

# Validación psicométrica de la EAEIA-GAD: Escala de Actitud Estudiantil hacia la Inteligencia Artificial Generativa en Actividades Docentes

**Luis Gerardo Meza-Cascante**

Escuela de Matemática  
Instituto Tecnológico de Costa Rica, Costa Rica  
✉ gемеza@itcr.ac.cr

**Melvin Ramirez Bogantes**

Escuela de Matemática  
Instituto Tecnológico de Costa Rica, Costa Rica  
✉ meramirez@itcr.ac.cr

**Leonel Chaves Salas**

Escuela de Matemática  
Instituto Tecnológico de Costa Rica, Costa Rica  
✉ leochaves@itcr.ac.cr

Fecha de recepción: 27 de setiembre del 2025 | Fecha de aprobación: 15 de abril del 2026

## Resumen

Este estudio presenta la validación de una escala diseñada para conocer cómo percibe el estudiantado universitario el uso de la inteligencia artificial generativa (IA-G) en actividades de aprendizaje. Para ello, se aplicó un cuestionario breve de 10 preguntas a 156 estudiantes de ingeniería, utilizando una escala de respuesta que permite expresar distintos niveles de acuerdo. Los resultados muestran que el instrumento es confiable y consistente, es decir, que las preguntas funcionan de manera coherente entre sí y permiten medir adecuadamente una misma idea: la actitud del estudiantado hacia el uso de IA en el aula. En términos sencillos, esto significa que la escala realmente capta si a los estudiantes les resulta útil, interesante o motivadora la incorporación de estas herramientas en su proceso de aprendizaje. Además, se encontró que las respuestas tienden a ser muy positivas, lo que indica una buena disposición hacia el uso de la IA, aunque también sugiere la necesidad de mejorar el instrumento para diferenciar mejor entre distintos niveles de entusiasmo. En conjunto, la escala propuesta constituye una herramienta práctica y fácil de aplicar que puede ayudar a docentes e investigadores a comprender cómo se reciben estas tecnologías en el aula y a tomar decisiones informadas sobre su uso educativo.

**Palabras clave:** inteligencia artificial generativa, actitudes, validación psicométrica, educación superior

## Abstract

This study presents the validation of a scale designed to assess how university students perceive the use of generative artificial intelligence (AI) in learning activities. A short 10-item questionnaire was administered to 156 engineering students using a response scale that captures different levels of agreement. The results show that the instrument is reliable and consistent, meaning that the questions work together coherently and effectively measure a single underlying idea: students' attitudes toward the use of AI in the classroom. In simple terms, the scale can capture whether students find these tools useful, engaging, and motivating for their learning. In addition, responses tended to be highly positive, indicating a favorable disposition toward the use of AI. However, this also suggests the need to refine the instrument to better distinguish between different levels of enthusiasm. Overall, the proposed scale is a practical and easy-to-use tool that can support teachers and researchers in understanding how these technologies are received in educational settings and in making informed decisions about their integration into teaching practices.

**Keywords:** generative artificial intelligence, attitudes, psychometric validation, higher education

## Introducción

En los últimos años, la incorporación de la inteligencia artificial generativa (IA-G) en la educación superior ha abierto nuevas posibilidades, pero también ha generado debates sobre su impacto pedagógico y ético. Si bien abundan estudios sobre la eficacia de estas herramientas, existe una necesidad creciente de comprender cómo se perciben en contextos docentes concretos.

No obstante, los instrumentos disponibles para medir actitudes hacia la IA —como la General Attitudes towards Artificial Intelligence Scale (GAAIS) [13] o escalas centradas en intenciones de uso y dependencia tecnológica [12]— capturan percepciones de carácter general que no reflejan necesariamente cómo reacciona el estudiantado frente a actividades instruccionales concretas diseñadas e implementadas por el docente en el aula. Esta brecha justifica el desarrollo de una escala situada, capaz de valorar la actitud estudiantil en escenarios de enseñanza específicos mediados por IA-G, diferenciándose así de los instrumentos existentes tanto en su foco constructual como en su aplicabilidad para el monitoreo formativo en cursos universitarios.

