Evaluación psicométrica de la Escala de Autoeficacia en Conductas Académicas en universitarios: una propuesta de optimización del instrumento

Psychometric evaluation of the Self-Efficacy Scale for Academic Behaviors in university students: a proposal for instrument optimization

Autores

M.C. Edgar Francisco Ordoñez Bencomo [1]
Dra. Susana Ivonne Aguirre Vásquez [1]
Dr. Jesús Enrique Peinado Pérez [1]
Dr. José René Blanco Ornelas [1]

[1] Universidad Autónoma de Chihuahua (México)

Autor de correspondencia:
Dra. Susana Ivonne Aguirre Vásquez
siaguirre@uach.mx

Recibido: 21-04-26
Aceptado: 30-05-26

Cómo citar en APA

Ordoñez Bencomo, E. F., Aguirre Vásquez, S. I., Peinado Pérez, J. E., & Blanco Ornelas, J. R. (2027). Evaluación psicométrica de la Escala de Autoeficacia en Conductas Académicas en universitarios: una propuesta de optimización del instrumento. *Retos, 86*, 122-132. https://doi.org/10.47197/retos.v86.119305

Resumen

Introducción: La autoeficacia es la creencia que una persona tiene acerca de su capacidad para realizar y concluir con éxito una tarea específica. Esta percepción favorece conductas autorreguladoras que contribuyen a mejorar el desempeño académico y el logro de metas educativas.
Objetivo: Analizar las propiedades psicométricas de la Escala de Autoeficacia en Conductas Académicas (EACA).
Método: Participaron 913 estudiantes (509 mujeres, 404 hombres), con edades entre 18 y 25 años, pertenecientes a la Universidad Autónoma de Chihuahua. Mediante análisis factorial confirmatorio se compararon dos modelos de medida Modelo 1 (M1) y Modelo 2 (M2). Los análisis se realizaron utilizando el software AMOS 21.
Resultados: El M1 presentó un ajuste aceptable (χ^2/gl = 5.231; GFI = .933; RMSEA = .068; CFI = .950; TLI=.939; SRMR=.037) y explicó el 69.36% de la varianza. No obstante, algunos ítems mostraron cargas factoriales débiles. El M2 mostró una mejora significativa en los índices de ajuste (χ^2/gl=2.623; GFI = .981; RMSEA = .042; CFI = .989; TLI= .984; SRMR=.0209) y explicó el 76.45 % de varianza.
Discusión: Aunque el M2 mejora el ajuste de la EACA y mantiene una fiabilidad adecuada, las altas intercorrelaciones entre los factores sugiere un posible solapamiento conceptual, lo que limita la validez discriminante del instrumento.
Conclusiones: El Modelo 2 muestra una mejora en los índices de ajuste y en la definición de los ítems, por lo que resulta preliminarmente adecuado para su uso en estudiantes de ciencias del deporte. No obstante, se requieren análisis adicionales para mejorar su estructura factorial y fortalecer su validez.

Palabras clave

Análisis factorial confirmatorio; autoeficacia académica; evaluación educativa; instrumentos de medición; propiedades psicométricas.

Abstract

Introduction: Self-efficacy is the belief that a person has about their ability to perform and successfully complete a specific task. This perception promotes self-regulatory behaviors that contribute to improving academic performance and the achievement of educational goals.
Objective: To analyze the psychometric properties of the Academic Behavior Self-Efficacy Scale (EACA).
Method: A total of 913 students (509 women, 404 men) aged between 18 and 25 from the Autonomous University of Chihuahua participated in the study. Two measurement models were compared using confirmatory factor analysis Model 1 (M1) and Model 2 (M2). The analyses were performed using AMOS 21 software.
Results: M1 showed an acceptable fit (χ^2/df= 5.231; GFI = .933; RMSEA = .068; CFI = .950; TLI=.939; SRMR=.037) and explained 69.36% of the variance. However, some items showed weak factor loadings. M2 showed a significant improvement in fit indices (χ^2/df =2.623; GFI = .981; RMSEA = .042; CFI = .989; TLI= .984; SRMR=.0209) and explained 76.45% of the variance.
Discussion: Although M2 improves the fit of the EACA while maintaining adequate reliability, the high intercorrelations among factors suggest possible conceptual overlaps, which limits the discriminant validity of the instrument.
Conclusions: Model 2 shows improved fit indices and better item definition, making it preliminarily suitable for use with sports science students. However, further analyses are required to improve its factor structure and strengthen its validity.

Keywords

Academic self-efficacy; confirmatory factor analysis; educational assessment; measurement instruments; psychometric properties.

Introducción

Las instituciones de educación superior muestran un interés creciente por promover aquellos factores que contribuyen al éxito en la culminación de una carrera universitaria. Durante este proceso, los estudiantes enfrentan el desafío de gestionar su tiempo de manera autónoma, desarrollar el pensamiento crítico y fortalecer las competencias propias de un profesional en formación. Para lograr un desempeño académico exitoso durante su paso por la institución de educación superior, resulta de suma importancia que los estudiantes tengan la convicción de que poseen las competencias y habilidades suficientes para afrontar y superar las exigencias de la formación universitaria (Borzone, 2017).

