed, finding a unidimensional 7-item solution (RMSEA = .07, SRMR = .05, CFI = .99, and TLI = .99). A Confirmatory Factor Analysis was conducted in the second sample, finding an optimal fit (RMSEA = .03, SRMR = .04, CFI = .96, TLI = .95). In both models, optimal reliability indices were obtained. The total score was crossed with socio-demographics, finding some real significant difference. Software Factor and *R* with different packages were used.

Keywords: *Colombia, cyber-harassment, psychometrics, attitudes, factor-analysis*

Resumen

Este estudio tuvo como objetivo validar la escala inicial de actitudes hacia la denuncia del ciberacoso en niños-adolescentes entre personas de la comunidad general colombiana. Se desarrolló un estudio instrumental, el cual se alcanzó en dos muestras no probabilísticas (n^1 = 220, n^2 = 204). Con la primera se realizó un Análisis Factorial Exploratorio Robusto (AFE), encontrando una solución unidimensional de 7 ítems (RMSEA = .07, SRMR = .05, CFI = .99 y TLI = .99). En la segunda muestra, se realizó un Análisis Factorial Confirmatorio, encontrando un ajuste óptimo (RMSEA = .03, SRMR = .04, CFI = .96, TLI = .95). En ambos modelos se obtuvieron índices de confiabilidad óptimos. La puntuación total se cruzó con la sociodemográfica, encontrando alguna diferencia significativa real. Se utilizó el software Factor y *R* con diferentes paquetes.

Palabras clave: *Colombia, ciber-acoso, psicometría, actitudes, análisis-factorial*

***Correspondence to:** Nicolás Martínez-Ramos, ncolasmramos@gmail.com or martineznicolas@javeriana.edu.co

How to cite: Martinez-Ramos, N., Arbeláez-Gómez, C., Martínez, W., Ramos, C., & Benavidez, J., (2024): Initial assessment of attitudes towards reporting cyber-harassment in children and adolescents. *Revista Evaluar, 24*(1), 88-102. Retrieved from https://revistas.unc.edu.ar/index.php/revaluar

Participaron en la edición de este artículo: Gloria Nieve, Eugenia Barrionuevo, Pablo Carpintero, Florencia Ruiz, Jorge Bruera.

Introduction

Digital development has allowed the advancement of different Information and Communication Technologies (ICT). However, there is a great concern about the new risks derived from their use, such as cyber-attacks, in which children and adolescents are more likely to be victims (Ayub & Malik, 2020; Beran et al., 2012; Cuenca-Pinqueras et al., 2020; Foley et al., 2015; Hassan et al., 2020; Jagayat & Choma, 2021; Tokunaga, 2010; van Laer, 2014). Cyber-harassment has been defined as an intentional and repetitive aggression in digital environments, characterized by the absence of geographical and temporal limits, where the victim finds it difficult to give a direct response to the aggressor (Ayub & Malik, 2020; Bégin, 2018; Foley et al., 2015; Neelam, 2017; Redondo-Pacheco et al. 2018; Zweig et al., 2013). The primary expressions of this phenomenon are:

1. Cyber-harassment with sexual connotations: It involves the "grooming" practice, in which an adult pretends to be a child or adolescent to establish a friendship in order to get the victim to send photographic material of a sexual nature (Panizo-Galence, 2011). It may involve disseminating sexual material without consent (Bégin, 2018; Cuenca-Pinqueras et al., 2020; Willard, 2007).

2. Doxing: Information is requested from someone and then disseminated without permission (Willard, 2007). It also encompasses spreading embarrassing photos, rumors, and false or harmful information about a person (Beran et al., 2012; Cowie, 2013).

3. Identity theft: A person appropriates another person's identity in digital media, which in some cases involves hacking into the victim's social networks (Baker et al., 2013).

4. Cyber-stalking: Repeatedly, a person seeks to establish contact with another without their consent (Jagavat & Choma, 2021). It involves monitoring and controlling a person's actions through social networks or other technological means without authorization (Cava et al., 2020; Zweig et al., 2013).

5. Cyber-violence: A digital violence can be seen in couple relationships (Cava & Buelga, 2018; Rodríguez-Domínguez et al., 2020), which involves manipulation, excess control, verbal, emotional or social abuse (Neelam, 2017).

6. Cyberbullying: An expression of cyber harassment between peers from the same school context (Neelam, 2017; Foley et al., 2015). It may involve making fun of the sexual orientation, race, religion or ethnic origin of another person through technological means (Rivadulla-López, & Rodríguez-Correa, 2019; Bégin, 2018); the "flaming" practice (online fights between the victim(s) and perpetrator(s)) (Willard, 2007), and the exclusion from different social networks (Cowie, 2013; Rincón-Rueda & Ávila-Díaz, 2014; Lee & Wu, 2018; Willard, 2007).

It has been found that victims of this kind of aggression tend to have adverse effects related to their mental health, like feelings of vulnerability, anger, sadness, shame, guilt, worry, fear and stress (Hassan et al., 2020; Choja & Nelson, 2016; Lu et al., 2018). Alternatively, these people could even show psychopathological symptoms related to depression and anxiety spectrums involving suicidal ideation (Cowie, 2013; Hassan et al., 2020; Hinduja & Patchin, 2010; Walker et al., 2011).

Regarding this situation, the victims are expected to report the aggression themselves (Ayub & Malik, 2020; Beran et al., 2012; Pereira et al., 2016). However, they face some difficulties in closing their cases, since they are asked to show evidence demonstrating the harassment and the whereabouts of the aggressor, elements which the victims do not always have (Halder & Jaishankar, 2011). In other cases, especially children and ad-

olescents tend to refrain from reporting their cases themselves (Hassan et al., 2020), commenting on the situation to someone they trust who may not be their parents (Kahn & Roshan, 2020; Neelam, 2017; Pereira et al., 2016; Wick et al., 2020).

