

Autoconcepto en matemáticas y estadística en estudiantes de secundaria: diseño y validación de dos cuestionarios breves

Self-concept in mathematics and statistics in high school students: design and validation of two short questionnaires

José Hernando Ávila-Toscano*¹ ; Leonardo José Vargas-Delgado¹ ;
Yurley Alejandra Badillo-Rueda¹ ; Kevin José Goenaga Pineda² 

¹Universidad del Atlántico, Colombia; ²Instituto de Ciências Matemáticas e de Computação,
Universidade Federal de São Carlos, Brasil

Resumen

El autoconcepto académico ha sido estudiado con asiduidad en psicología y educación, considerándolo un predictor relevante de logro educativo. Los teóricos recomiendan que este constructo se evalúe diferenciando dominios específicos como la matemática, la estadística y otras áreas, para garantizar una valoración puntual y contextualizada. Este estudio fue desarrollado con el fin de diseñar y validar dos cuestionarios breves para estudiantes de secundaria, que evalúan dominios específicos de autoconcepto en matemáticas y estadística. A través de un diseño instrumental se construyó un banco de ítems que posteriormente configuraron cuestionarios breves de seis reactivos. Mediante dos estudios independientes con 326 y 381 participantes respectivamente, se probó la estructura factorial de los cuestionarios. El Análisis Factorial Confirmatorio identificó que ambos instrumentos muestran adecuadas propiedades psicométricas, con buenos índices de ajuste, así como alta consistencia interna. Este estudio pone al servicio de educadores, profesionales de la psicología e investigadores de estas y otras áreas afines, dos instrumentos breves de dominio específico de autoconcepto matemático y estadístico, de fácil comprensión, que permite un diligenciamiento rápido y eficiente, los cuales pueden ser de utilidad para la investigación aplicada.

Palabras clave: Evaluación de la educación, Instrumento de medida, Autoconcepto, Psicología de la educación, Análisis factorial.

* Para correspondencia: José Hernando Ávila-Toscano, PhD. joseavila@mail.uniatlantico.edu.co

Abstract

Academic self-concept has been studied assiduously in psychology and education, being considered as a relevant predictor of educational achievement. Theorists recommend that this construct be assessed by differentiating specific domains such as mathematics, statistics and other areas, to ensure a timely and contextualized assessment. This study was developed in order to design and validate two brief questionnaires for high school students, which assess specific domains of self-concept in mathematics and statistics. An instrumental design was used to construct a bank of items that subsequently formed short questionnaires consisting of six items. The factorial structure of the two questionnaires was tested by means of two independent studies with 326 and 381 participants, respectively. The Confirmatory Factor Analysis identified that both instruments show adequate psychometric properties, with good fit indices, as well as high internal consistency. This study makes available to educators, psychology professionals and researchers in these and other related areas, two brief domain-specific instruments of mathematical and statistical self-concept, easy to understand, allowing a quick and efficient completion, which can be useful for applied research.

Keywords: Educational evaluation, Measuring instruments, Self-concept, Educational psychology, Factor analysis.

1. INTRODUCCIÓN

El autoconcepto es de interés para disciplinas interconectadas como educación y psicología, dadas sus notables implicaciones con el aprendizaje. Aunque no ha sido sencillo el camino para alcanzar una definición, hoy se cuenta con modelos teóricos para comprender su acepción general y las delimitaciones específicas.

El autoconcepto general se entiende como la perspectiva que una persona tiene de sí en contacto con un entorno que le es relevante (Shavelson et al., 1976; Marsh y Hau, 2004; Wu et al., 2021). Suma, por lo tanto, todas las percepciones personales en interacción con el mundo exterior (González et al., 2016). Esta capacidad de análisis innata e individual, implica distintos niveles de habilidades y destrezas entrelazadas a partir de las cuales la persona pondera su desempeño en relación con expectativas personales y sociales, lo que le otorga un carácter evaluativo o descriptivo (Filiz et al., 2020). El autoconcepto representa una retroalimentación completa del desempeño a partir de la comparación social (Niepel et al., 2014) y consigue ser un mediador que afecta a procesos mentales y de acción (Möller et al, 2020).

El autoconcepto se entiende a partir de un modelo teórico multidimensional basado en una estructura jerárquica que lo divide en dos grandes subcategorías, el académico y el no académico (Shavelson et al., 1976). El primero, es considerado una capacidad de juicio o razón de todas aquellas competencias escolares o de un dominio específico (Scherrer y

Preckel, 2019), mientras que el segundo se centraliza en una visión de aspectos sociales y de la vida en contexto (Goble et al., 2021).

Cada dimensión tiene subdivisiones que le otorgan el carácter jerárquico (Shavelson et al., 1976). El autoconcepto académico se divide en materias específicas y luego en factores más precisos de cada materia, mientras que el no académico se divide en físico, social y emocional, que a su vez se subdividen en factores más específicos. Esto indica que el constructo es cada vez más particular y varía según la cultura, la edad y las experiencias obtenidas en cada situación vivida, por lo que también tiene una connotación evolutiva (Bellmore y Chillessen, 2006; Marsh et al., 1984).