El objetivo de este trabajo es aportar evidencia psicométrica inicial sobre una escala que evalúa actitudes del estudiantado hacia la IA-G en situaciones instruccionales concretas de aula creadas por la persona docente. La contribución es doble: (a) proveer una medida parsimoniosa que pueda usarse para monitoreo formativo en cursos con IA-G, y (b) delimitar el constructo actitud frente a otros relacionados, como la intención de uso, la dependencia y la autoeficacia. El estudio se guía por tres preguntas centrales: ¿la puntuación total exhibe fiabilidad y estructura interna coherentes con un factor general?, ¿qué ítems requieren revisión por bajo desempeño relativo?, y ¿qué ajustes permiten aumentar la sensibilidad del instrumento en niveles altos del rasgo?

## Marco teórico

### 1. Actitudes y dominio afectivo en educación

El estudio sistemático de las actitudes en educación tiene una larga trayectoria, en la que se reconoce como pionero [1] al destacar la importancia del dominio afectivo en los procesos de aprendizaje. No obstante, los análisis actuales sobre actitudes hacia la tecnología y la inteligencia artificial (IA) se fundamentan en investigaciones recientes que muestran cómo estas predisposiciones influyen en la adopción y el uso crítico de innovaciones educativas. En educación superior, se ha demostrado que las actitudes, junto con percepciones de utilidad, confianza y facilidad de uso, determinan en gran medida la disposición del estudiantado a incorporar tecnologías emergentes [2]-[4]. En este marco, comprender las actitudes hacia la IA resulta clave para interpretar los procesos de aceptación, resistencia o uso crítico de estas herramientas en la enseñanza universitaria.

### 2. Actitudes hacia la IA-G en educación superior

En los últimos años, la investigación sobre actitudes hacia la IA-G en educación superior ha crecido de manera significativa. En general, los estudios coinciden en señalar una valoración positiva por parte del estudiantado, destacando su utilidad para el aprendizaje, la motivación y el apoyo en tareas académicas [5]-[9].

No obstante, esta valoración favorable coexiste con reservas importantes. Investigaciones recientes evidencian preocupaciones relacionadas con la autoría de los trabajos, la calidad de las respuestas generadas y el riesgo de dependencia tecnológica [6], [10], [11]. Estas tensiones entre potencial pedagógico y desafíos éticos configuran el contexto en el que resulta necesario desarrollar instrumentos de medición más específicos y situados.

#### 2.1 Actitudes del estudiantado hacia el uso de la IA-G en actividades docentes

La evidencia reciente con estudiantes universitarios confirma que las actitudes hacia la IA-G en docencia son en promedio favorables. En [5] se validó una escala que identifica tres componentes de estas actitudes: cognitivo, afectivo y comportamental, concluyendo que existe una disposición positiva hacia la integración de la IA en la enseñanza y el aprendizaje. Las actitudes varían en función de para qué se usa la IA (por ejemplo., exploración conceptual, retroalimentación, generación de borradores), cómo se enmarca pedagógicamente (rutas de verificación, criterios de autoría) y qué riesgos se hacen explícitos (sesgos, dependencia, calidad de fuentes). Este patrón obliga a desplazar la discusión desde un "sí/no a la IA" hacia condiciones de uso que activan o inhiben la aceptación crítica.

En contextos concretos, como el uso de ChatGPT en actividades académicas, los hallazgos revelan simultáneamente beneficios y reservas. Diversos estudios coinciden en señalar que el estudiantado reconoce la utilidad académica de estas herramientas, pero expresa reservas sobre la autoría, la fiabilidad de las respuestas y los riesgos de dependencia tecnológica [6], [10], [11].

En esa misma línea, en [12] se desarrolló y validó la Scale for Dependence on Artificial Intelligence (DAI) con estudiantes universitarios, aportando evidencias de fiabilidad, estructura interna e invarianza por género. Esta herramienta constituye un referente clave para medir el grado de dependencia estudiantil frente al uso de IA en educación superior, fortaleciendo la necesidad de diferenciar entre actitudes generales hacia la IA y actitudes específicas frente a su incorporación en actividades docentes.

### 3. Instrumentos de medición de actitudes hacia la IA-G

Para captar estas percepciones se han desarrollado instrumentos psicométricos con respaldo empírico. Entre ellos destaca la General Attitudes towards Artificial Intelligence Scale (GAAIS), validada en [13], que permite medir actitudes generales hacia la IA mediante ítems de valoración positiva y negativa. Esta escala ha mostrado propiedades métricas robustas en diferentes contextos, consolidándose como un referente en la medición de actitudes globales.