La autoeficacia es la creencia que tienen las personas sobre su capacidad para ejecutar una tarea concreta y concluirla favorablemente (Bandura, 1977). Desarrollada en el marco de la teoría del aprendizaje social, esta percepción influye en la forma en que se enfrentan los desafíos; cuando una persona confía en su capacidad para realizar una tarea, incluso si esta es compleja, es más probable que la aborde en lugar de evitarla. En el contexto académico, esta creencia favorece la formación de conductas autorreguladoras que contribuyen a mejorar el desempeño de los estudiantes. De hecho, la autorregulación y la autoeficacia están ampliamente relacionadas y permiten predecir el logro de metas académicas de los estudiantes (Rosales-Ronquillo & Hernández-Jácquez, 2020; Yupanqui-Lorenzo et al., 2021). Además, la autoeficacia se correlaciona de forma significativa con el compromiso académico, y los alumnos con niveles altos muestran mejor rendimiento (Casanova et al., 2024) y mejores calificaciones (Sánchez-Anguita y Pulido, 2022). Estas relaciones adquieren una dimensión particular en las ciencias de la actividad física y el deporte, donde la práctica física forma parte de la vida cotidiana y de la formación del estudiantado. La literatura sugiere que la actividad física no solo se vincula con beneficios físicos, sino también con creencias de competencia en el plano académico en universitarios, mayores niveles de ejercicio se han asociado con mayor autoeficacia (Wang et al., 2022), y estudios recientes reportan una relación positiva, aunque de magnitud pequeña, entre el nivel de actividad física y la autoeficacia para el aprendizaje (Adigue y Adigue, 2025).

Dada la importancia que tiene la autoeficacia en el rendimiento académico, es importante contar con mejores instrumentos que midan este constructo, que sean válidos y confiables y que permitan evaluarlo con más precisión. Para ello es imprescindible que las escalas utilizadas cuenten con propiedades psicométricas sólidas. La validez de constructo es clave en este proceso, ya que indica en qué medida los indicadores realmente reflejan el concepto que se busca medir. La validez discriminante y convergente suelen medirse juntas. La validez convergente se refiere a qué tan bien los indicadores de un mismo constructo están relacionados entre sí, mientras que la validez discriminante asegura que un constructo es realmente diferente de los demás dentro del modelo. En otras palabras, se espera que las mediciones de un mismo factor tengan una fuerte conexión entre ellas, sin mezclarse demasiado con otros constructos. Para comprobar esto, una buena práctica es calcular la varianza media extraída (AVE), que debería ser mayor a .50 y el Alfa de Cronbach, el cual es una medida de consistencia interna cuyos valores por encima de .70 indican una buena confiabilidad. Sin embargo, este coeficiente por sí solo no es suficiente para demostrar validez convergente (Bagozzi & Yi, 2012).

El realizar este tipo de análisis permite garantizar que los ítems realmente miden lo que se pretende medir. Además, al llevar a cabo las validaciones de los instrumentos de una forma continua, permite adaptar los instrumentos a diferentes poblaciones y contextos (Blanco et al., 2016). Se ha demostrado que la revisión de la validez convergente y discriminante, junto con la confiabilidad, es esencial para optimizar las herramientas de evaluación (Dominguez-Lara, 2016).

Diversas Investigaciones han desarrollado o adaptado escalas ya existentes, tomando en cuenta el análisis de validez convergente y discriminante de los ítems y los factores de los instrumentos utilizados para medir la autoeficacia percibida. En este sentido, García-Méndez y Rivera-Ledesma (2021) desarrollaron la Escala Autoeficacia en la Vida Académica, la cual está formada por 102 ítems distribuidos en 14 factores (logros profesionales, liderazgo, habilidades de investigación, responsabilidad académica, concentración y comprensión académicas, creatividad e innovación, socialización y compañerismo, exposición frente a grupo, organización académica, control de impulsos, planeación académica, aprendizaje fluido, trabajo bajo presión, asimilación de información) con una confiabilidad Alfa de Cronbach de .77 a .90 para los factores y .89 para la estructura general.

Por otra parte, Robles (2020) desarrolló la Escala de Autoeficacia Académica (ESAA) con 28 ítems y dos factores expectativa de la situación y expectativas personales. Este último se subdivide a su vez en expectativas de resultados y expectativas de la autoeficacia percibida. Se reporta el Alfa de Cronbach en .831 y .923 para los dos factores y .812 y .902 para las subdivisiones respectivamente, arrojando una autoeficacia general de .877, lo cual indica una consistencia interna alta.

Así mismo, el estudio de García-Álvarez et al. (2022) en el cual se analizó la Escala de Autoeficacia Percibida Específica de Situaciones Académicas (EAPESA), la cual está constituida por 7 ítems y con una estructura unifactorial, reportó una fiabilidad de .88 para el alfa de Cronbach.

En México, así como en diversos contextos latinoamericanos, la Escala de Autoeficacia en Conductas Académicas (EACA) desarrollada por Blanco et al. (2011) ha sido empleada en diversas investigaciones para medir la autoeficacia en estudiantes universitarios. Entre los estudios que han utilizado este instrumento se encuentran los realizados por Ornelas et al. (2013), Gutiérrez-García y Landeros-Velázquez (2018) y Ávalos et al. (2018), entre otros.

De igual manera, se han llevado a cabo distintos análisis psicométricos del instrumento. Entre ellos destacan los estudios de Rangel (2015) y Blanco et al. (2016). En conjunto, estos y otros autores han reportado que la escala presenta propiedades psicométricas adecuadas, considerándose un instrumento válido y confiable, además de ser de fácil aplicación debido a su formato breve.

Para el presente estudio, se decidió utilizar la versión de la EACA propuesta por Rangel (2015), quien realizó un análisis mediante modelos de ecuaciones estructurales, proponiendo una estructura del instrumento compuesta por 16 ítems organizados en cuatro factores: comprensión, comunicación, atención y excelencia. Los resultados obtenidos en dicho estudio reportaron índices de ajuste adecuados y niveles satisfactorios de consistencia interna, lo que respalda la pertinencia y solidez de esta estructura factorial.

A pesar de estos avances, resulta necesario seguir evaluando la estructura factorial y las propiedades psicométricas de la escala en distintas muestras universitarias. Aunque la EACA ha sido validada en estudios previos (Blanco et al., 2016; Blanco et al., 2011; Rangel, 2015), se han identificado inconsistencias en su estructura factorial, como saturaciones débiles en algunos ítems e incertidumbre sobre el número óptimo de factores, lo que constituye el problema psicométrico central que justifica continuar evaluando sus propiedades en otras poblaciones. Por lo tanto, el objetivo del presente estudio fue analizar las propiedades psicométricas de la Escala de Autoeficacia en Conductas Académicas (EACA) en estudiantes universitarios mexicanos, con el propósito de confirmar su estructura factorial y aportar evidencia empírica sobre su validez y confiabilidad.