Este modelo jerárquico sustenta bien las relaciones entre formas de autoconcepto específico y constructos relacionados, verbigracia, el rendimiento académico correlaciona mejor con el autoconcepto académico que con los componentes globales y no académicos del autoconcepto (Marsh y Hau, 2004). Ajustado a esa lógica, posteriormente Marsh formuló ajustes al modelo jerárquico proponiendo una estructura con dominios específicos dentro del autoconcepto académico, que distingue entre autoconcepto matemático y verbal (Marsh, 1993; Marsh y Hau, 2004; Marsh et al., 1984; Marsh y O'Neill, 1984). Su propuesta describe el autoconcepto como una dimensión psicológica multidimensional, organizada jerárquicamente (autoconcepto general y autoconceptos específicos) y estable en el tiempo, aunque en la medida que se hace más específico es más susceptible a sufrir cambios (Marsh y Seeshing, 1997).

En el campo educativo diferentes estudios han demostrado que los individuos comparan sus logros o desempeños para autoevaluarse y generar autoconceptos de sí mismos, lo que le otorga importancia en el entorno escolar (Marsh, 1993; Marsh y Hau, 2004) debido a que las calificaciones se relacionan altamente con el autoconcepto académico (Ghazvini, 2011). Este se desarrolla en función de situaciones vividas y logros adquiridos, operando como predictor del logro académico y del aprendizaje matemático (von Keyserlingk et al., 2020).

Por ello, en lugar de evaluar el autoconcepto general (Huang, 2011) es necesario realizar mediciones de dominios específicos, especialmente en áreas históricamente complejas como matemática y estadística, en las cuales los autoconceptos de los estudiantes están fuertemente relacionados (Sproesser et al., 2016). La tradición investigativa ha enfatizado en la matemática a partir de los desarrollos del modelo ajustado de Marsh previamente mencionado, sin embargo, es necesario distinguir el autoconcepto en los dos dominios, pues si bien son materias con contenidos similares, se diferencian en función del uso de los datos y la aplicación contextual (Burrill y Biehler, 2011).

El autoconcepto matemático implica el conjunto de habilidades relacionadas con esta disciplina (Marsh y Hau 2004), el cual se forma debido a comparaciones sociales (Marsh et al., 2019) que impactan el rendimiento en matemáticas y a su vez se ve afectado recíprocamente por dicho rendimiento (Marsh, 2023). Mientras que el autoconcepto

estadístico se ha asumido como un componente de las actitudes hacia la estadística, por ello la metodología más común de medida radica en cuestionarios de actitudes (Filiz et al., 2020), siendo necesario explorar mecanismos evaluativos que garanticen la separación del autoconcepto estadístico de otros constructos, y que ayuden a recabar evidencia con el mismo grado de profundidad que actualmente se conoce para el autoconcepto matemático.

Este estudio tiene por objetivo diseñar y validar dos cuestionarios breves de dominio específico para autoconcepto en matemáticas y estadística en estudiantes de secundaria. Los cuestionarios emplean el mismo conjunto de ítems, diferenciándose únicamente en el contexto específico de cada dominio.

2. MÉTODO

Para diseñar un conjunto de ítems que respondiera a la medición del autoconcepto de dominio específico en estadística y matemáticas, fueron revisados otros cuestionarios de autoconcepto como dominio general y de autoconcepto académico como dominio particular, además de la revisión de fuentes que describen los elementos conceptuales del autoconcepto. La escogencia de instrumentos para la revisión de ítems se basó en los siguientes criterios: instrumentos cuantitativos; autoinformes; cuestionarios o subescalas breves (≤ 20 ítems); con buenas propiedades psicométricas; con aplicación en múltiples estudios.

Esto condujo a seleccionar los siguientes instrumentos o subescalas de otros instrumentos más amplios: Academic Self-Concept questionnaire (ASC, Liu et al., 2005), Self-Concept Questionnaire (SCQ-s, Saraswat, 1984), Self-Description Questionnaire (Marsh et al., 1984), Self Description Questionnaire III: subescalas Academic y Mathematics (SDQIII, Marsh & O'Neill, 1984), Autoconcepto forma 5 (AF-5, García y Musitu, 2014), Math Self-Concept Scale (Benson, 1989).

Dentro de la literatura sobresalieron elementos clave para construir una definición del autoconcepto de dominio específico, tales como la multidimensionalidad del constructo (Byrne, 1984), sus propiedades descriptivas y evaluativas (Marsh y Seeshing, 1997), la consideración de marcos de referencia personal (Marsh 1984; Marsh y O'Neill, 1984), la comparación social (Marsh y Seeshing, 1997) y la consideración del desempeño en el medio académico (García y Musitu, 2014; Liu & Wang, 2005; Marsh, 1993). A partir de esa revisión se definieron los elementos nomológicos que debían caracterizar la medición del constructo, procurando que la redacción de los ítems respondiera a seis dominios identificados como aspectos que deben evaluarse en la medida específica de autoconcepto estadístico y matemático (imagen personal, opinión docente percibida, desempeño en tareas, conocimiento, opinión de pares percibida, contexto de la clase). En este estudio la definición de autoconcepto empleada fue la siguiente: el autoconcepto de dominio específico (estadística / matemática) es la percepción individual de un estudiante sobre

su conocimiento y sus estándares de desempeño en la asignatura, formada a partir de la experiencia educativa personal, las interacciones y opiniones de pares y educadores, y la forma como el estudiante las interpreta.

El continuo del constructo (Tay y Jebb, 2018) se definió de tal forma que puntuaciones altas indican autoconcepto de dominio específico elevado, y puntuaciones bajas indican un bajo autoconcepto en el dominio. Se trata entonces de polos diferentes de un mismo constructo en lugar de dos constructos distintos dados por las puntuaciones extremas. Al tratarse de un constructo unipolar se seleccionó una escala de respuesta para los ítems basada en la fuerza de la creencia, en la cual las etiquetas verbales de las respuestas definen el grado de acuerdo / desacuerdo con el contenido de las oraciones. Dicha escala se recomienda para medir continuos conceptuales unipolares (Tay y Jebb, 2018).