En el ámbito universitario, los estudios más recientes han señalado la necesidad de contar con instrumentos más específicos, que aborden percepciones sobre el uso de la IA en actividades

académicas concretas. Algunos trabajos han avanzado en esa dirección, abordando actitudes hacia ChatGPT en contextos académicos [6] y el desarrollo de escalas de dependencia tecnológica en estudiantes universitarios [12]. Estos aportes reflejan un esfuerzo emergente por transitar desde instrumentos de carácter general hacia escalas situadas, capaces de captar con mayor precisión las actitudes del estudiantado en escenarios de enseñanza universitaria mediados por IA.

## Metodología

Se aplicó un cuestionario de 10 afirmaciones a 156 estudiantes de ingeniería. Este tamaño muestral equivale a una razón de 15.6 participantes por ítem, lo que supera el criterio habitual de 5 a 10 participantes por ítem recomendado para análisis psicométrico exploratorio, y proporciona suficiencia estadística adecuada para los análisis de fiabilidad y estructura interna realizados. Para responder, se utilizó una escala Likert de 7 puntos (una escala de valoración numerada del 1 al 7, donde 1 significa “totalmente en desacuerdo” y 7 significa “totalmente de acuerdo”), lo que permite capturar con precisión el grado de acuerdo o desacuerdo de cada persona. Los estudiantes fueron seleccionados mediante muestreo por conglomerados, es decir, se trabajó con grupos completos ya formados (clases del curso), en lugar de elegir a cada estudiante de forma individual. Las afirmaciones del cuestionario evaluaron dimensiones como interés, claridad, motivación, utilidad, dificultad, pensamiento crítico, colaboración, retroalimentación, disposición a repetir la experiencia y mejora en comprensión conceptual. El uso de escalas tipo Likert es habitual en investigación educativa por su simplicidad y capacidad de capturar percepciones cuantitativas [14].

Los ítems del instrumento aplicado son los siguientes:

1. La actividad mediada por IA me resultó interesante.
2. Las instrucciones y objetivos de la actividad mediada por IA fueron claros.
3. La actividad mediada por IA me motivó a aprender.
4. Considero que la actividad fue útil para mi aprendizaje.
5. El nivel de dificultad de la actividad fue adecuado para mi nivel.
6. La actividad fortaleció mi pensamiento crítico.
7. La actividad promovió la colaboración entre compañeros.
8. La retroalimentación recibida fue oportuna y útil.
9. Repetiría actividades de este tipo en futuras lecciones.
10. El uso de IA en la actividad mejoró mi comprensión de los contenidos.

La Tabla 1 presenta los diez ítems que conforman la EAEIA-GAD, organizados junto con la dimensión actitudinal que cada uno evalúa, la puntuación media obtenida (en una escala de 1 a 7) y la correlación ítem-total corregida, que indica qué tan bien contribuye cada pregunta a medir el constructo global de la escala.

**Tabla 1.** Ítems de la EAEIA-GAD, dimensión evaluada, puntuación media y correlación ítem–total corregida (n = 156)

N°	Enunciado del ítem	Dimensión evaluada	Media	r ítem-total
1	La actividad mediada por IA me resultó interesante.	Interés	5.98	0.75
2	Las instrucciones y objetivos de la actividad mediada por IA fueron claros.	Claridad	5.81	0.58
3	La actividad mediada por IA me motivó a aprender.	Motivación	5.67	0.72
4	Considero que la actividad fue útil para mi aprendizaje.	Utilidad	5.89	0.78
5	El nivel de dificultad de la actividad fue adecuado para mi nivel.	Dificultad	5.43	0.45
6	La actividad fortaleció mi pensamiento crítico.	Pensamiento crítico	5.72	0.65
7	La actividad promovió la colaboración entre compañeros.	Colaboración	5.48	0.29*
8	La retroalimentación recibida fue oportuna y útil.	Retroalimentación	5.65	0.52
9	Repetiría actividades de este tipo en futuras lecciones.	Disposición a repetir	5.91	0.68
10	El uso de IA en la actividad mejoró mi comprensión de los contenidos.	Comprensión	6.03	0.71

\* Por debajo del umbral recomendado de 0.30.  $r$  = correlación ítem–total corregida.  $n$  = 156.