Any of the cyber-harassment manifestations occur in the presence of unlimited third-party observers (Garaigordobil & Aliri, 2013; Lee & Wu, 2018), paradoxically, the report incidence is underreported regarding the general community (Carter, 2013). In the literature, it has been a questioned about the psychological aspects that may involve the reporting of this type of aggression, taking into account that it is not suitable to assess directly the assessment of "intentions" of the general community to report this type of situation due to the influence of social desirability (Krumpal, 2013; van der Schyff et al., 2022). Because of this, some authors have highlighted the role of attitudes, which influence an individual's intentions in reporting harassment events (Foster & Fullagar, 2018). Traditionally, attitudes have been defined as psychological tendencies, manifested by the evaluations of a particular entity with any degree of favor or disfavor, which brings together other constructs relating to cognition, emotion and behavior (Eagly & Chaiken, 2007; Rosenberg, 1956; Rosenberg, 1960; Ostrom, 1969).

In the literature, no study was found which assesses this same construct, as the case of Walsh et al.'s study (2010), which assessed the teachers' attitudes towards reporting cases of sexual abuse in children; and the study of Cesario et al. (2018), in which Sexual Harassment Reporting Attitudes were evaluated. Furthermore, the studies of Barlett et al. (2016) developed an attitude scale towards cyberbullying, and Gomes-Cavalcanti et al. (2021) validated this scale in Brazil. Other questionnaires concerning cyber-aggression have been validated through studies in other Latin American countries (e.g. Best et al., 2021).

In Colombia, for example, a prevalence of 30% of cyber-harassment in children and adolescents has been recognized (Redondo-Pacheco et al. 2018). However, it was not found any study that has measured attitudes toward reporting cyber-harassment cases. This country has been allied with other Latin-American countries related to the first Safe Internet Center in Latin America. This center has the skills and experience to protect, prevent, help and guide children, adolescents and young people on their journey through the Internet. It has the support of *Te Protejo*, *Red PaPaz,* and *INHOPE* (RedPaPaz, 2023. Viguías. https://viguias.org). Through this center, cases of children or adolescents who have experienced cyber-harassment can be reported.

This study was set to establish an initial instrument to assess attitudes towards reporting cases of cyber-harassment in children and adolescents directed to the Colombian general community. Once reached, it was aimed to confirm this structure and establish possible significant differences between any socio-demographic variable.

Method

An instrumental study was conducted between March 2021 and June 2021. A sequential, non-probabilistic sample was obtained involving Colombian participants. The sample size for the first validation study followed parameters for conducting Classic Theory Test analysis (CTT) and multivariate reduction analysis, defining at least 200 participants (Ferrando et al., 2022; Izquierdo et al., 2014; Osborne & Costello, 2004). For the confirmatory phase of the study, the sample size was estimated by the ratio 1:20 for conducting structural equation modeling analysis (Kline, 2011).

All participants were contacted via social media, and all of them gave their consent to par-

ticipate in the study, as one of the inclusion criteria for taking into account was the fact of being more than 18 years old and having access to technological devices with the internet.

Instrument development

The guidelines proposed by Muñiz & Fonseca-Pedrero (2019) were followed, under which a search of specialized literature was initiated. Once the theoretical framework was consolidated, 34 unidimensional items with a 5-point Likert scale (1 = *Totally disagree*, 2 = *Disagree*, 3 = *Not Agree neither Disagree*, 4 = *Agree*, 5 = *Totally agree*) were created. In order to obtain a content validity evidence that could show the items' representativeness (Muñiz, 1998; Roebianto et al., 2023), the scale was revised by eight judges (with expertise in psychometrics, developmental psychology, and cyber-psychology), who followed the recommendations by Skjong & Wentworth (2000). All of them evaluated items from 1 to 4, considering the criteria of relevance (the degree to which the items assess the content proposed), coherence (the logical relationship between items and construct) and clarity (content precision).

After obtaining the degree of agreement between judges, an Exploratory Factor Analysis (EFA) was performed to establish a preliminary test version. Then, a Confirmatory Factor Analysis (CFA) was performed to confirm the resulting structure. With these results, the socio-demographic variables were crossed to seek potential significant differences in the construct measured.

Data analysis

The degree of agreement was estimated using the Content Validity Index (CVI) proposed by Lawshee (1975). Previous to the EFA, a descriptive analysis was conducted, and items with kurtosis and skewness superior to the absolute value of 7 and 2, respectively, were excluded from the analysis for having extremely abnormal behavior (Fabrigar et al., 1999; Ferguson & Cox, 1993). In addition, the Homogeneity Index (HI) was calculated, and the Measure of Sampling Adequacy (MSA) values below .35 and .70 were set as cutoff criteria to exclude items (Blum et al., 2013; Lorenzo-Seva & Ferrando, 2021). Furthermore, the polychoric Kaiser-Meyer-Oklin (KMO) sampling adequacy test and Bartlett's sphericity test were obtained to establish the accuracy of data to perform an EFA (Méndez-Martínez & Rondón-Sepúlveda, 2012).

For the EFA, considering the sample size, the nature of data and the abnormal distribution of data, a parallel analysis was performed with an optimal implementation based on polychoric correlations. The models were appraised with a Diagonally Weighted Least Squares (DWLS) estimator for this analysis and the CFA. To accept the models, the following cutoff values were considered: RMSEA $\leq$ .08, SRMR $\leq$ .08, TLI $\geq$ .95, and CFI $\geq$ .95 (Hu & Bentler, 1999; Kline, 2011; Xia & Yang, 2019). Reliability estimates were obtained using Cronbach's Alpha (α) and McDonald's Omega (ω), which aimed at establishing the items' internal consistency. As complementary indices, the Factor Determinacy (FD) and the Latent Construct Reliability (HLat) were obtained. The first one shows how the factor scores are reasonable estimates of individual differences given a specific factor (Grice, 2001; Rodríguez et al., 2016), and the second one assesses how well the factor can be identified by the continuous latent response variables that underlie the observed item scores (Ferrando & Lorenzo-Seva, 2017).