Se redactaron ítems directos para evitar contaminación de puntuaciones inversas en escalas de continuo unipolar (Jebb et al., 2021; Tay y Jebb, 2018), procurando aumentar la comprensión, dada la edad y el nivel formativo de la población. Para cada dominio se redactaron tres ítems que posteriormente fueron revisados por cinco evaluadores expertos para juzgar apariencia y contenido. Su revisión se cumplió mediante grupos focales para la discusión de las propiedades de los ítems, además de ofrecer sugerencias de mejora en la redacción, como recomienda la literatura (Cabaguing, 2018).

La revisión del grupo de jueces condujo a definir un ítem por dominio para conformar un cuestionario breve unidimensional de seis reactivos. En la Tabla 1 se resume el contenido de los ítems seleccionados, así como el dominio al cual responden. La estructura lingüística de los ítems de ambas escalas fue la misma, mientras que la diferenciación entre estadística y matemática marcó la contextualización para cada campo disciplinar.

Tabla 1. Ítems diseñados para las escalas.

Ítems	Dominio
Soy bueno en estadística/matemática	Imagen personal
Para mis profesores soy un buen estudiante en estadística/ matemática	Opinión docente percibida
Realizo bien los trabajos y actividades de estadística/ matemática	Desempeño en tareas
Tengo buen nivel de conocimiento sobre estadística/ matemática	Conocimiento
Mis compañeros me consideran hábil para la estadística/ matemática	Opinión de pares percibida
En clases de estadística/matemática trabajo bastante	Contexto de la clase

Se realizó un pilotaje con 10 estudiantes que diligenciaron ambos cuestionarios con el propósito de conocer la facilidad de comprensión de las oraciones. Posteriormente se aplicó la versión final de cada escala junto con otros instrumentos que miden constructos relacionados para probar la convergencia de las puntuaciones. La aplicación fue colectiva, en instituciones educativas con autorización directiva, consentimiento de padres o acudientes (adultos responsables) y asentimiento informado de los estudiantes. La evaluación de las propiedades psicométricas se realizó en dos estudios independientes.

3. ESTUDIO 1

3.1.1 Participantes

Se evaluó a 326 estudiantes colombianos de secundaria divididos en 159 hombres (48.8%) y 167 mujeres (51.2%), con edades entre 15 y 18 años ($M = 15.62 \pm .885$). La media de edad de los chicos fue 15.66 ($\pm .86$) y la de las chicas fue 15.60 ($\pm .73$).

3.1.2 Instrumentos

Cuestionario breve de autoconcepto estadístico (CBAE). Compuesto por seis reactivos tipo Likert (1=total desacuerdo, 5=total acuerdo) de puntuación directa, que responden a un modelo unidimensional. Puntuaciones mayores indican autoconcepto estadístico más elevado.

Autoconcepto Forma 5 (AF-5). Compuesto por 30 ítems que evalúan cinco dimensiones del autoconcepto. Cada dimensión se compone de seis reactivos, en este estudio se empleó la primera dimensión (en adelante AF5-AA) que mide el autoconcepto en el escenario educativo y laboral asumiendo que ambos son contextos de trabajo (García y Musitu, 2014). Las puntuaciones indican la percepción del evaluado sobre la calidad del desempeño de su rol.

Survey of Attitudes Toward Statistics (SATS). Cuestionario de amplio uso internacional para medir actitudes hacia la estadística (Schau et al., 1995). Consta de 28 ítems tipo Likert que miden cuatro dimensiones actitudinales: *afectiva* (seis ítems: sentimientos positivos/negativos hacia la estadística), *cognitiva* (seis ítems: actitudes hacia los conocimientos cuando se aplica estadística), *valor* (nueve ítems: relevancia y utilidad de la estadística para la vida personal y profesional), *dificultad* (siete ítems: grado de dificultad percibido en la estadística).

3.1.3 Análisis de datos

Se aplicó Análisis Factorial Confirmatorio (AFC) con el programa RStudio y el paquete lavaan, empleando matrices policóricas. El ajuste matricial se calculó con KMO y prueba de esfericidad de Bartlett, mientras que la estructura factorial se probó con el método Diagonally Weighted Least Squares (WLSMV), útil para matrices policóricas con datos ordinales distribuidos de forma no normal (Wang y Cunningham, 2005), y arroja resultados menos sesgados y más precisos al estimar las cargas factoriales.

El ajuste del modelo se evaluó con chi-cuadrado (χ^2) (se esperan valores no significativos) y chi-cuadrado grados de libertad (χ^2/df ; < 5 aceptable, ≤ 2 bueno, $= 1$ muy bueno) (West et al., 2012). También se calcularon índices de ajuste como Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), Nonnormed Fit Index (NNFI), para los cuales se definen valores aceptables $>.90$ y buenos $>.95$. Adicionalmente, se calculó el Root Mean Square Error of Approximation (RMSEA) cuyos valores esperados son $<.08$ con IC90%, y el Root Mean Square Residuals (SRMR), con valores $<.08$ (Hu y Bentler, 1999; West et al., 2012).