Como se aprecia en la tabla 1, la gran mayoría de los ítems obtuvieron puntuaciones medias elevadas y correlaciones ítem–total aceptables. La única excepción es el ítem 7 (colaboración), cuya correlación de 0.29 se ubica levemente por debajo del umbral recomendado de 0.30, lo que sugiere la necesidad de revisar su formulación en versiones futuras del instrumento.

El análisis estadístico incluyó los siguientes procedimientos: (a) fiabilidad mediante  $\alpha$  de Cronbach (alfa de Cronbach: un índice que indica qué tan consistentes son las respuestas a los distintos ítems del cuestionario; va de 0 a 1, y valores por encima de 0.70 se consideran aceptables) y  $\omega$  de McDonald (omega de McDonald: un índice similar al alfa pero más preciso en ciertos contextos), recomendados para valorar la consistencia interna [15]–[17]; (b) los índices de adecuación muestral KMO (Kaiser-Meyer-Olkin: una medida que indica si los datos son adecuados para hacer un análisis de factores; valores cercanos a 1 son excelentes) y la prueba de esfericidad de Bartlett (una prueba estadística que verifica si las respuestas a los ítems están suficientemente relacionadas entre sí como para justificar el análisis factorial) [18], [19]; (c) un análisis paralelo de Horn (criterio de Horn: una técnica estadística que ayuda a determinar cuántos factores o dimensiones subyacen en los datos, comparando los resultados reales con valores generados al azar) para determinar el número de factores [20]; y (d) correlaciones ítem–total corregidas (un indicador que muestra qué tanto contribuye cada pregunta a medir lo mismo que el conjunto de ítems; valores por encima de 0.30 son deseables). Los análisis se realizaron con SPSS, versión 29 [21]. En términos prácticos, esto significa que las respuestas del estudiantado siguen un patrón coherente, es decir, las preguntas del cuestionario apuntan en la misma dirección. Además, las preguntas están lo suficientemente relacionadas entre sí como para analizarlas como un conjunto, y los resultados obtenidos muestran que los datos cumplen con las condiciones necesarias para realizar este tipo de análisis con confianza.

## Resultados

La EAEIA-GAD mostró una consistencia interna elevada, con un alfa de Cronbach ( $\alpha$ ) de 0.89 y un omega de McDonald ( $\omega$ ) de 0.80 —ambos indicadores de qué tan coherentemente miden lo mismo las preguntas del cuestionario—, valores que superan el umbral de 0.70 habitualmente recomendado en ciencias sociales, lo que indica que la escala presenta una alta consistencia interna: los distintos ítems miden de forma coherente el mismo constructo (la actitud hacia la IA en actividades docentes). Por ejemplo, un alfa de 0.89 es comparable al nivel de consistencia que se esperaría de un termómetro que da lecturas muy similares al medirse varias veces seguidas. Las correlaciones ítem–total corregidas (indicador de qué tanto aporta individualmente cada pregunta al conjunto) oscilaron entre 0.29 y 0.78; la mayoría de los ítems se ubicaron por encima de 0.30, salvo el enunciado “La actividad promovió la colaboración entre compañeros”, que alcanzó un valor de 0.29, indicando que ese ítem tiene una relación ligeramente más débil con el resto de la escala. En síntesis, las respuestas del estudiantado son consistentes entre sí.