Método

Se empleó un estudio instrumental (Montero & León, 2007), cuyo propósito es analizar las propiedades psicométricas de un instrumento de medición. Este tipo de estudios se orienta al desarrollo, adaptación y validación de pruebas o escalas, así como a la evaluación de fiabilidad y validez.

Se contrastaron dos modelos de medición para establecer cuál ofrecía la mejor calidad de ajuste estadístico y permitía una comprensión más clara de los mecanismos que conforman la autoeficacia académica. Esta estrategia metodológica proporcionó una revisión minuciosa de la consistencia del instrumento, obteniendo como resultado una propuesta de mejora de la escala para su uso en el ámbito educativo en universitarios.

Participantes

En este estudio participaron 913 estudiantes universitarios, de la Facultad de Ciencias de la Cultura Física (FCCF) de la Universidad Autónoma de Chihuahua (UACH), pertenecientes a las licenciaturas en Educación Física y Motricidad Humana. Del total de la muestra, el 55.8% correspondió a mujeres (n = 509) y el 44.2% a hombres (n = 404), distribuidos de manera aproximadamente proporcional entre los distintos semestres. Las edades oscilaron entre los 18 y los 25 años, con una media de 20.53 años (DE = 1.92). Los participantes fueron seleccionados mediante un muestreo por conveniencia, garantizando la representación de los diferentes semestres de ambas licenciaturas. El tamaño muestral se determinó

con base en los criterios metodológicos propuestos por Ruiz et al. (2010) para la implementación de modelos de ecuaciones estructurales.

Como criterios de inclusión, los estudiantes debían estar inscritos en la FCCF de la UACH y aceptar participar de manera voluntaria en la investigación, para lo cual firmaron el consentimiento informado, conforme a los lineamientos propuestos por Carracedo et al. (2017). Se excluyeron del análisis aquellos cuestionarios que presentaron respuestas incompletas.

El estudio se desarrolló en apego a lo establecido en el Reglamento de la Ley General de Salud en Materia de Investigación para la Salud en México (Secretaría de Salud, 2024), la cual, en su artículo 17, clasifica como investigación sin riesgo a los estudios que emplean métodos retrospectivos o documentales, sin alterar deliberadamente variables fisiológicas, psicológicas o sociales en los participantes. Este tipo de investigación incluye la aplicación de cuestionarios o entrevistas, siempre que se garantice el anonimato de los participantes y no se aborden aspectos sensibles de su comportamiento. Así mismo, el proyecto fue aprobado por la Secretaría de Investigación y Posgrado de la Facultad de Ciencias de la Cultura Física de la Universidad Autónoma de Chihuahua con folio de registro: 11112024-033.

Procedimiento

Se invitó a participar a estudiantes de las licenciaturas de Educación Física y Motricidad Humana de la Facultad de Ciencias de la Cultura Física de la UACH. Aquellos que aceptaron participar firmaron previamente el consentimiento informado, en el cual se les indicó que su participación era voluntaria y que podían retirarse del estudio en cualquier momento. Posteriormente, los instrumentos se aplicaron en formato digital mediante computadoras equipadas con el sistema de evaluación estandarizada, en sesiones realizadas en centros de cómputo de las instalaciones de la FCCF con una duración aproximada de 25 minutos.

Al iniciar cada una de las sesiones, se comunicó de forma verbal el objetivo y la importancia del estudio, así como las instrucciones para acceder y responder el cuestionario. Se reiteró a los participantes que podían retirarse si no deseaban continuar con la evaluación. Así mismo, se enfatizó la importancia de responder con sinceridad y se garantizó la confidencialidad y el anonimato de la información proporcionada.

Al finalizar la aplicación, se agradeció a los estudiantes su participación. Posteriormente, los datos obtenidos fueron procesados mediante el software denominado generador de resultados del editor de esca-las, versión 2.0 (Blanco et al., 2013).

Instrumento

El instrumento utilizado para este estudio corresponde a la Escala de Autoeficacia en Conductas Académicas (EACA), propuesta por Rangel (2015). Se trata de un instrumento de autorreporte conformado por 16 ítems distribuidos en cuatro factores que representan dimensiones diferenciadas, pero complementarias, de la autoeficacia en el contexto académico: Comprensión: (3 ítems) se refiere a la habilidad para entender los contenidos académicos y aplicarlos en contextos prácticos. Comunicación: (4 ítems) evalúa la capacidad para expresar ideas con claridad, realizar comentarios relevantes, participar en discusiones con los docentes y sentirse cómodo al hablar en público, ya sea ante un grupo reducido o numeroso. Atención (5 ítems) mide la habilidad para atender las explicaciones del profesorado, formular preguntas a los compañeros y contribuir activamente en las discusiones grupales. Excelencia (4 ítems) examina el nivel de compromiso con la realización de tareas, la preparación para los exámenes, la entrega puntual de trabajos, la asistencia a clases y la puntualidad. El instrumento evalúa el grado de confianza que los estudiantes perciben tener para ejecutar conductas académicas específicas, como organizar su tiempo de estudio, participar en clase, estudiar de manera autónoma y cumplir con las tareas escolares.

Cada ítem se responde en tres escenarios distintos, lo que permite obtener tres indicadores de autoeficacia: a) la autoeficacia percibida actualmente, derivada de las respuestas al escenario actual; b) la autoeficacia deseada, derivada de las respuestas al escenario ideal; y c) la autoeficacia alcanzable en el futuro, derivada de las respuestas al escenario de cambio. Las respuestas se registran en una escala tipo Likert de 0 a 10 (0 = nada capaz; 10 = totalmente capaz). Se eligió este formato porque los estudiantes universitarios en México están cultural y académicamente familiarizados con este sistema de calificación, lo que facilita la comprensión de los reactivos y favorece una mayor consistencia en las respuestas.