La consistencia interna se analizó con Omega de McDonald (ω), que realiza el cálculo a partir de las cargas factoriales sin depender del número de ítems. También se calculó la varianza común explicada (ECV), que identifica la fuerza del factor general, es decir, su unidimensionalidad. La convergencia de las puntuaciones se calculó mediante correlación de Spearman (tamaño de efecto z_{Fisher} pequeño $= .10$, mediano $= .30$, grande $= .50$) con los resultados del SATS y el AF5-AA.

3.1.4. Resultados estudio 1

La matriz policórica demostró buen comportamiento ($KMO = .888$, $Bartlett \chi^2_{[15]} = 1148.5$, $p < .001$) que permitió adelantar el análisis factorial. En la Tabla 2 se resumen los estadísticos descriptivos por ítem señalando porcentajes de respuesta por cada opción de la escala. Ningún ítem acumula respuestas en los valores máximos o mínimos. Las correlaciones de cada ítem con el resto fueron altas, al igual que saturaciones y comunalidades.

Tabla 2. Análisis de ítems del CBAE: correlación, distribución de respuesta y saturaciones.

Ítem	Distribución porcentual por respuesta							Correlaciones		Cargas	
	Media	DT	1	2	3	4	5	CIT	R ²	λ	h ²
1	3.337	.949	4.6	12.3	35.6	39.9	7.7	.735	.694	.821	.675
2	3.276	.879	5.2	8.3	44.8	37.1	4.6	.732	.661	.810	.656
3	3.736	.879	2.8	5.2	23.0	53.7	15.3	.733	.658	.818	.668
4	3.377	.968	4.0	14.1	31.0	42.0	8.9	.779	.787	.860	.739
5	3.135	1.002	7.7	14.4	41.4	29.8	6.7	.741	.650	.814	.663
6	3.405	1.056	7.1	1.4	29.4	41.1	12.0	.711	.612	.784	.615

CIT = correlación ítem-test, λ = saturaciones, h² = comunalidades.

Se calcularon dos modelos factoriales, el primero integraba los seis ítems, pero mostró bajo ajuste con valores $\chi^2/gl > 5$ y RMSEA $>.08$. La revisión de ítems condujo a suprimir el cuarto reactivo por presentar elevada covarianza de los errores con los reactivos 1, 2 y 3. El resultado fue una versión de cinco ítems que mostró indicadores favorables de bondad de ajuste con χ^2/gl aceptable ($> 2 \mid \chi^2/gl \mid < 5$) al igual que el RMSEA ($<.08$), y buenos índices CFI, TLI y NNFI. Este modelo obtuvo una varianza total explicada del 61%.

Tabla 3. Estadísticas de bondad de ajuste de los modelos probados.

Modelo	χ^2	gl	χ^2/gl	CFI	TLI	NNFI	RMSEA [IC90%]	SRMR
CBAE 6 ítems	69.773	9	7.75	.985	.974	.999	.144 [.114, .177]	.039
CBAE 5 ítems	14.984	5	2.99	.996	.992	.992	.078 [.035, .126]	.021

El análisis de fiabilidad reportó valores altos del Omega total ($\omega =.92$) que sugieren elevada consistencia interna, mientras que la variación común explicada ($ECV =.86$) indica que la solución unidimensional es admisible. Para probar la convergencia se emplearon las puntuaciones del modelo de mejor ajuste (5 ítems) y se calcularon las correlaciones con las puntuaciones del AF5-AA y el SATS empleando bootstrapping para 1000 muestras. Se identificaron correlaciones fuertes entre las puntuaciones del CBAE y el AF5-AA con tamaño de efecto grande ($rho =.650, p <.001, z_{Fisher} =.77$), así como correlaciones moderadas y efecto mediano con las actitudes Afectiva ($rho =.384, p <.001, z_{Fisher} =.40$) y Cognitiva ($rho =.421, p <.001, z_{Fisher} =.44$), correlaciones bajas con la actitud Valor ($rho =.246, p <.001, z_{Fisher} =.25$) y no se identificó relación con la Dificultad.

3.1.5 Discusión estudio 1

La distribución de los porcentajes de respuesta desacredita la existencia de efecto suelo y techo, de modo que no se aprecia acumulación de respuestas en las opciones mínima y máxima por ítem. Las puntuaciones oscilan entre las opciones de respuesta 2 y 4. El comportamiento general de los ítems del CBAE fue adecuado exceptuando el número 4 “*Tengo buen nivel de conocimiento sobre estadística*”, cuya supresión condujo a la ostensible mejora del ajuste del modelo. Aunque el modelo resultante explica 61% de la varianza total, debe considerarse que el ajuste es aceptable producto del valor relativamente elevado del χ^2/gl . Pese a ello, el comportamiento general del instrumento es favorable y ofrece medidas de consistencia interna altas, además, las puntuaciones correlacionan con otros constructos de naturaleza afectiva dentro de la educación matemática como se observa en el análisis de correlación con las puntuaciones SATS y AF5-AA.

3.2 Estudio 2

3.2.1 Participantes

Se evaluó a 381 estudiantes divididos en 242 chicos (63.5%) y 139 chicas (36.5%), con un promedio de edad de 15.27 años ($\pm .99$). La media en años de los chicos fue de 15.34 (± 1.04) y en las chicas de 15.07 ($\pm .86$).

3.2.2 Instrumentos

Cuestionario breve de autoconcepto matemático (CBAM). Bajo la misma composición del CBAE, posee seis ítems de escala tipo Likert que responden a una sola dimensión. Todos los ítems son positivos, de modo que puntuaciones altas indican más autoconcepto matemático.