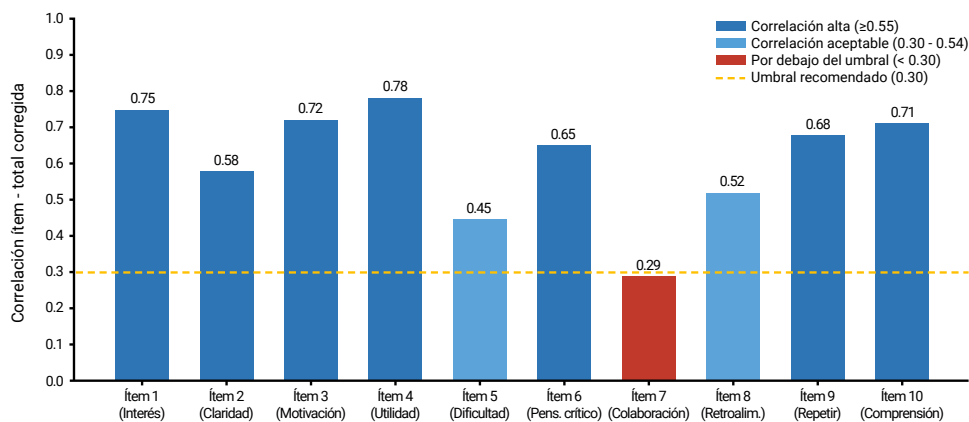
A continuación, la Tabla 2 ofrece una vista consolidada de todos los indicadores psicométricos calculados para la EAEIA-GAD, organizados en tres bloques: fiabilidad, validez factorial y discriminación de ítems. Para cada indicador se presenta el valor obtenido, el criterio de referencia utilizado en ciencias sociales y una interpretación.

**Tabla 2.** Resumen de indicadores psicométricos de la EAEIA-GAD (n = 156)

Indicador	Valor obtenido	Criterio de referencia	Interpretación
Fiabilidad			
Alfa de Cronbach ( $\alpha$ )	<b>0.89</b>	> 0.70	Alta consistencia interna: los ítems miden el mismo constructo de forma coherente.
Omega de McDonald ( $\omega$ )	<b>0.80</b>	> 0.70	Confirma la alta consistencia; más robusto que el alfa en ciertos contextos.
Validez factorial			
Índice KMO	<b>0.91</b>	> 0.80	Adecuación muestral excelente; los datos son idóneos para el análisis factorial.
Prueba de Bartlett	$\chi^2(45) = 794.65, p < 0.001$	$p < 0.05$	Matriz de correlaciones significativa; justifica la aplicación del análisis factorial.
Primer autovalor real	<b>5.37</b>	> 1.33 (P95)	Supera el umbral aleatorio del análisis paralelo de Horn.
Umbral aleatorio (P95)	<b>1.33</b>	—	Valor de referencia del análisis paralelo para 10 ítems.
Factores retenidos	<b>1</b>	—	Estructura unidimensional: todos los ítems miden un único constructo general.
Discriminación y efecto techo			
Rango correlaciones ítem-total	<b>0.29 – 0.78</b>	> 0.30	9 de 10 ítems superan el umbral; el ítem 7 (colaboración) se ubica en 0.29.
Rango de medias por ítem	<b>5.43 – 6.03</b>	—	Efecto techo: respuestas concentradas en los valores altos de la escala (1–7). Los valores de la tabla anterior permiten apreciar, de un solo vistazo, que la escala cumple satisfactoriamente con todos los criterios de validez y fiabilidad evaluados. A continuación se amplía la interpretación de los indicadores de validez factorial y se presentan las figuras correspondientes.

La adecuación de la matriz de correlaciones se confirmó con un índice KMO de 0.91 —valor considerado excelente, pues indica que los datos son muy apropiados para realizar un análisis factorial— y una prueba de esfericidad de Bartlett estadísticamente significativa ( $\chi^2(45) = 794.65, p < 0.001$ ), lo que confirma que las respuestas del estudiantado a los distintos ítems están suficientemente interrelacionadas como para extraer factores comunes. Ambos resultados, en conjunto, dan luz verde al análisis factorial. El análisis paralelo de Horn indicó que únicamente el primer autovalor —el componente principal de variación en los datos— (5.37) superó el umbral aleatorio correspondiente al percentil 95 (1.33), lo que respalda que la escala mide una sola dimensión o constructo general (estructura unidimensional), es decir, todas las preguntas apuntan hacia un mismo concepto subyacente: la actitud hacia el uso de la IA en actividades docentes. Esto significa que las preguntas están suficientemente relacionadas como para analizarse en conjunto.