En estudios previos, la EACA ha mostrado estructuras diferentes; en su versión inicial, Blanco et al. (2011) encontraron una estructura de tres factores (Comunicación, Atención y Excelencia) con 13 ítems, un ajuste óptimo en el análisis factorial confirmatorio (GFI = .973; RMSEA = .043) y coeficientes de fiabilidad entre .78 y .83. Posteriormente, Blanco et al. (2016) replicaron una estructura de cuatro factores (Comprensión, Comunicación, Atención y Excelencia) con 12 ítems, con un buen ajuste y omegas entre .65 y .85.

Análisis de datos

La evaluación psicométrica del instrumento comenzó con el análisis descriptivo de cada ítem, calculando la media, la desviación típica y los índices de asimetría y curtosis, con el fin de examinar la distribución de las respuestas. Se consideraron aceptables valores de asimetría dentro del rango ±2 y de curtosis dentro del rango ±3 (Byrne, 2016). La normalidad multivariada se evaluó mediante el coeficiente de Mardia. Posteriormente, para el Modelo 2, la eliminación de ítems se basó en tres criterios: (a) cargas factoriales bajas, (b) índices de modificación que indicaran un ajuste inadecuado a la estructura factorial y (c) el análisis de su contenido, considerando la redundancia con otros reactivos del mismo factor o el planteamiento de situaciones específicas distintas de las del resto de los ítems del factor.

El análisis comparó dos modelos de medición: el Modelo 1 (M1), propuesto por Rangel (2015), que seguía la estructura original de cuatro factores según la disposición inicial de los reactivos y el Modelo 2 (M2), que conservaba esta misma organización factorial, pero eliminaba selectivamente los reactivos que no alcanzaron un adecuado ajuste psicométrico en el M1, y en el que, en su caso, se especificaron correlaciones entre errores sugeridas por los índices de modificación y justificadas por la similitud de contenido de los ítems.

Se realizaron análisis factoriales confirmatorios por medio del programa AMOS 21. Las varianzas de los términos de error se definieron como parámetros no restringidos y, para cada constructo latente (factor), se normalizó una de las estimaciones de la relación estructural asignándole un valor unitario, con el propósito de equiparar su escala con la de las variables manifiestas (ítems). Como técnica de estimación se optó por el método de máxima verosimilitud (ML). Aunque los datos no cumplieron el supuesto de normalidad multivariada (coeficiente de Mardia), se mantuvo este método debido al tamaño de la muestra y a que los índices univariados de asimetría y curtosis se encontraron dentro de los límites en los que ML se considera robusto (Kline, 2016). Siguiendo la recomendación de Thompson (2004), según la cual no solo debe valorarse el ajuste del modelo teórico, sino también contrastar los indicadores de bondad de ajuste con los de modelos alternativos, se compararon dos modelos de medición: el M1 y el M2.

El ajuste del modelo se evaluó mediante múltiples indicadores: (1) índices de ajuste absoluto, como el estadístico Chi-cuadrado (χ^2), el índice de bondad de ajuste (GFI), el índice ajustado de bondad de ajuste (AGFI), la raíz media cuadrática residual estandarizada (SRMR) y el error cuadrático medio de aproximación (RMSEA); (2) índices de ajuste incremental, que incluyeron el índice Tucker-Lewis (TLI) y el índice de ajuste comparativo (CFI); y (3) índices de parsimonia, representados por la razón chi-cuadrado/grados de libertad (χ^2/gl) y el criterio de información de Akaike (AIC), en concordancia con las recomendaciones de Byrne (2016) y Gelabert et al. (2011).

La validez convergente se evaluó mediante la significancia y la magnitud de las cargas factoriales de los ítems en cada factor y mediante la varianza media extraída (AVE), considerando adecuados valores superiores a .50 (Fornell y Larcker, 1981). La validez discriminante se evaluó con el criterio de Fornell y Larcker (1981) según el cual la raíz cuadrada del AVE de cada factor debe ser mayor que sus correlaciones con los demás factores. Finalmente, la fiabilidad de todas las dimensiones de los modelos se estimó con el Coeficiente Alpha de Cronbach y el Coeficiente Omega (Bentler, 2017; Deng y Chan, 2017; Viladrich et al., 2017).

Resultados

Análisis descriptivos

Los análisis descriptivos de los 16 ítems del cuestionario revelaron que las respuestas a todos los ítems presentan puntuaciones medias que varían entre 7.57 y 8.75. En cuanto a la desviación estándar, los

valores son superiores a 1.47 en todos los casos, considerando un rango de respuesta de 0 a 10. Los valores de asimetría se ubicaron dentro del rango ± 2, y la mayoría de los valores de curtosis se encontraron dentro del rango ± 3, lo que sugiere que las variables se aproximan razonablemente a una distribución normal. Estos resultados indican normalidad univariada y respaldan la viabilidad de aplicar un análisis factorial confirmatorio (Byrne, 2016). Sin embargo, el coeficiente de Mardia (356.07) evidenció la ausencia de normalidad multivariada (Rodríguez y Ruiz, 2008).

Análisis factorial confirmatorio

Los resultados globales del análisis factorial confirmatorio (GFI= .933; RMSEA= .068; CFI= .950) para el Modelo 1 (M1) indican que el modelo de medición es aceptable (Tabla 1).

Tabla 1. Índices absolutos, incrementales y de parsimonia para los modelos generados.

	Índices absolutos				Índices incrementales			Índices de parsimonia	
Modelo	χ^2	GFI	RMSEA	SRMR	AGFI	TLI	CFI	χ^2/gl	AIC
M1	512.643*	.933	.068	.0370	.906	.939	.950	5.231	588.643
M2	97.053*	.981	.042	.0209	.967	.984	.989	2.623	155.053

Nota: * p < .05; GFI = índice de bondad de ajuste; RMSEA = error cuadrático medio de aproximación; SRMR = raíz media cuadrática residual estandarizada; AGFI = índice corregido de la bondad de ajuste; TLI = índice de Tucker-Lewis; CFI = índice de ajuste comparativo; χ^2/gl= razón de chi-cuadrado sobre los grados de libertad; AIC = criterio de información de Akaike.