Escala de Metas de Rendimiento en Matemática (E-MRM). Consta de 10 ítems tipo Likert de cinco puntos (desacuerdo/acuerdo) con solución unidimensional ($\alpha = .86$). Demuestra estadísticos robustos calculados mediante AFC que señalan un nivel de ajuste aceptable del modelo unifactorial ($GFI = .90$; $CFI = .93$, $RMSEA = .09$) (Cupani, 2010).

Escala de Expectativas Rendimiento en Matemática (E-ERM). Diseñada con estructura análoga a la anterior, la componen nueve ítems tipo Likert (desacuerdo/acuerdo) retenidos en una solución unidimensional con buen nivel de consistencia interna y ajuste factorial ($\alpha = .85$, $GFI = .95$; $CFI = .96$, $RMSEA = .06$) (Cupani, 2010).

Subescala de Autoeficacia lógico-matemática (ALM). Incluida en el *Inventario de Autoeficacia para Inteligencias Múltiples Revisado (IAIM-R)* (Pérez y Medrano, 2007). La ALM corresponde a la segunda de ocho dimensiones y está constituida por seis ítems medidos en una escala de 1 (nada seguro) a 10 (totalmente seguro), los cuales evalúan la

valoración personal frente al razonamiento lógico-matemático. La consistencia interna es buena ($\alpha = .89$).

3.2.3 Análisis de datos

El análisis siguió el mismo protocolo desarrollado con el CBAE. Se realizó AFC con pruebas chi-cuadrado (χ^2), chi-cuadrado grados de libertad (χ^2/df), CFI, TLI, NNFI, RMSEA y SRMR, mientras que la consistencia interna se calculó con Omega de McDonald (ω) y con la varianza común explicada (ECV). La convergencia de las puntuaciones se calculó con correlación de Spearman con las puntuaciones del E-MRM, E-ERM y ALM.

3.2.4 Resultados estudio 2

La Tabla 4 contiene el análisis descriptivo de los ítems junto con las correlaciones y las cargas factoriales. A pesar de que todos los ítems presentaron saturaciones altas, los reactivos con menores comunalidades fueron 3 y 6, los cuales también acumularon más del 20% de respuestas en la opción máxima, sugiriendo posible efecto techo.

Tabla 4. Análisis de ítems del CBAM: correlación, distribución de respuesta y saturaciones.

Ítem	Distribución porcentual por respuesta							Correlaciones		Cargas	
	Media	DT	1	2	3	4	5	CIT	R ²	λ	h ²
1	3.349	.985	5.0	11.3	38.3	34.6	10.8	.688	.714	.836	.700
2	3.386	.941	3.2	11.8	39.4	34.6	11.0	.546	.416	.649	.421
3	3.759	1.013	4.7	5.5	21.3	46.2	22.3	.514	.365	.604	.365
4	3.318	.963	3.2	15.7	37.5	33.3	10.2	.656	.613	.778	.606
5	3.110	1.135	8.1	21.5	34.9	22.0	13.4	.646	.589	.766	.587
6	3.698	1.008	3.7	7.3	26.2	40.9	21.8	.502	.330	.573	.328

CIT = correlación ítem-test, λ = saturaciones, h² = comunalidades.

El modelo con seis ítems mostró ajuste pobre con valores χ^2/df y RMSEA elevados. La inspección de ítems sugirió eliminar el reactivo 6 por presentar la más baja comunalidad y correlación con el resto de ítems, además, aumentaba la covarianza de los errores con los ítems 1 y 3. Se probó un modelo de cinco ítems que explica 53.5% de la varianza mostrando buen ajuste en todos los índices calculados (Tabla 5). El Omega total fue de .84 y la variación común explicada fue de .83.

Tabla 5. Bondad de ajuste de los modelos probados.

Modelo	χ^2	gl	χ^2/gl	CFI	TLI	NNFI	RMSEA [IC90%]	SRMR
CBAM 6 ítems	36.384	9	4.02	.985	.975	.995	.089 [.060, .121]	.042
CBAM 5 ítems	9.871	5	1.97	.997	.994	.994	.051 [.000, .097]	.025

Para evaluar la convergencia se calculó correlación de Spearman con bootstrapping de 1000 muestras para CBAM, E-MRM, E-ERM y ALM, transformando los valores a puntuaciones Z dadas las diferencias en sus escalas de medida. Las puntuaciones del CBAM mostraron correlaciones débiles con E-ERM ($\rho = .216$, $p < .001$, $z_{Fisher} = .21$), moderadas con E-MRM ($\rho = .360$, $p < .001$, $z_{Fisher} = .37$) y fuertes con ALM ($\rho = .605$, $p < .001$, $z_{Fisher} = .70$).

3.2.5 Discusión estudio 2

El comportamiento del CBAM con cinco ítems muestra ajuste favorable. Las puntuaciones se distribuyen uniformemente entre las opciones de respuesta, de hecho, tras eliminar el sexto ítem solo se observa acumulación de más del 20% de respuestas en la opción máxima para el reactivo 3 (“*Realizo bien los trabajos y actividades de matemáticas*”). A pesar de ello, la corrección del modelo con cinco ítems responde a adecuados niveles de ajuste factorial en todos los índices calculados. Las puntuaciones han obtenido evidencia de validez concurrente al presentar correlaciones con las demás variables evaluadas.