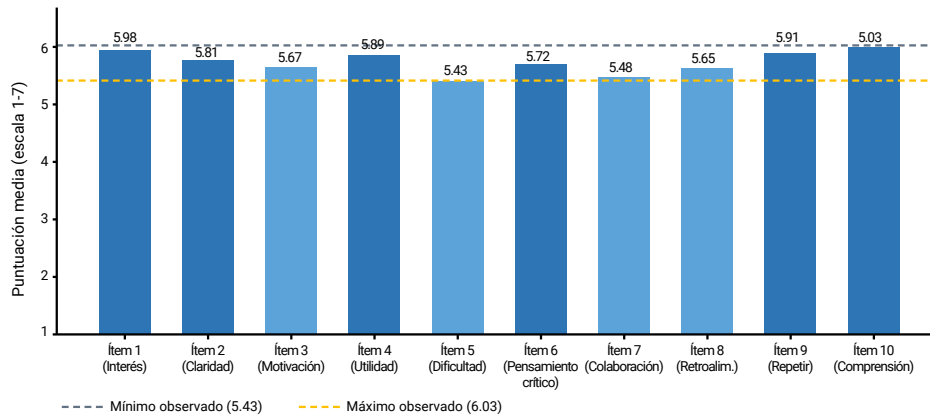
La Figura 1 ilustra gráficamente las correlaciones ítem–total corregidas de cada uno de los diez ítems. Las barras permiten comparar visualmente el aporte de cada pregunta al instrumento: azul oscuro para correlaciones altas ( $\geq 0.55$ ), azul claro para correlaciones aceptables (0.30–0.54) y rojo para el único ítem que se ubica por debajo del umbral mínimo recomendado de 0.30, indicado con la línea naranja discontinua.



**Figura 1.** Correlaciones ítem–total corregidas de la EAEIA-GAD (n = 156). Las barras rojas indican ítems por debajo del umbral mínimo recomendado de 0.30 (línea naranja discontinua).

La Figura 2 muestra la distribución de las puntuaciones medias por ítem. La zona sombreada y las líneas de referencia permiten visualizar con claridad el rango en que se concentran todas las respuestas (5.43–6.03), lo que evidencia el efecto techo descrito anteriormente.

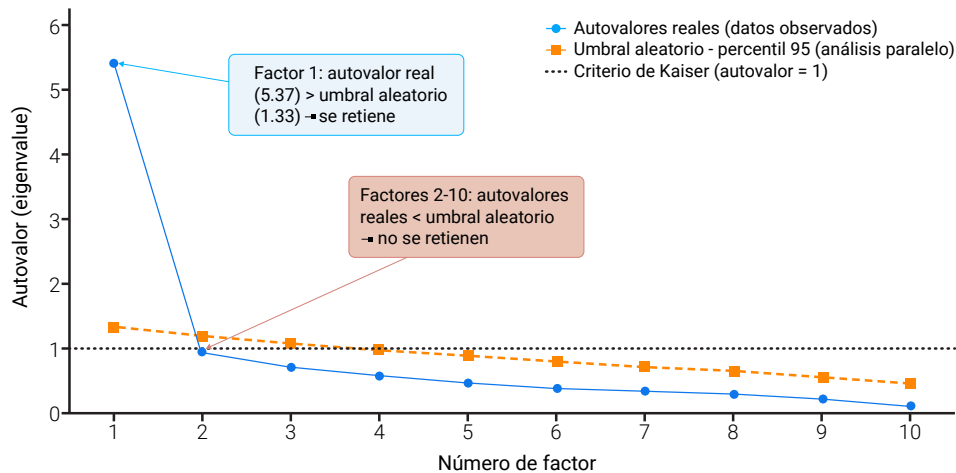
Las medias (promedios) de los ítems oscilaron entre 5.43 y 6.03 en una escala de 1 a 7, como se muestra en la Figura 2. Esta concentración de valores en los rangos altos sugiere la presencia de un efecto techo, un fenómeno que ocurre cuando la mayoría de las personas responden con puntuaciones muy altas, lo que reduce la capacidad del instrumento para distinguir entre quienes tienen actitudes moderadamente positivas y quienes las tienen muy positivas. En términos prácticos, es como si un examen fuera tan fácil que casi todos sacaran 9 o 10, sin poder distinguir quién domina mejor el tema.



**Figura 2.** Puntuaciones medias por ítem de la EAEIA-GAD (escala 1–7, n = 156). La zona sombreada representa el rango de variación observado (5.43–6.03), evidenciando el efecto techo en las respuestas del estudiantado.