Results of the final structure were crossed with the sociodemographic variables in search of

differences via a Kruskal-Wallis Test (H-test) (if an abnormal distribution was found *via* a Shapiro-Wilk-Francia normality test). The size effect was calculated with the epsilon squared coefficient (E^2_R) (Ventura-León, 2019).

For all analyses, a *p*-value of .05 was considered for statistical significance. The EFA was implemented with FACTOR 12.02.01 (Ferrando & Lorenzo-Seva, 2017), and the rest of the descriptive and bivariate analyses were performed in *R* software (R Core Team , 2022) using the packages: "stats" (R Core Team , 2022) "readxl" (Wickham & Bryan, 2023), "dplyr" (Wickham et al., 2023), "summarytools" (Comtois, 2022), "psychometric" (Fletcher, 2022), "usf" (Peters & Gruijters, 2021), "tm" (Feinerer & Hornik, 2023), "psych" (Revelle, 2024), "MVN" (Korkmaz et al., 2014), "lavaan" (Rosseel, 2012), "semTools" (Jorgensen et al., 2022).

Results

Sample- Participants

Two samples were obtained, one for the EFA and another for the CFA. Table 1 shows the main results; and a proportion test was performed for both samples to check the equivalence between the two samples.

Table 1
Descriptive sample statistics.

Characteristics		**Sample 1 (EFA)**	**Sample 2 (CFA)**	***p-value***
		220	**204**	
Age	Early adulthood	30(14%)	33(16%)	0.462
	Adulthood	181(82%)	162(79%)	0.532
	Late adulthood	9(5%)	9(4%)	1.000
Sex	Male	66 (30%)	76 (37%)	0.139
	Female	154 (70%)	128 (63%)	
City of residence	Bogotá	118 (54%)	138(68%)	**0.004***
	Other	102(46%)	66 (32%)	
Income level	Low	54(26%)	48(24%)	0.895
	Medium	121(55%)	125 (61%)	0.226
	High	45(20%)	31(15%)	0.199
Educational level	Elementary – High school	33(15%)	35(17%)	0.636
	Technical	33(15%)	25(12%)	0.496
	Bachelor	68(31%)	68(33%)	0.667
	Postgraduate	86(39%)	76(37%)	0.773
Working experience with children	Yes	88 (40%)	109(53%)	**0.007***
	No	132 (60%)	95(47%)	
Previous episode of cyber-harassment	Yes	16 (7%)	17(8%)	0.681
	No	203(93%)	185(92%)	
Knowledge of official reporting mechanisms of cyber-harassment	Yes	69(31%)	72(35%)	0.450
	No	151(69%)	132(65%)	

Note. Significant differences were found in the variables: city of residence and previous working experience with children.

Content validity evidence

Scores above .75 on the CVI were sufficient to consider the essentiality of an item (Lawshe, 1975). In this case, 22 were retained without any modifications (64%), 3 were modified according to the judge's criteria (9%) and 8 items were excluded from the scale's final version (24%). Table 2 shows the main results obtained.

Table 2
Content Validity Evidence Results.

Item	Clarity	Relevance	Coherence	Average CVI	Decision
I1	1,00	1,00	1,00	1,00	M
I2	0,50	0,50	0,75	0,58	Ex
I3	0,75	1,00	1,00	0,92	M
I4	0,50	1,00	1,00	0,83	M
I5	0,75	1,00	0,75	0,83	M
I6	0,75	1,00	1,00	0,92	M
I7	0,75	1,00	1,00	0,92	M
I8	0,75	0,75	1,00	0,83	M
I9	1,00	1,00	1,00	1,00	M
I10	1,00	1,00	1,00	1,00	M
I11	0,50	-0,50	-0,75	-0,25	Ex
I12	1,00	0,50	0,75	0,75	Ad
I13	0,75	0,75	0,75	0,75	Ad
I14	1,00	1,00	1,00	1,00	M
I15	0,50	1,00	1,00	0,83	M
I16	0,75	0,25	0,25	0,51	Ex
I17	0,50	0,75	1,00	0,75	Ad
I18	0,75	0,75	0,50	0,66	Ex
I19	0,75	1,00	1,00	0,92	M
I20	0,50	0,50	0,25	0,40	Ex
I21	0,50	0,75	0,75	0,66	Ex
I22	0,75	0,00	0,00	0,25	Ex
I23	0,50	0,75	0,50	0,58	Ex
I24	1,00	1,00	1,00	1,00	M
I25	0,75	1,00	1,00	0,92	M
I27	1,00	0,75	0,75	0,83	M
I28	1,00	1,00	1,00	1,00	M
I29	0,75	1,00	1,00	0,92	M
I30	1,00	1,00	1,00	1,00	M
I31	1,00	0,75	0,75	0,83	M
I32	1,00	0,50	0,75	0,75	M
I33	1,00	1,00	1,00	1,00	M
I34	1,00	1,00	1,00	1,00	M

Note. *I* = Items; *M* = Maintained without any modifications; *Ad* = Adjusted to judge's recommendations and Ex = Excluded from the instrument.

Exploratory Factor Analysis

Regarding the results of shape forms, 13 items were excluded from the analysis, leaving 12 items. After the HI analysis, another 3 more items were excluded, and for the MSA, one additional item was excluded, leaving a total of 8 items for the EFA. An optimal KMO (.88) and a significant Bartlett's Sphericity test ($p < .001$) were obtained, confirming the polychoric matrix adequacy to perform an EFA. The Mardia multivariate normality test showed an abnormal data distribution ($p > .05$).

The parallel analysis yielded 1 main factor; one item showed a high presence of residual, being excluded from the final version. The resulting structure explains 64.27% of the variance; its fit was obtained RMSEA = .07, SRMR = .05, CFI = .99 and TLI = .99. All items had adequate communalities (Mean = .60, Min = .30, Max = .89) and optimal factor loadings (Mean = .76, Min = .55, Max = .95), as well as good internal consistency indices (α = .90, ω = .91, FD = .98, HLat = .96). Table 3 shows the resulting scale structure, which includes the final 7 items, its factors loadings and communalities, and the respective descriptive statistics, as it has been suggested by the specialized literature (Muñíz & Fonseca-Pedrero, 2019).