4. DISCUSIÓN GENERAL

El autoconcepto es un constructo multidimensional dentro del cual se ha abordado con asiduidad su dimensión académica en investigación educativa y psicológica. El autoconcepto académico ha sido ampliamente identificado como una variable relacionada con el logro educativo (Chen et al., 2013; Bakadorova y Raufelder, 2020), y dentro de la educación matemática se le ha considerado como una variable que integra el dominio afectivo con implicaciones relevantes sobre el logro (Shaalvik y Shaalvik, 2006).

Aunque se entiende que autoconcepto estadístico y matemático son diferentes (González et al., 2016), en este estudio se ha propuesto un conjunto de ítems con estructura lingüística común organizada por dominios de análisis, lo que facilita emplear los ítems para analizar autoconcepto estadístico y matemático reemplazando el contexto disciplinar en las oraciones. A partir del diseño de seis reactivos, los estudios aplicados en esta investigación han conducido a diferenciar dos modelos de medida para autoconcepto estadístico y matemático, ambos con cinco ítems, de los cuales comparten cuatro.

Esta apuesta no es nueva, literatura describe instrumentos que evalúan constructos diversos basados en una estructura de ítems que comparte el contenido, siendo modificados ciertos elementos para contextualizar puntualmente la medida. En la medición de ansiedad como rasgo/estado psicológico encontramos un ejemplo con el STAI, que comparte ítems entre las dimensiones que evalúan el rasgo y el estado ansioso. Otro ejemplo lo tenemos en psicología social con el Humiliation Inventory (Hartling y Luchetta, 1999), que dispone de dos conjuntos de ítems idénticos para evaluar humillación acumulada y miedo a ser humillado, y para los cuales la variación en la estructura para diferenciar las dos dimensiones radica en la consigna de medida, pues para la humillación acumulada se le pide al evaluado considerar desde el pasado hasta el presente, mientras que la segunda dimensión emplea los mismos ítems, pero solicitando considerar una orientación hacia el futuro.

En este estudio, el diseño de ítems se basó en evaluar el autoconcepto desde su concepción multidimensional, por lo cual los reactivos abordan los dominios de imagen personal, opinión docente, desempeño en tareas, conocimiento, opinión de pares y contexto de la clase. Para el caso del autoconcepto estadístico, el modelo factorial retuvo todos los ítems exceptuando el relativo a conocimiento, mientras que en el modelo de autoconcepto matemático el ítem excluido es el relativo al contexto de la clase. Los modelos factoriales han mostrado buen ajuste, con proporciones de varianza explicada que superan el 50%. Adicionalmente, la medida de consistencia interna reporta alta fiabilidad en las puntuaciones de ambos cuestionarios, así como ajuste apropiado de unidimensionalidad.

Las puntuaciones responden favorablemente a la evidencia de validez concurrente expresando relaciones con las mediciones calculadas con instrumentos diversos. En el caso del CBAM los resultados coinciden con la literatura (Shaalvik y Shaalvik, 2006) al registrar relaciones directas con expectativas de logro, metas en matemáticas y evaluación autoinformada del razonamiento lógico-matemático. Análogamente, las puntuaciones del CBAE se asociaron con autoconcepto académico y con actitudes ante la estadística medidas con el SATS.

La medida del autoconcepto de escolares de secundaria en estadística y matemática con el CBAE y el CBAM ofrece varias ventajas para la investigación aplicada. Primero, al ser cuestionarios cortos garantizan que la evaluación goce de rapidez (Schrepp et al., 2017) y contribuyen a que la investigación aplicada sea más eficiente, aumentando la relación costo-beneficio. Adicionalmente, las mediciones breves permiten que los investigadores incluyan un mayor número de medidas en sus estudios, lo que ayuda a que puedan aplicar más instrumentos para recabar otros datos relevantes. En tercer lugar, las medidas abreviadas pueden tener efectos positivos sobre la validez al reducir el riesgo de fatiga y el aburrimiento en los informantes (Burisch, 1984). También existe evidencia de la favorabilidad en la validez de criterio de escalas breves, incluso más que en versiones similares más extensas (Thalmayer et al., 2011).

El diseño de estas escalas afronta limitaciones. Este estudio se basó en muestras no probabilísticas cuyas limitantes relacionadas con la generalización de resultados son ampliamente conocidas, siendo importante replicar estudios de validez en los que se seleccione a los participantes con muestras aleatorias, preferiblemente estratificadas por la naturaleza jerárquica de los grados académicos.

También es necesario incluir nuevos análisis que ofrezcan medidas de fiabilidad test-retest, incluso, en la evaluación del desempeño de instrumentos de autoconcepto a nivel general se ha sugerido la importancia de realizar análisis de tipo longitudinal, dado que el autoconcepto de los escolares puede sufrir cambios por los efectos de la edad (Wu et al., 2021). Por último, los cuestionarios no han incluido ítems que identifiquen el autoconcepto en relación con el aprendizaje mediado con tecnología, lo cual ha sido sugerido por algunos autores (Anastasiadou, 2011), dada su utilidad para estudiar contenidos numéricos y resolver problemas.

5. CONCLUSIONES

CBAE y CBAM parecen dos recursos solventes para la medida del autoconcepto de población escolar de secundaria en el campo de la educación matemática. Su brevedad y la simpleza de las oraciones (ítems), contribuyen al rápido entendimiento y diligenciamiento, por lo cual su uso en investigación puede facilitar el análisis del papel del autoconcepto matemático y estadístico en fenómenos diversos como ansiedad, logro académico, actitudes, entre otros.