La Figura 3 presenta el resultado del análisis paralelo de Horn. La línea azul representa los autovalores reales obtenidos con los datos del estudio, mientras que la línea naranja discontinua indica el umbral generado aleatoriamente (percentil 95). Solo cuando la línea azul supera a la naranja se justifica retener ese factor; en este caso únicamente ocurre con el Factor 1, lo que confirma la estructura unidimensional de la escala.

En conjunto, los resultados presentados en las tablas y figuras anteriores aportan evidencia robusta sobre la calidad psicométrica de la EAEIA-GAD, confirmando que el instrumento es válido y confiable para evaluar las actitudes del estudiantado universitario hacia el uso de la inteligencia artificial generativa en actividades docentes.



**Figura 3.** Análisis paralelo de Horn: autovalores observados vs. umbral aleatorio (percentil 95). Solo el Factor 1 supera el umbral, respaldando la estructura unidimensional de la EAEIA-GAD.

## Discusión de resultados

Los análisis psicométricos confirmaron la solidez de la EAEIA-GAD. La consistencia interna, con un alfa de Cronbach ( $\alpha$ ) de 0.89 y un omega de McDonald ( $\omega$ ) de 0.80 —ambos por encima del mínimo recomendado de 0.70—, indica que el instrumento mide de forma confiable y coherente lo que pretende medir. En términos prácticos, si se aplicara el cuestionario en condiciones similares, se esperarían resultados consistentes. Se utilizaron ambos indicadores porque el omega es considerado más robusto en ciertas condiciones estadísticas [15]–[17]. En otros términos, estos resultados indican que las respuestas del estudiantado son consistentes entre sí.

La validez del instrumento —entendida como la evidencia de que mide adecuadamente el constructo que se propone— se examinó a partir de su estructura interna. En primer lugar, el índice KMO alcanzó un valor de 0.91, clasificado como excelente, lo que indica que los datos presentan condiciones adecuadas para identificar patrones comunes entre las respuestas del estudiantado. En segundo lugar, la prueba de esfericidad de Bartlett ( $\chi^2(45) = 794.65$ ,  $p < 0.001$ ) resultó significativa, lo que confirma que las respuestas a los distintos ítems están relacionadas entre sí de forma consistente y no por azar. En conjunto, estos resultados muestran que la información recolectada es adecuada para aplicar un análisis factorial, es decir, para estudiar cómo se organizan las respuestas del cuestionario como un todo coherente [18], [19].

Los resultados muestran que la escala presenta una estructura unidimensional: todos los ítems apuntan hacia un único concepto general, que es la actitud del estudiantado hacia la IA en actividades docentes. En otras palabras, las diez preguntas no miden aspectos independientes entre sí, sino que convergen en captar una misma predisposición general. Esto es esperable en contextos universitarios donde el uso de la IA es relativamente reciente, ya que el estudiantado tiende a valorarla de forma global —por ejemplo, en términos de utilidad o agrado— más que a diferenciar dimensiones específicas. Aunque la escala podría reflejar matices como motivación, utilidad o confianza, estos no se distinguen con suficiente claridad como para justificar subdivisiones estables del instrumento.

El análisis paralelo de Horn confirmó esta estructura de un solo factor: al comparar los resultados reales con valores generados al azar, únicamente el primer componente de variación (autovalor = 5.37) superó el umbral aleatorio correspondiente (1.33). En términos sencillos, esto indica que solo una dimensión explica de manera clara las respuestas del estudiantado. Este hallazgo respalda la existencia de una única dimensión subyacente y coincide con estudios que han validado escalas breves de actitudes hacia la IA con estructuras unidimensionales, como el AIAS-4 [22], lo que aporta evidencia adicional de la coherencia interna del instrumento.

En cuanto a las correlaciones ítem–total corregidas —indicador de qué tanto contribuye individualmente cada pregunta al puntaje total—, los valores oscilaron entre 0.29 y 0.78. La mayoría superó el umbral de 0.30 que se considera aceptable en psicometría [23], lo que significa que casi todas las preguntas aportan de manera suficiente al instrumento. La excepción fue el ítem “La actividad promovió la colaboración entre compañeros”, que obtuvo un valor de 0.29, ligeramente por debajo del umbral. Esto podría deberse a que en algunas actividades con IA el trabajo fue individual y no colaborativo, lo que haría que ese ítem midiera algo un poco diferente al resto. No obstante, esta debilidad no invalida el instrumento; simplemente sugiere que ese enunciado podría revisarse o contextualizarse mejor en versiones futuras.