El modelo M1, compuesto por cuatro factores, explicó el 69.36% de la varianza total, calculada a partir de los autovalores de la matriz de correlaciones, lo que respalda una estructura factorial adecuada. Sin embargo, tres de los 16 ítems (02, 01 y 16) presentan saturaciones inferiores a .70 en su factor correspondiente, lo que podría afectar la solidez del modelo. Además, se observan intercorrelaciones altas entre algunos factores, particularmente entre el Factor 1 (Comprensión) y el Factor 3 (Atención) (r = .85), entre el Factor 1(Comprensión) y el Factor 4 (Excelencia) (r = .88), y entre el Factor 3 (Atención) y el Factor 4 (Excelencia) (r = .91). Este último valor sugiere una posible falta de validez discriminante entre los factores 3 y 4, ya que su alta correlación indica que podrían estar midiendo constructos muy similares. Por lo tanto, aunque el modelo presenta una estructura clara, la diferenciación conceptual entre algunos de sus factores podría no ser completamente adecuada.

Los resultados generales del análisis factorial confirmatorio para el segundo modelo evaluado, el Modelo 2 (M2), muestran un ajuste óptimo (GFI = .981; RMSEA = .042; CFI = .989). Este modelo es una versión revisada del Modelo 1, en la que se eliminaron los ítems 09, 02, 03, 01 y 11, debido a que no fueron explicados de manera adecuada por el modelo original o, según los índices de modificación, no se ajustaban correctamente a la estructura factorial. Desde el punto de vista del contenido, los ítems 02 y 03 presentan similitud con otros reactivos de atención que se conservaron (08 y 12), los ítems 01 y 11 comparten contenido entre sí (cumplimiento y entrega de tareas) y el ítem 09 alude a una situación específica de desacuerdo con los profesores, distinta del resto de los ítems de Comunicación. Asimismo, con base en los índices de modificación, se especificó una correlación entre los errores de los ítems 08 y 12, justificada por la similitud de su contenido (ambos se refieren a prestar atención en clase). Los resultados indican que el Modelo 2 ofrece una mejor representación de los datos en comparación con el Modelo 1.

En el Modelo 2, se mantienen los cuatro factores, los cuales explican en conjunto aproximadamente el 76.45% de la varianza total, lo que indica una mayor capacidad explicativa que el M1. Ninguno de los 11 ítems restantes muestra saturaciones inferiores a .74 en su factor correspondiente, lo que indica una mejor definición de las dimensiones. No obstante, el factor Excelencia quedó conformado por solo dos ítems (10 y 16). Aunque las intercorrelaciones entre los factores siguen siendo elevadas, se logró reducir la correlación entre Factor 1 (Comprensión) y Factor 4 (Excelencia) (r = .88 a r = .82) y entre Factor 3 (Atención) y Factor 4 (Excelencia) (r = .91 a r = .84), lo que representa una mejora respecto al Modelo 1. Sin embargo, la correlación entre Comprensión y Atención (F1–F3) aumentó ligeramente (de r = .85 a r = .87), lo que sugiere que estos dos factores podrían estar midiendo aspectos similares. Como se detalla en el apartado de validez discriminante, el criterio de Fornell y Larcker (1981) tampoco se cumplió en los pares F1–F4 y F3–F4, por lo que la diferenciación entre factores sigue siendo limitada.

Tabla 2. Cargas factoriales de los ítems en los modelos M1 y M2 después del análisis factorial confirmatorio de la escala de autoeficacia en conductas académicas.

	M1				M2			
	F1	F2	F3	F4	F1	F2	F3	F4
Ítem	Cargas Factoriales							
13 Identificar las ideas principales de un texto.	.77				.77			
14 Escribir de manera coherente y organizada.	.74				.74			
15 Relacionar los conceptos nuevos que estoy estudiando con otros que ya conozco.	.82				.81			
04 Expresar mis ideas con claridad.		.81				.83		
05 Hacer comentarios y aportaciones pertinentes.		.77				.76		
06 Hablar o expresarme enfrente de una clase o grupo de gente.		.76				.77		
09 Entablar un diálogo con mis profesores en caso de desacuerdo con ellos.		.73				-		
02 Escuchar con atención cuando el profesor aclara una duda a un compañero.			.69				-	
03 Escuchar con atención las preguntas y aportaciones de mis compañeros.			.74				-	
07 Poner atención cuando los profesores dan la clase.			.82				.78	
08 Poner atención cuando un compañero expone en clase.			.74				.78	
12 Escuchar con atención las preguntas y comentarios de mis profesores.			.83				.86	
01 Cumplir con las tareas que se me asignan.				.65				-
10 Preparar mis exámenes apoyándome en los apuntes de clase, el texto del curso y lecturas adicionales.				.78				.82
11 Entregar puntualmente los trabajos que se me encargan.				.71				-
16 Dedicar más horas al estudio de las que había planificado; cuando se aproxima un examen.				.69				.75
Correlaciones Factoriales								
F1	-				-			
F2	.78	-			.74	-		
F3	.85	.70	-		.87	.65	-	
F4	.88	.72	.91	-	.82	.65	.84	-

Nota: F1= Comprensión, F2 = Comunicación, F3 = Atención, F4= Excelencia.