Table 3
Final Scale Structure.

	Items	L	C	HI	M	Sk	K
A1	I am responsible for reporting a cyber-harassment situation that occurs to any child and/or adolescent [*Soy responsable de reportar una situación de ciberacoso que se le presente a cualquier niño, niña y/o adolescente*].	0.76	0.57	0.58	4.50	-2.16	4.33
A2	Depending on the severity, I would report a case of cyber-harassment [*Dependiendo de la gravedad, reportaría un caso de ciberacoso*].	0.66	0.43	0.43	4.44	-1.93	2.93
A3	When seeing a material in which a boy, girl or adolescent is ridiculed on social networks, I would report it [*Al ver un material en el que se ridiculice a un niño, niña o adolescente en redes sociales lo reportaría*].	0.94	0.88	4.56	4.56	-2.14	4.82
A4	I feel safe to report a case of cyber-harassment, even if I am not completely clear about the situation [*Me siento seguro/a de reportar un caso de ciberacoso, aunque no tenga total claridad de la situación*].	0.61	0.37	3.54	3.54	-0.45	-0.87
A5	The content of social networks in which a child or adolescent is mocked must be reported [*El contenido de redes sociales en el que se hace burla de algún niño, niña o adolescente debe ser reportado*].	0.95	0.89	4.61	4.61	-2.38	6.21
A6	If I see on social networks that the publication of a child or adolescent receives offensive comments, I would report it [*Si veo en redes sociales que la publicación de un niño, niña o adolescente recibe comentarios ofensivos lo reportaría*].	0.88	0.78	4.37	4.37	-1.56	1.66
A7	I would feel more at ease reporting a case of cyberbullying to the authorities rather than to the victim's family [*Me sentiría más tranquilo/a informando un caso de ciberacoso ante las autoridades en lugar de hacerlo ante la familia de la víctima*].	0.55	0.30	3.75	3.76	-0.66	-0.62

Note. *L* = Factor Loadings; *C* = Communality; *HI* = Homogeneity Index; *M* = Mean; *Sk* = Skewness; *K* = Kurtosis.

Confirmatory Factor Analysis

After the resulting structure, a new sample was collected (see Table 1) and a multivariate Mardia normality test was performed, resulting in an abnormal distribution ($p > .05$). A CFA confirmed the structure found in the EFA, RMSEA = .03, SRMR = .04, CFI = .96, TLI = .95. All items had adequate factor loadings (Mean = .63, Min = .438, Max = .816). This model had good internal consistency indices (α = .80, ω = .89). The results can be seen in the following SEM graphic (Figure 1).

Bivariate analysis

Regarding this analysis, a Shapiro-Wilk-Francia normality test was first performed, finding an abnormal distribution ($p < .05$). A Kruskal Wallis test was conducted (with the whole sample, $n = 424$), finding differences between age, education level, working experience with children and knowledge of the official report mechanisms. However, it was found to be almost a null effect size. Table 4 shows the results.

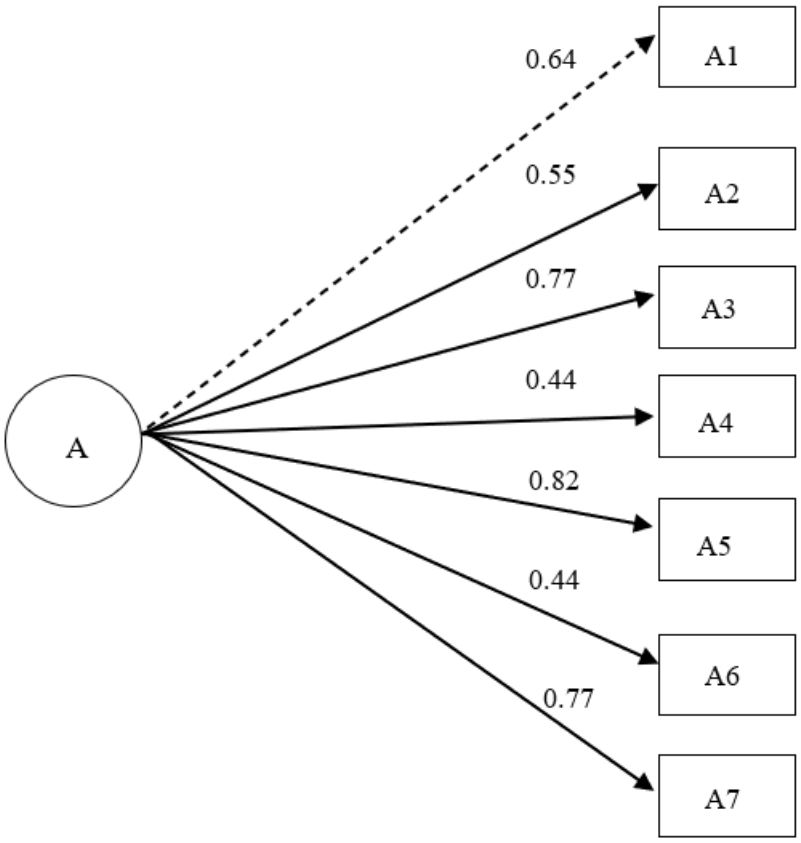


Figure 1.
Graphical CFA representation.

Table 4
Bi-variate analysis.

Characteristics		Mean	*p-value*	Size effect
Age	Early adulthood	27.11	**< 0.001***	0.06
	Adulthood	30.30		
	Late adulthood	28.33		
Sex	Male	29.43	0.738	-
	Female	29.90		
City of residence	Bogotá	29.87	0.485	-
	Other	29.54		
Income level	Low	30.35	0.236	-
	Medium	29.75		
	High	28.89		
Educational level	Elementary – High school	30.35	**0.016***	0.02
	Technical	29.98		
	Bachelor	28.89		
	Postgraduate	30.11		
Working experience with children	Yes	29.47	**0.026***	0.01
	No	29.97		
Previous episode of cyber-harassment	Yes	30.06	0.639	-
	No	29.70		
Knowledge of official reporting mechanisms of cyber-harassment	Yes	29.03	**0.013***	0.01
	No	30.81		

Note. Early adulthoods had the least favorable level of attitudes, the size effect only was calculated for the variables which were significant.