REFERENCIAS

- Anastasiadou, S. (2011). Reliability and validity testing of a new scale for measuring attitudes toward learning statistics with technology. *Acta Didactica Napocensia*, 4(1), 1-10. <https://cutt.ly/9wErufZS>
- Bakadorova, O., & Raufelder, D. (2020). The relationship of school self-concept, goal orientations and achievement during adolescence. *Self and Identity*, 19(2), 235-249. <https://doi.org/10.1080/15298868.2019.1581082>
- Bellmore, A., & Cillessen, A. (2006). Reciprocal influences of victimization, perceived social preferences, and self-concept in adolescence. *Self and Identity*, 5, 209-229. <https://doi.org/10.1080/15298860600636647>
- Benson, J. (1989). Structural components of statistical test anxiety in adults: An exploratory model. *Journal of Experimental Education*, 57, 247-261. <https://www.jstor.org/stable/20151773>

- Burisch, M. (1984). Approaches to personality inventory construction: A comparison of merits. *American Psychologist*, *39*, 214-227. <https://doi.org/10.1037/0003-066X.39.3.214>
- Burrill, G., & Biehler, R. (2011). Fundamental statistical ideas in the school curriculum and in training teachers. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics-Challenges for teaching and teacher education* (pp. 57-69). Springer. <https://cutt.ly/VwEQYj9N>
- Byrne, B. (1984). The general/academic self-concept nomological network: A review of construct validation research. *Review of Educational Research*, *54*, 427-456. <https://doi.org/10.3102/00346543054003427>
- Cabaguino, A. (2018). Adaptation and validation of Academic Self-Concept Questionnaire (ASCQ) for College Students. *International Journal of Multidisciplinary Approach and Studies*, *5*(1), 57-66. <https://cutt.ly/bwEruF0C>
- Chen, S.-K., Yeh, Y.-C., Hwang, F.-M., & Lin, S. (2013). The relationship between academic self-concept and achievement: A multicohort–multioccasion study. *Learning and Individual Differences*, *23*, 172-178. <https://doi.org/10.1016/j.lindif.2012.07.021>
- Cupani. (2010). Validación de una nueva escala de Expectativas de Resultados y Metas de Rendimiento en Matemática. *Interdisciplinaria*, *27*(1), 111-127. <https://cutt.ly/awEQTH56>
- Filiz, M., Early, E., Thurston, A., & Miller, S. (2020). Measuring and improving university students' statistics self-concept: A systematic review. *International Journal of Educational Research Open*. <https://doi.org/100020>. 10.1016/j.ijedro.2020.100020
- García, F., & Musitu, G. (2014). *AF-5. Autoconcepto forma 5* (4ª ed.). TEA Ediciones.
- González, A., Rodríguez, Y., Faílde, J., & Carrera, M. V. (2016). Anxiety in the statistics class: Structural relations with self-concept, intrinsic value, and engagement in two samples of undergraduates. *Learning and Individual Differences*, *45*, 214-221. <https://doi.org/10.1016/j.lindif.2015.12>
- Ghazvini, S. D. (2011). Relationships between academic self-concept and academic performance in high school students. *Procedia - Social and Behavioral Sciences*, *15*, 1034-1039. <https://doi.org/10.1016/j.sbspro.2011.03.235>.
- Goble, G., Van Ooyik, J., Robertson, T., & Roberts, G. (2021). Effects on students' academic and non-academic outcomes and student participation in theatre arts: A research synthesis. *Educational Research: Theory and Practice*, *32*(3), 1-2. <http://files.eric.ed.gov/fulltext/EJ1329015.pdf>
- Hartling, L., & Luchetta, T. (1999). Humiliation: Assessing the impact of derision, degradation, and debasement. *The Journal of Primary Prevention*, *19*, 259-278. <https://doi.org/10.1023/A:1022622422521>

- Huang, C. (2011). Self-concept and academic achievement: A meta-analysis of longitudinal relations. *Journal of School Psychology, 49*(5), 505-528. <https://doi.org/10.1016/j.jsp.2011.07.001>
- Hu, L., & Bentler, P. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1-55. 10.1080/10705519909540118
- Jebb, A., Vincent, Ng., & Tay, L. (2021). A review of key likert scale development advances: 1995-2019. *Frontiers in Psychology, 12*. <https://doi.org/10.3389/fpsyg.2021.637547>
- Liu, W., & Wang, C. (2005). Academic self-concept: A cross-sectional study of grade and gender differences in a Singapore Secondary School. *Asia Pacific Education Review, 6*(1), 20-27.
- Liu, W., Wang, C., & Parkins, E. (2005). A longitudinal study of students' academic self-concept in a streamed setting: The Singapore context. *British Journal of Educational Psychology, 75*(4), 567-586. 10.1348/000709905x42239
- Marsh, H. (1993). Academic Self-Concept: Theory, measurement and research. En J. Suls, *Psychological Perspectives on the Self: the Self in Social Perspective* (Vol. 4) (pp. 59-98). Taylor & Francis. <https://cutt.ly/4wEQRT00>
- Marsh, H. (2023). Extending the reciprocal effects model of math self-concept and achievement: Long-term implications for end-of-high-school, age-26 outcomes, and long-term expectations. *Journal of Educational Psychology, 115*(2), 193-211. <https://doi.org/10.1037/edu0000750>
- Marsh, H., & Hau, K. (2004). Explaining paradoxical relations between academic self-concepts and achievements: Cross-cultural generalizability of the internal/external frame of reference predictions across 26 countries. *Journal of Educational Psychology, 96*(1), 56-67. <https://doi.org/10.1037/0022-0663.96.1.56>
- Marsh, H. W., & O'Neill, R. (1984). Self-Description Questionnaire III: The construct validity of multidimensional self-concept ratings by late adolescents. *Journal of Educational Measurement, 21*(2), 153-174. <https://doi.org/10.1111/j.1745-3984.1984.tb00227.x>
- Marsh, H. W., Barnes, J., Cairns, L., & Tidman, M. (1984). Self-Description Questionnaire: Age and sex effects in the structure and level of self-concept for preadolescent children. *Journal of Educational Psychology, 76*(5), 940-956. <https://doi.org/10.1037/0022-0663.76.5.940>
- Marsh, H., Pekrun, R., Parker, P., Murayama, K., Guo, J., Dicke, T., & Arens, A. (2019). The murky distinction between self-concept and self-efficacy: Beware of lurking jingle-jangle fallacies. *Journal of Educational Psychology, 111*(2), 331. <https://doi.org/10.1037/edu0000281>