Fiabilidad y validez

La fiabilidad de la escala completa fue adecuada en ambos modelos, en el M1 se obtuvo Ω = .955 y α = .933, y en el M2, Ω = .948 y α = .912. Así mismo, los coeficientes de fiabilidad por factor fueron adecuados en ambos modelos (Tabla 3). En el M1, el omega (Ω) osciló entre .801 y .876 y el alfa (α) entre .792 y .874, mientras que en el M2 el omega (Ω) osciló entre .763 y .849 y el alfa (α) entre .763 y .834. En general, el M2 mostró valores ligeramente menores que el M1, siendo el factor Excelencia (F4) el de menor fiabilidad. En cuanto a la validez convergente, la varianza media extraída (AVE) superó el criterio de .50 en todos los factores, con valores entre .503 y .604 en el M1 y entre .599 y .652 en el M2 (Fornell y Larcker, 1981). Cabe señalar que el AVE de F4 en el M1 fue apenas superior al umbral. Respecto a la validez discriminante, conforme al criterio de Fornell y Larcker (1981), la raíz cuadrada del AVE de cada factor debe ser mayor que sus correlaciones con los demás factores. Este criterio no se cumplió en la mayoría de los pares de factores del M1 (Tabla 2 y 3), solo se cumplió entre Comunicación y Atención (F2 y F3). En el M2, el criterio se cumplió en la mitad de los pares y se incumplió entre Comprensión y Atención (F1-F3), Comprensión y Excelencia (F1-F4) y Atención y Excelencia (F3-F4). Por tanto, aunque el M2 mostró una mejora, la validez discriminante fue limitada en ambos modelos.

Tabla 3. Coeficientes Omega, Alfa y AVE de los factores en los modelos M1 y M2 del análisis factorial confirmatorio.

Factor	M1			M2		
	Ω	α	AVE	Ω	α	AVE
F1	.821	.819	.604	.817	.819	.599
F2	.852	.848	.590	.830	.828	.620
F3	.876	.874	.587	.849	.834	.652
F4	.801	.792	.503	.763	.763	.617
Escala total	.955	.933		.948	.912	

Nota: Ω= Omega, α = Alfa, AVE = Varianza media extraída. F1= Comprensión, F2 = Comunicación, F3 = Atención, F4= Excelencia.

Discusión

Los resultados obtenidos reflejan una mejora considerable en el ajuste global y la capacidad explicativa del modelo al pasar de M1 a M2. Esto deja en claro la importancia de reconocer reactivos problemáticos, ya sea por cargas factoriales bajas o por problemas identificados en los índices de modificación, que tal

vez deban ser eliminados porque no contribuyen adecuadamente al modelo. Esto concuerda con lo mencionado por Kline (2016), quien apoya la eliminación de reactivos que causan tensión en el modelo, estos identificados mediante los índices de modificación, e indica que un modelo con menos ítems, pero con mejores indicadores de ajuste (RMSEA, CFI, GFI) es preferible. Ferrando et al. (2022) mencionan que los ítems con problemas y que permanecen en el modelo, producen distorsiones y sesgos que afectan incluso a los ítems que funcionan bien, y que aquellos reactivos con valores menores de .50 se comportan casi de forma aleatoria, por lo que es mejor eliminarlos. La estructura de cuatro factores obtenida en este estudio coincide con la propuesta por Rangel (2015) y con la replicada por Blanco et al. (2016); al igual que en este trabajo el modelo reducido mostró un mejor ajuste que el modelo original. Así mismo, se observan altas intercorrelaciones entre los factores, resultados que coinciden con Blanco et al. (2016). A pesar de la mejora del modelo, la persistencia de estas intercorrelaciones plantea interrogantes sobre la diferenciación conceptual entre los factores. Este patrón, junto con la estructura de tres factores identificada por Blanco et al. (2011), sugiere que la distinción entre dimensiones podría no ser la más parsimoniosa, aunque es evidente la necesidad de un análisis más profundo. Los resultados aportan evidencia favorable, aunque parcial sobre la estructura interna del instrumento. El M2 reproduce las cuatro dimensiones de autoeficacia académica propuestas previamente (Blanco et al., 2016; Rangel, 2015). Sin embargo, las altas intercorrelaciones indican una validez discriminante limitada, por lo que las dimensiones podrían reflejar un constructo más general de autoeficacia académica y no facetas plenamente diferenciadas.

Estos hallazgos podrían contribuir a desarrollar un modelo más preciso para comprender cómo los estudiantes perciben su propia capacidad en el ámbito académico. La autoeficacia juega un papel clave en el éxito universitario, ya que influye en la confianza con la que enfrentan desafíos y en su capacidad para mantenerse motivados a lo largo de su formación. Para completar sus estudios, no solo necesitan buenas estrategias de aprendizaje, sino también creer en sus propias habilidades y potencial (Dominguez-Lara et al., 2023). El análisis factorial confirmatorio permite evaluar la calidad de los instrumentos de medición, al verificar si las variables se agrupan en un número menor de factores y a cuál se asocia cada una según el ajuste del modelo (Maldonado, 2019).

Dado que el instrumento ya cuenta con evidencia de validación en población mexicana (Rangel, 2015), este estudio se centró en la evidencia basada en la estructura interna, por lo que otras fuentes de evidencia, como la relación con variables externas, la invarianza factorial o la estabilidad temporal, deberán abordarse en estudios posteriores. En este sentido, la propuesta derivada del análisis factorial confirmatorio aporta evidencia empírica sobre la estructura interna del instrumento.

Para las ciencias de la actividad física y del deporte, la versión depurada (M2) ofrece una medida más breve de la autoeficacia académica que puede utilizarse para identificar necesidades de apoyo académico durante su formación. No obstante, su uso debe limitarse por ahora a fines de investigación y tamizaje, hasta contar con evidencia en una muestra independiente y con otras fuentes de validez. Estos resultados son relevantes para la formación en ciencias de la actividad física y del deporte, donde la autoeficacia académica puede asociarse con la práctica deportiva y con el desempeño en asignaturas prácticas (Adigue y Adigue, 2025; Wang et al., 2022), una relación que futuras investigaciones podrían examinar con este instrumento.