Discussion

The present study presents the first initial validation study of an attitude scale towards reporting cases presented in children and adolescents directed to the general community. As a strength of the study, the resulting scale has a content validity evidence, a piece of evidence that is not usually reported by health and social sciences studies (Zapata-Ospina & García-Valencia, 2022), which helped to adjust items and remove the ones that did fit the expert judge's criteria for not being a good reflect of the measured construct.

Along the same lines, this study has two main internal structure validity evidences, as suggested when conducting psychometric studies (Lloret-Segura et al., 2014), which confirmed the unidimensional structure of the construct of the study, having optimal reliability indices. To accomplish the model fit, it was beneficial to include only direct items, as suggested by specialists in the literature (Suárez-Álvarez, et al., 2018), and many item filters before the EFA, such as the shape forms and the HI and MSA. This structure is theoretically coherent with the conceptualization of attitudes, which states that the attitudes' components are complicated to segregate, suggesting the existence of a unidimensional structure to comprehend this construct (Valdivieso-Taborga, 2013; Makanyeza, 2014). Moreover, this scale has

a similar length and structure to another construct related scales (e.g. Barlett et al., 2016; Cesario et al., 2018; Gomes-Cavalcanti et al., 2021; Walsh et al., 2010).

Regarding the bivariate analysis, an hypothesis was to find significant differences between age (Sakellari et al., 2022; Stefani, 2005), sex (Garaigordobil & Aliri, 2013; Cava et al., 2020; Rivadulla-López & Rodríguez-Correa, 2019; Rotundo et al., 2001; Yee et al., 2015), education and income level (Sánchez-Díaz, 2019). Even though initial differences between age and education level were found, the size effect was almost null. The analysis included variables relating to working experience with children, previous episodes of cyber-harassment and knowledge of official reporting mechanisms of cyber-harassment because it has been stated that the direct experiences with aspects of the attitudes object may have an incidence in the attitude formation (Verplanken & Orbell, 2022). However, this variable did not present any significant difference. The last results help to conclude that there may be no differences of attitude level in any socio-demographic characteristics.

Our study may help implement preventive interventions involving cases of cyber-harassment among the general community, taking into account the important role that attitudes play in reporting cases of harassment (Verplanken & Orbell, 2022). However, some limitations, such as the lack of convergent and divergent validity evidence, can be considered. It has been stated that regard may be an influential variable in attitude formation (Foster & Fullagar, 2018), even social desirability (Krumpal, 2013) or acquiescence (Valentini, 2017). Furthermore, the sampling method is an element that limits the results obtained, reflected in the differences found between the samples. For future research, more robust sampling methods would be needed. In addition, it is recommended to establish association models to find which variables may be included in the attitudes towards reporting cases of cyber harassment in children and adolescents.

Data availability

The datasets generated during the project and analyzed during the current study are available in the Repository Open Science Framework: https://osf.io/z2w9y.

Funding

This study did not receive any economic funding.

Acknowledgements

To Milton Bermúdez Jaimes who encouraged this project, to Mylene Sánchez and Gema Castillo who helped recollect data, and to the judges who revised the scale: Lucila Cárdenas Niño, Lina Guevara Bedoya, Flor Ángela León Grisales, Laura Álvarez Graciano, Mauricio Herrea López, Leonardo Rodríguez Cely and Rosa Isabel Galvis Vargas.

References

Ayub, S., & Malik, F. (2020). Experiences of cyber harassment and social adjustment in female university students: Moderating role of self-efficacy. *Pakistan Journal of Psychological Research, 35*(3), 523-544. https://doi.org/10.33824/PJPR.2020.35.3.28

Baker, L., Campbell, M., Barreto, E., Campbell, M., & Barreto, E. (2013). Understanding technology-relat-

ed violence against women: Types of violence and women's experiences. *Learning Network Brief, 6*, 1-7. https://gbvlearningnetwork.ca/our-work/briefs/briefpdfs/LB-06.pdf

Barlett, C. P., Helmstetter, K., & Gentile, D. A. (2016). The development of a new cyberbullying attitude measure. *Computers in Human Behavior, 64*, 906-913. http://doi.org/10.1016/j.chb.2016.08.013

Bégin, M. (2018). Cyberbulling. Una revisión de investigaciones internacionales sobre representaciones, prevalencias, efectos y explicaciones del fenómeno. Re-presentaciones. Periodismo. *Comunicación y Sociedad, 10*, 52-78. https://www.revistas.usach.cl

Beran, T. N., Rinaldi, C., Bickham, D. S., & Rich, M. (2012). Evidence for the need to support adolescents dealing with harassment and cyber-harassment: Prevalence, progression, and impact. *School Psychology International, 33*(5), 562-576. https://doi.org/10.1177/0143034312446976

Best, S., Ré, N., Corcoran, L., & McGuckin, C. (2021). Validez de contenido del Cuestionario de Ciberagresión. *Revista Evaluar, 21*(2), 1-16. https://revistas.unc.edu.ar

Blum, D., Auné, S., Galibert, M. S., & Attorresi, H. (2013). Criterios para la eliminación de ítems de un Test de Analogías Figurales. *Summa Psicológica UST, 10*(2), 49-56. https://doi.org/10.18774/448x.2013.10.140

Carter, M. A. (2013). Third party observers witnessing cyber bullying on social media sites. *Procedia-Social and Behavioral Sciences, 84*, 1296-1309. https://doi.org/10.1016/j.sbspro.2013.06.747

Cava, M-. J., & Buelga, S. (2018). Propiedades psicométricas de la escala de Ciber-Violencia en Parejas Adolescentes (Cib-VPA). *Suma Psicológica, 25*(1), 51-61. http://doi.org/10.14349/sumapsi.2018.v25.n1.6

Cava, M-. J., Martínez-Ferrer, B., Buelga, S., & Carrascosa, L. (2020). Sexist attitudes, romantic myths, and offline dating violence as predictors of cyber dating violence perpetration in adolescents. *Computers in Human Behavior, 111*. https://doi.org/10.1016/j.chb.2020.106449

Cesario, B., Parks-Stamm, E., & Turgut, M. (2018). Initial assessment of the psychometric properties of the Sexual Harassment Reporting Attitudes Scale. *Cogent Psychology, 5*(1). https://doi.org/10.1080/23311908.2018.1517629

Choja, O., & Nelson, O. (2016). Psychological Violence and the bane of Cyber-Harassment against Women: An Experiential Inquest on Facebook. *Gender & Behaviour, 14*(3), 7589–7608.