- Marsh, H., & Seeshing, A. (1997). Causal effects of academic self-concept on academic achievement: structural equation of longitudinal data. *Journal of Educational Psychology*, 89(1), 41-54. <https://psycnet.apa.org/doi/10.1037/0022-0663.89.1.41>
- Möller, J., Zitzmann, S., Helm, F., Machts, N., & Wolff, F. (2020). Uma meta-análise das relações entre realização e autoconceito. *Revisão de Pesquisa Educacional*. <https://doi.org/10.3102/0034654320919354>
- Niepel, C., Brunner, M., & Preckel, F. (2014). Achievement goals, academic self-concept, and school grades in mathematics: Longitudinal reciprocal relations in above average ability secondary school students. *Contemporary Educational Psychology*, 39(4), 301-313. <https://doi.org/10.1016/j.cedpsych.2014.07.002>
- Pérez, E., & Medrano, L. (2007). Inventario de Autoeficacia para Inteligencias Múltiples Revisado: Un estudio de validez de criterio. *Avances en Medición*, 5, 105-114. <https://cutt.ly/xwEQEYwj>
- Saraswat, R. (1984). Self-Concept Questionnaire (SCQ-s). *National Psychological Cooperation, India*.
- Schau, C., Stevens, J., Dauphine, T., & Del Vecchio, A. (1995). The development and validation of the survey of attitudes toward statistics. *Educational and Psychological Measurement*, 55(5), 868-875. <https://doi.org/10.1177/0013164495055005022>
- Shaalvik, E.M., & Shaalvik, (2006). Self-concepto and self-efficacy in mathematics: Relations with mathematics motivation and achievement. En A. Prescott (Ed.), *The concept of self in education, family, and sports* (pp. 51-75). Nova Science Publishers. <https://cutt.ly/2wEQWRub>
- Shavelson, R., Hubner, J., & Stanton, G. (1976). Self-concept: Validation of construct interpretations. *Review of Educational Research*, 46(3), 407-441. <https://doi.org/10.3102/00346543046003407>
- Scherrer, V., & Preckel, F. (2019). Development of motivational variables and self-esteem during the school career: A meta-analysis of longitudinal studies. *Review of educational research*, 89(2), 211-258. <https://doi.org/10.3102/0034654318819127>
- Schrepp, M., Hinderks, A., & Thomaschewski, J. (2017). Design and evaluation of a Short Version of the User Experience Questionnaire (UEQ-S). *International Journal of Interactive Multimedia and Artificial Intelligence*, 4(6), 103-108. <https://doi.org/10.9781/ijimai.2017.09.001>
- Sproesser, U., Engel, J., & Kuntze, S. (2016). Fostering self-concept and interest for statistics through specific learning environments. *Statistics Education Research Journal*, 15(1), 28-54. <https://doi.org/10.52041/serj.v15i1.256>

- Tay, L., & Jebb, A. (2018). Establishing construct continua in construct validation: the process of continuum specification. *Advances in Methods and Practices in Psychological Science*, 1-14. <https://doi.org/10.1177/2515245918775707>
- Thalmayer, A., Saucier, G., & Eigenhuis, A. (2011). Comparative validity of Brief to Medium-Length Big Five and Big Six Personality Questionnaires. *Psychological Assessment*, 23(4), 995–1009. <https://doi.org/10.1037/a0024165>
- von Keyserlingk, L., Becker, M., Jansen, M., & Maaz, K. (2020). Leaving the pond-Choosing an ocean: Effects of student composition on STEM major choices at university. *Journal of Educational Psychology*, 112(4), 751–764. <https://doi.org/10.1037/edu0000378>
- Wang, W., & Cunningham, E. (2005). Comparison of alternative estimation methods in Confirmatory Factor Analyses of the General Health Questionnaire. *Psychological Reports*, 97(1), 3-10. <https://doi.org/10.2466/pr0.97.1.3-10>
- West, S., Taylor, A., & Wu, W. (2012). Model fit and model selection in structural equation modeling. In R. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 209–231). The Guilford Press. <https://cutt.ly/mwEQQuPb>
- Wu, H., Guo, Y., Yang, Y., Zhao, L., & Guo, C. (2021). A Meta-analysis of the longitudinal relationship between academic self-concept and academic achievement. *Educational Psychology Review*, <https://doi.org/10.1007/s10648-021-09600-1>

Recibido: 27 de abril de 2024

Aceptado: 10 de agosto de 2024