Conclusiones

El análisis factorial confirmatorio realizado permitió evaluar la calidad y el ajuste de dos modelos de medición propuestos. Los resultados del modelo inicial (M1) indicaron un ajuste aceptable (GFI = .933; RMSEA = .068; CFI = .950) y una capacidad explicativa del 69.36% de la varianza total. Sin embargo, se identificaron limitaciones importantes, como la presencia de tres ítems con saturaciones inferiores al umbral de .70 (ítems 02, 01 y 16), esto también fue identificado por Blanco et al. (2016), y la existencia de intercorrelaciones muy altas entre los factores: F1-F3 con un coeficiente de correlación de .85, entre los factores F1-F4 con un coeficiente de .88 y entre F3-F4 con coeficiente de correlación de .91. Esto evidencia que, aunque la fiabilidad global de la escala completa del M1 es adecuada (Ω = .955; α = .933), existen problemas de validez discriminante, ya que los factores no se diferenciaban claramente entre sí.

En el M2 se eliminaron cinco ítems (09, 02, 03, 01 y 11) que mostraron problemas de ajuste o de redundancia con otros reactivos, y se especificó una correlación entre los errores de los ítems 08 y 12, sugerida por los índices de modificación. Con estos cambios se obtuvo un mejor ajuste (GFI = .981; RMSEA = .042; CFI = .989) y un mayor porcentaje de varianza explicada (76.45%). Además, ninguno de los ítems restantes presentó saturaciones inferiores a .74, lo que indica una mejora sustancial en la relación entre los ítems y sus respectivas dimensiones. No obstante, se observaron nuevamente intercorrelaciones de moderadas a muy altas entre los factores (r = .65 a .87), por lo que la validez discriminante sigue siendo una limitación de este modelo. Aunque la fiabilidad global del M2 fue ligeramente menor que la del M1, se mantuvo adecuada (Ω = .948; α = .912), por lo que la eliminación de ítems no afectó de manera sustancial la fiabilidad de la escala.

Limitaciones

Este estudio presenta tres limitaciones principales. Como primera restricción, la muestra estuvo conformada de manera exclusiva por estudiantes de la Universidad Autónoma de Chihuahua, lo que limita la generalización de los resultados con otras instituciones y contextos educativos. En segundo lugar, el uso de cuestionarios autoaplicados podría haber introducido sesgos de respuesta, como la deseabilidad social. Esto podría afectar la calidad de los datos obtenidos. Tercero, las altas intercorrelaciones entre algunos factores y la validez discriminante limitada sugieren que la estructura de cuatro factores podría no ser la más parsimoniosa. Dado que solo se compararon los modelos M1 y M2, no se evaluaron estructuras alternativas, como la unifactorial, la bifactorial o la de segundo orden. Futuras investigaciones deberían contrastarlas comparando su ajuste y su interpretabilidad teórica y en una etapa posterior, confirmar la versión reducida del instrumento (M2) en una muestra independiente, evaluando su estructura factorial, su fiabilidad, su validez y la representatividad de su contenido.

Referencias

Adigue, M. L. P., & Adigue, N. (2025). From motion to motivation: The relationship of physical activity on Filipino students' self-efficacy for learning. *Sportis Scientific Journal of School Sport, Physical Education and Psychomotricity, 11*(4), 1–18. https://doi.org/10.17979/sportis.2025.11.4.11967

Ávalos, M. L., Oropeza, R., Ramírez, J. C., & Platos, M. Ú. (2018). Percepción de autoeficacia y rendimiento académico en estudiantes de bachillerato. *Caleidoscopio, 22*(39), 33–46. https://doi.org/10.33064/39crscsh1323

Bagozzi, R. P., & Yi, Y. (2012). Specification, evaluation, and interpretation of structural equation models. *Journal of the Academy of Marketing Science*, *40* (1), 8–34. https://doi.org/10.1007/s11747-011-0278-x

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*(2), 191–215. https://doi.org/10.1037/0033-295X.84.2.191

Bentler, P. M. (2017). Specificity-enhanced reliability coefficients. *Psychological Methods*, *22*(3), 527–540. https://doi.org/10.1037/met0000092

Blanco, H., Aguirre, J. F., Barrón, J. C., & Blanco, J. R. (2016). Composición factorial de la escala de autoeficacia académica en universitarios mexicanos. *Formación Universitaria*, *9*(2), 81–88. https://doi.org/10.4067/S0718-50062016000200009

Blanco, H., Martínez, M., Zueck, M. C., & Gastélum, G. (2011). Análisis psicométrico de la escala autoeficacia en conductas académicas en universitarios de primer ingreso. *Revista Electrónica Actualidades Investigativas en Educación, 11*(3), 1–27. https://doi.org/10.15517/aie.v11i3.10214

Blanco, H., Ornelas, M., Tristán, J. L., Cocca, A., Mayorga-Vega, D., López-Walle, J., & Viciana, J. (2013). Editor for creating and applying computerise surveys. *Procedia - Social and Behavioral Sciences, 106*, 935–940. https://doi.org/10.1016/j.sbspro.2013.12.105

Borzone, M. A. (2017). Autoeficacia y vivencias académicas en estudiantes universitarios. *Acta Colombiana de Psicología, 20*(1), 266–274. https://doi.org/10.14718/ACP.2017.20.1.13

Byrne, B. M. (2016). *Structural equation modeling with AMOS: Basic concepts, applications, and programming* (3rd ed.). Routledge. https://doi.org/10.4324/9781315757421

Carracedo, M., Sánchez, D., & Zunino, C. (2017). Consentimiento informado en investigación. *Anales de la Facultad de Medicina*, 4(Supl.2), 16–21.