Comtois, D. (2022). *summarytools: Tools to quickly and neatly summarize data. R* package (Version 1.0.1). [Software de cómputo]. https://CRAN.R-project.org/package=summarytools

Cowie, H. (2013). El impacto emocional y las consecuencias del ciberacoso. *Revista Digital de la Asociación CONVIVES, 3*, 16-24. http://convivenciaenlaescuela.es

Cuenca-Piqueras, C., Fernández-Prados, J. S., & González-Moreno, M. J. (2020). Face-to-face versus online harassment of European women: Importance of date and place of birth. *Sexuality & Culture, 24*(1), 157-173. https://doi.org/10.1007/s12119-019-09632-4

Eagly, A. H., & Chaiken, S. (2007). The advantages of an inclusive definition of attitude. *Social Cognition, 25*(5), 582-602. https://doi.org/10.1521/soco.2007.25.5.582

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*(3), 272-299. https://doi.org/10.1037/1082-989X.4.3.272

Feinerer, I., & Hornik, K. (2023). *tm: Text Mining Package. R* package (Version 0.7-11). [Software de cómputo]. https://CRAN.R-project.org/package=tm

Ferrando, P. J., & Lorenzo-Seva, U. (2017). Program FACTOR at 10: Origins, development and future directions. *Psicothema, 29*(2), 236-240. https://doi.org/10.7334/psicothema2016.304

Ferrando, P. J., Lorenzo-Seva, U., Hernández-Dorado, A., & Muñiz, J. (2022). Decálogo para el análisis factorial de los ítems de un test [Decalogue for the factor anal-

ysis of test items]. *Psicothema, 34*(1), 7-17. https://doi.org/10.7334/psicothema2021.456

Ferguson, E., & Cox, T. (1993). Exploratory factor analysis: A users' guide. International *Journal of Selection and Assessment, 1*(2), 84-94. https://doi.org/10.1111/j.1468-2389.1993.tb00092.x

Fletcher, T. D. (2022). *psychometric: Applied psychometric theory. R* package (Version 2.3). [Software de cómputo]. https://CRAN.R-project.org/package=psychometric

Foley, C., May, D. C., Blevins, K. R., & Akers, J. (2015). An exploratory analysis of cyber-harassment of K-12 teachers by parents in public school settings. *Educational Policy, 29*(6), 879-902. https://doi.org/10.1177/0895904814550071

Foster, P. J., & Fullagar, C. J. (2018). Why don't we report sexual harassment? An application of the theory of planned behavior. *Basic and Applied Social Psychology, 40*(3), 148-160. https://doi.org/10.1080/01973533.2018.1449747

Garaigordobil, M., & Aliri, J. (2013). Ciberacoso ("cyberbullying") en el país vasco: Diferencias de sexo en víctimas, agresores y observadores. *Behavioral Psychology/Psicología Conductual, 21*(3), 461-474. https://www.behavioralpsycho.com

Gomes-Cavalcanti, J., de Lima-Pinto, A. V., Pimentel, C. E., & Fernandes de Melo, L. M. (2021). Psychometric parameters of the Cyberbullying Attitude Scale in Brazil. *Current Psychology, 40*, 3897-3907. http://doi.org/10.1007/s12144-019-00326-6

Grice, J. W. (2001). Computing and evaluating factor scores. *Psychological Methods, 6*(4), 430-450. https://doi.org/10.1037/1082-989X.6.4.430

Halder, D., & Jaishankar, K. (2011). Cyber gender harassment and secondary victimization: A comparative analysis of the United States, the UK, and India. *Victims & Offenders, 6*(4), 386-398. http://dx.doi.org/10.1080/15564886.2011.607402

Hassan, F. M., Khalifa, F. N., El Desouky, E. D., Salem, M. R., & Ali, M. M. (2020). Cyber violence pattern and related factors: Online survey of females in Egypt. *Egyptian Journal of Forensic Sciences, 10*, article number 6. http://doi.org/10.1186/s41935-020-0180-0

Hinduja, S., & Patchin, J. W. (2010). Bullying, cyberbullying, and suicide. *Archives of Suicide Research, 14*(3), 206-221. https://doi.org/10.1080/13811118.2010.494133

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

Izquierdo, I., Olea, J., & Abad, F. J. (2014). Exploratory factor analysis in validation studies: Uses and recommendations. *Psicothema, 26*(3), 395-400. https://doi.org/10.7334/psicothema2013.349

Jagayat, A., & Choma, B. L. (2021). Cyber-aggression towards women: Measurement and psychological predictors in gaming communities. *Computers in Human Behavior, 120*. https://doi.org/10.1016/j.chb.2021.106753

Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., & Rosseel, Y. (2022). *semTools: Useful tools for structural equation modeling. R package* (Version 0.5-6). [Software de cómputo]. https://CRAN.R-project.org/package=semTools

Khan, A. U., & Roshan, R. (2020). Youth and cyber harassment: Reporting mechanism of cyber harassment by the KPK universities students. *The Dialogue, 15*(1), 74-82. https://journals.qurtuba.edu.pk/ojs/index.php/thedialogue/about

Kline, R. B. (2011). *Principles and practice of structural equation modeling* (3rd ed.). Guilford Press.

Korkmaz, S., Goksuluk, D., & Zararsiz, G. (2014). MVN: An R package for assessing multivariate normality. *The R Journal, 6*(2), 151-162. https://journal.r-project.org

Krumpal, I. (2013). Determinants of social desirability bias in sensitive surveys: A literature review. *Quality & Quantity, 47*(4), 2025-2047. http://doi.org/10.1007/s11135-011-9640-9

Lee, Y. C., & Wu, W. L-. (2018). Factors in cyber bully-

ing: The attitude-social influence-efficacy model [Factores en el acoso cibernético: El modelo de actitud-influencia social, eficacia]. *Annals of Psychology/ Anales de Psicología, 34*(2), 324-331. https://doi.org/10.6018/analesps.34.2.295411

Lloret-Segura, S., Ferreres-Traver, A., Hernández-Baeza, A., & Tomás-Marco, I. (2014). El análisis factorial exploratorio de los ítems: Una guía práctica, revisada y actualizada [Exploratory ítem factor analysis: A practical guide, revised and updated.]. *Anales de Psicología/Annals of Psychology, 30*(3), 1151-1169. https://doi.org/10.6018/analesps.30.3.199361

Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology, 28*(4), 563-575. https://doi.org/10.1111/j.1744-6570.1975.tb01393.x

Lorenzo-Seva, U., & Ferrando, P. J. (2021). MSA: The forgotten index for identifying inappropriate items before computing exploratory item factor analysis. *Methodology, 17*(4), 296-306. https://doi.org/10.5964/meth.7185

Lu, Y., Van Ouytsel, J., Walrave, M., Ponnet, K., & Temple, J. R. (2018). Cross-sectional and temporal associations between cyber dating abuse victimization and mental health and substance use outcomes. *Journal of Adolescence, 65(*1), 1-5. https://doi.org/10.1016/j.adolescence.2018.02.009

Makanyeza, C. (2014). Measuring consumer attitude towards imported poultry meat products in a developing market: An assessment of reliability, validity and dimensionality of the tri-component attitude model. *Mediterranean Journal of Social Sciences, 5*(20), 874-881. http://dx.doi.org/10.5901/mjss.2014.v5n20p874

Méndez-Martínez, C. & Rondón-Sepúlveda, M. A. (2012). Introducción al análisis factorial exploratorio [Introduction to exploratory factor analysis (EFA)]. *Revista Colombiana de Psiquiatría, 41*(1), 197-207. https://doi.org/10.1016/S0034-7450(14)60077-9

Muñiz, J. (1998). La medición de lo psicológico. *Psicothema, 10(*1), 1-21. https://www.psicothema.com

Muñiz, J., & Fonseca-Pedrero, E. (2019). Diez pasos para la construcción de un test [Ten steps to make a test]. *Psicothema, 31*(1), 7-16. https://doi.org/10.7334/psicothema2018.291

Neelam, A. (2017). Law relating to cyber harassment in India. International Journal of Recent *Research Aspects, 4(*4), 44-46. https://www.ijrra.net

Osborne, J. W., & Costello, A. B. (2004). Sample size and subject to item ratio in principle components analysis. *Practical Assessment, Research and Evaluation, 9(*11), 1-9. https://doi.org/10.7275/ktzq-jq66

Ostrom, T. M. (1969). The relationship between the affective, behavioral, and cognitive components of attitude. *Journal of Experimental Social Psychology, 5(*1), 12-30. https://doi.org/10.1016/0022-1031(69)90003-1

Panizo-Galence, V. (2011). El ciber-acoso con intención sexual y el child-grooming. *Quadernos de Criminología: Revista de Criminología y Ciencias Forenses, 15*, 22-33. https://revistaqdc.es

Pereira, F., Spitzberg, B. H., & Matos, M. (2016). Cyber-harassment victimization in Portugal: Prevalence, fear and help-seeking among adolescents. *Computers in Human Behavior, 62*, 136-146. http://doi.org/10.1016/j.chb.2016.03.039

Peters, G. J. Y., & Gruijters, S. (2021). *ufs: A collection of utilities. R package* (Version 0.4.5.). [Software de cómputo]. https://r-packages.gitlab.io/ufs

Redondo-Pacheco, J., Luzardo-Briceño, M., Inglés-Saura, C. J., & Rivas-Rivero, E. (2018). Ciberacoso en una muestra de adolescentes de instituciones educativas de Bucaramanga. *Psychologia, 12*(1), 35-44. http://doi.org/10.21500/19002386.3366

RedPaPaz (2023). *Viguías*. https://viguias.org

Revelle, W. (2024). *psych: Procedures for psychological, psychometric, and personality research* (Version 2.4.3). [Software de cómputo]. https://CRAN.R-project.org/package=psych

Rincón-Rueda, A. I., & Ávila-Díaz, W. D. (2014). Simbiosis vital para describir el ciberbullying en Colombia. *Revista Científica General José María Córdova 12*(14), 149-164. https://doi.org/10.21830/19006586.61

Rivadulla-López, J. C., & Rodríguez-Correa, M. (2019). Ciberacoso escolar: Experiencias y propuestas de jóvenes universitarios. *Revista Iberoamericana de Educación a Distancia. 22*(2), 179-201. http://doi.org/10.5944/ried.22.2.23541

R Core Team. (2022). *RStudio: Integrated development environment for R* (2022). (Version 12.0) [Software de cómputo]. https://www.rstudio.com

Roebianto, A., Savitiri, S. I., Aulia, I., Suciyana, A., & Mubarokah, L. (2023). Content validity: Definition and procedure of content validation in psychological research. *Testing, Psychometrics, Methodology in Applied Psychology, 30*(1), 5-18. http://doi.org/10.4473/TPM30.1.1

Rodriguez, A., Reise, S. P., & Haviland, M. G. (2016). Applying bifactor statistical indices in the evaluation of psychological measures. *Journal of Personality Assessment, 98*(3), 223-237. https://doi.org/10.1080/00223891.2015.1089249