Casanova, J., Sinval, J., & Almeida, L. (2024). Academic success, engagement and self-efficacy of first-year university students: Personal variables and first-semester performance. *Anales de Psicología*, *40*(1), 44–53. https://doi.org/10.6018/analesps.479151

Deng, L., & Chan, W. (2017). Testing the difference between reliability coefficients alpha and omega. *Educational and Psychological Measurement*, *77(*2), 185–203. https://doi.org/10.1177/0013164416658325

Dominguez-Lara, S. (2016). Valores normativos de una escala de autoeficacia académica en estudiantes universitarios de Lima. *Interacciones*, *2*(2), 91–98. https://doi.org/10.24016/2016.v2n2.31

Dominguez-Lara, S., Alarcón-Parco, D., Campos-Uscanga, Y., Tamayo-Agudelo, W., Merino-Soto, C., Tumino, M. C., Quinde, J. M., Barroca de Moura, G., & de Almeida Baptista, M. das G. (2023). Psychometric properties and measurement invariance of an academic self-efficacy scale in college students from five Latin American countries. *Ciencias Psicológicas*, *17*(1), e3051. https://doi.org/10.22235/cp.v17i1.3051

Ferrando, P. J., Lorenzo-Seva, U., Hernández-Dorado, A., & Muñiz, J. (2022). Decálogo para el análisis factorial de los ítems de un test. *Psicothema*, *34*(1), 7–17. https://doi.org/10.7334/psicothema2021.456

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/10.1177/002224378101800104

García-Álvarez, D., Cobo-Rendón, R., & Hernández-Lalinde, J. (2022). Validez, fiabilidad e invarianza factorial de las escalas de autoeficacia general y autoeficacia académica en estudiantes universitarios. *Retos*, *46*, 1093–1104. https://doi.org/10.47197/retos.v46.94281

García-Méndez, R., & Rivera-Ledesma, A. (2021). Escala de autoeficacia en la vida académica: Propiedades psicométricas en estudiantes de nuevo ingreso al nivel universitario. *Revista Electrónica Educare*, *25*(2), 1–28. https://doi.org/10.15359/ree.25-2.1

Gelabert, E., García-Esteve, L., Martín-Santos, R., Gutiérrez, F., Torres, A., & Subirà, S. (2011). Psychometric properties of the Spanish version of the Frost Multidimensional Perfectionism Scale in women. *Psicothema*, *23*(1), 133–139

Gutiérrez-García, A. G., & Landeros-Velázquez, M. G. (2018). Autoeficacia académica y ansiedad, como incidente crítico, en mujeres y hombres universitarios. *Revista Costarricense de Psicología, 37*(1), 1–25. https://doi.org/10.22544/rcps.v37i01.01

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

Maldonado, J. A. (2019). Introducción al análisis factorial confirmatorio: Un enfoque aplicado*. INEVA en Acción, 14*(1), 1–4.

Montero, I., & León, O. G. (2007). A guide for naming research studies in Psychology. *International Journal of Clinical and Health Psychology, 7*(3), 847–862. https://www.aepc.es/ijchp/GNEIP07_es.pdf

Ornelas, M., Blanco, H., Gastélum, G., & Muñoz, F. (2013). Perfiles de autoeficacia en conductas académicas en alumnos de primer ingreso a la universidad. *Revista Electrónica de Investigación Educativa, 15*(1), 17–28. https://www.scielo.org.mx/pdf/redie/v15n1/v15n1a2.pdf

Rangel, Y. S. (2015). *Validación de una escala de autoeficacia académica: Análisis de un modelo mediante ecuaciones estructurales* [Tesis doctoral no publicada]. Universidad Autónoma de Chihuahua, México.

Robles, H. (2020). Escala de autoeficacia académica en estudiantes universitarios en Lima. *Avances en Psicología*, *28*(1), 99–107. https://doi.org/10.33539/avpsicol.2020.v28n1.2115

Rodríguez, M. N., & Ruiz, M. Á. (2008). Atenuación de la asimetría y de la curtosis de las puntuaciones observadas mediante transformaciones de variables: Incidencia sobre la estructura factorial. *Psicológica, 29*(2), 205–227. https://www.redalyc.org/articulo.oa?id=16929206

Rosales-Ronquillo, C., & Hernández-Jácquez, L. (2020). Autoeficacia académica y su relación con el rendimiento académico en estudiantes de nutrición. *Revista Electrónica Educare*, *24*(3), 1–17. https://doi.org/10.15359/ree.24-3.7

Ruiz, M. A., Pardo, A., & San Martín, R. (2010). Modelos de ecuaciones estructurales. *Papeles del Psicólogo*, *31*(1), 34–45. https://www2.papelesdelpsicologo.es/pdf/1794.pdf

Sánchez-Anguita, Á., & Pulido, M. F. (2022). Liderazgo y autoeficacia como predictores del rendimiento académico. *International Journal of Developmental and Educational Psychology*, *1*(1), 453–462. https://doi.org/10.17060/ijodaep.2022.n1.v1.2403

Secretaría de Salud. (2024). *Reglamento de la Ley General de Salud en Materia de Investigación para la Salud.* Diario Oficial de la Federación. https://salud.gob.mx/unidades/cdi/nom/compi/rlgsmis.html

Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications. American Psychological Association*. https://doi.org/10.1037/10694-000

Viladrich, C., Angulo-Brunet, A., & Doval, E. (2017). Un viaje alrededor de alfa y omega para estimar la fiabilidad de consistencia interna. *Anales de Psicología*, *33*(3), 755–782. https://doi.org/10.6018/analesps.33.3.268401

Wang, K., Li, Y., Zhang, T., & Luo, J. (2022). The relationship among college students' physical exercise, self-efficacy, emotional intelligence, and subjective well-being. *International Journal of Environmental Research and Public Health, 19*(18), 11596. https://doi.org/10.3390/ijerph191811596

Yupanqui-Lorenzo, D. E., Mollinedo, F. M., & Montealegre, A. C. (2021). Modelo explicativo de la autoeficacia académica: Autorregulación de actividades, afecto positivo y personalidad. *Propósitos y Representaciones*, *9*(2), e755. https://doi.org/10.20511/pyr2021.v9n2.755

Datos de los/as autores/as y traductor/a:

Edgar Francisco Ordoñez Bencomo	eordonez@uach.mx	Autor
Susana Ivonne Aguirre Vásquez	siaguirre@uach.mx	Autora
Jesús Enrique Peinado Pérez	jpeinad@uach.mx	Autor
José René Blanco Ornelas	jblanco@uach.mx	Autor