Rodríguez-Domínguez, C., Pérez-Moreno, P. J., & Durán, M. (2020). Cyber dating violence: A review of its research methodology [Ciberviolencia en las relaciones de pareja: Una revisión sobre su metodología de investigación]. *Annals of Psychology/Anales de Psicología, 36*(2), 200-209. https://doi.org/10.6018/analesps.370451

Rosenberg, M. J. (1956). Cognitive structure and attitudinal affect. *The Journal of Abnormal and Social Psychology, 53*(3), 367-372. https://doi.org/10.1037/h0044579

Rosenberg, M. J. (1960). A structural theory of attitude dynamics. *Public Opinion Quarterly, 24*(2), 319-340. http://doi.org/10.1086/266951

Rotundo, M., Nguyen, D. -H., & Sackett, P. R. (2001). A meta-analytic review of gender differences in perceptions of sexual harassment. *Journal of Applied Psychology, 86*(5), 914-922. http://doi.org/10.1037//0021-9010.86.5.914

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

Sakellari, E., Berglund, M., Santala, E., Bacatum, C. M. J., Sousa, J. E. X. F., Aarnio, H., Kubiliuté, L., Prapas, C., & Lagiou, A. (2022). The perceptions of sexual harassment among adolescents of four european countries. *Children, 9*(10). https://doi.org/10.3390/children9101551

Sánchez-Díaz, L. (2019). Street harassment perception and its relations with self-objectification of women. *Inter Disciplina, 7*(17), 153-170. http://doi.org/10.22201/ceiich.24485705e.2019.17.67534

Skjong, R., & Wentworth, B. H. (2000). *Expert judgment and risk perception*. Paper presented in the Eleventh International Offshore and Polar Engineering Conference, Stavanger, Norway. https://onepetro.org

Stefani, D. (2005). Teoría de la acción razonada: Una propuesta de evaluación cuali-cuantitativa de las creencias acerca de la institucionalización geriátrica. *Revista Evaluar, 5*(1), 22-37. http://doi.org/10.35670/1667-4545.v5.n1.538

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

Valdivieso-Taborga, C. E. (2013). Comparación de los modelos formativo, reflexivo y de antecedentes de evaluación estudiantil del servicio de docencia. *Revista de Métodos Cuantitativos para la Economía y la Empresa, 16*, 95-120. https://doi.org/10.46661/revmetodoscuanteconempresa.2182

Tokunaga, R. S. (2010). Following you home from school: A critical review and synthesis of research on cyberbullying victimization. *Computers in Human Behavior, 26(*3), 277-287. http://doi.org/10.1016/j.chb.2009.11.014

Valentini, F. (2017). Influência e controle da aquiescência na análise fatorial. *Avaliação Psicológica, 16*(2), 120-251. http://doi.org/10.15689/ap.2017.1602.ed

van der Schyff, K., Flowerday, S., & Renaud, K. (2022). Socially desirable responding within the context of privacy-related research: A personality perspective.

South African Journal of Information Management, 24(1), a1507. https://doi.org/10.4102/sajim.v24i1.1507

van Laer, T. (2014). The means to justify the end: Combating cyber harassment in social media. *Journal of Business Ethics, 123*(1), 85-98. http://doi.org/10.1007/s10551-013-1806-z

Ventura-León, J. L. (2019). Tamaño del efecto para Kruskal-Wallis: Aportes al artículo de Domínguez-González et al. *Investigación en Educación Médica, 8*(30), 135-136. http://doi.org/10.1016/j.riem.2017.07.002

Verplanken, B., & Orbell, S. (2022). Attitudes, habits, and behavior change. *Annual Review of Psychology, 73*, 327-352. https://doi.org/10.1146/annurev-psych-020821-011744

Walker, C. M., Sockman, B. R., & Koehn, S. (2011). An exploratory study of cyberbullying with undergraduate university students. *TechTrends, 55*(2), 31-38. https://doi.org/10.1007/s11528-011-0481-0

Walsh, K., Rassafiani, M., Mathews, B., Farrell, A., & Butler, D. (2010). Teachers' attitudes toward reporting child sexual abuse: Problems with existing research leading to new scale development. *Journal of Child Sexual Abuse, 19*(3), 310-336. https://doi.org/10.1080/10538711003781392

Wick, S. E., Nguyen, A. P., West, J. T., Nagoshi, C. T., Jordan, C., & Lehmann, P. (2020). Cyber harassment, coping, and psychological maladjustment in college students. *College Student Journal, 54*(1), 77-87. https://projectinnovationaustin.com/college-student-journal

Wickham, H., & Bryan, J. (2023). *readxl: Read Excel Files. R package* (Version 1.4.2). [Software de cómputo]. https://CRAN.R-project.org/package=readxl

Wickham, H., François, R., Henry, L., Müller, K., & Vaughan, D. (2023). *dplyr: A grammar of data manipulation. R package* (Version 1.1.2). [Software de cómputo]. https://CRAN.R-project.org/package=dplyr

Willard, N. E. (2007). *Cyberbullying and cyberthreats: Responding to the challenge of online social aggression, threats, and distress*. Research Press.

Xia, Y., & Yang, Y. (2019). 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*(1), 409-428. https://doi.org/10.3758/s13428-018-1055-2

Yee, M. W., Alagappar, P. N., & Ngeow, Y. M. (2015). Differences in the perception of sexual harassment by gender and ethnicity among selected Malaysian undergraduates. *Gender, Technology and Development, 19*(2), 204-230. https://doi.org/10.1177/0971852415578042

Zapata-Ospina, J. P., & García-Valencia, J. (2022). Validity based on content: A challenge in health measurement scales. *Journal of Health Psychology, 27*(2), 481-493. https://doi.org/10.1177/1359105320953477

Zweig, J. M., Dank, M., Yahner, J., & Lachman, P. (2013). The rate of cyber dating abuse among teens and how it relates to other forms of teen dating violence. *Journal of Youth and Adolescence, 42*(7), 1063-1077. http://doi.org/10.1007/s10964-013-9922-8