En términos prácticos, los resultados ofrecen orientaciones concretas para el docente que desee aplicar la EAEIA-GAD. Un puntaje promedio elevado (por ejemplo, superior a 5.5 en la escala de 1 a 7) puede interpretarse como una señal de que el estudiantado recibe favorablemente las actividades mediadas por IA diseñadas por el docente, lo que valida pedagógicamente la propuesta instruccional. Por el contrario, puntuaciones bajas en ítems específicos —como los relacionados con claridad de instrucciones (ítem 2) o adecuación del nivel de dificultad (ítem 5)— señalan áreas concretas de ajuste. Adicionalmente, el ítem de colaboración (ítem 7) debe interpretarse con cautela en contextos de trabajo individual. En conjunto, la escala ofrece al docente información diagnóstica rápida y accionable para mejorar el diseño de sus actividades con IA en tiempo real, sin requerir formación especializada en psicometría.

Respecto al efecto techo observado —el hecho de que los promedios se concentraron entre 5.43 y 6.03 en una escala del 1 al 7— indica que la mayoría del estudiantado respondió con puntuaciones muy altas. Si bien esto refleja actitudes favorables hacia la IA, también limita la capacidad del instrumento para distinguir entre distintos grados de entusiasmo en el rango alto. Es como si en una competencia todos llegaran muy cerca del primer lugar: difícilmente se puede determinar quién ganó. Para afinar este aspecto en futuras versiones, se recomienda: (1) incluir preguntas que requieran evidencias concretas de práctica más allá del simple acuerdo general; (2) agregar ítems sobre aspectos como la fiabilidad de la IA o el riesgo de depender demasiado de ella, redactados de forma no punitiva; y (3) revisar las descripciones de los extremos de la escala para evitar que las personas respondan automáticamente con la opción más alta [24]–[26].

## Conclusiones, aportes prácticos y proyecciones futuras

Los resultados de este estudio aportan evidencia de validez y confiabilidad de la EAEIA-GAD como instrumento para evaluar las actitudes del estudiantado universitario hacia el uso de la inteligencia artificial en actividades docentes. En particular, los indicadores de consistencia interna —alfa de Cronbach (0.89) y omega de McDonald (0.80)— superan el umbral mínimo recomendado de 0.70, lo que evidencia un funcionamiento coherente del cuestionario. Asimismo, el análisis de la estructura interna respalda una organización unidimensional, es decir, que las preguntas del instrumento se articulan en torno a un único constructo general.

En conjunto, estos resultados sustentan la validez psicométrica inicial de la escala en el contexto analizado, en el sentido de que sus ítems se comportan de manera consistente y permiten interpretar las puntuaciones como una medida global de actitud hacia la inteligencia artificial en actividades docentes.

Desde el punto de vista práctico, la EAEIA-GAD constituye una herramienta breve y de fácil aplicación que puede apoyar la exploración de las percepciones estudiantiles sobre el uso de la inteligencia artificial en el aula, así como el seguimiento de su incorporación en contextos educativos universitarios.

En particular, su potencial impacto en la práctica educativa se manifiesta en tres planos. Primero, permite al docente tomar decisiones informadas sobre el diseño y la mejora continua de sus actividades con IA-G, a partir de evidencia cuantitativa sobre las percepciones del grupo. Segundo, facilita la comparación sistemática entre distintas cohortes, cursos o intervenciones pedagógicas, contribuyendo a una cultura de evaluación formativa del uso de la IA en el aula. Tercero, ofrece a los equipos académicos y a las instituciones un indicador estandarizado para dar seguimiento a la integración de estas tecnologías en sus planes de estudio, apoyando decisiones de política educativa fundamentadas en evidencia.

No obstante, se identifican aspectos susceptibles de mejora en futuras versiones del instrumento. En particular, el ítem relacionado con la colaboración entre estudiantes mostró un desempeño ligeramente inferior al esperado, y se observó un efecto techo en las respuestas, lo que sugiere la conveniencia de ajustar algunos enunciados para mejorar la capacidad de discriminación de la escala.

Como proyección, se recomienda aplicar la EAEIA-GAD en muestras más amplias y diversas, así como examinar su comportamiento en distintos grupos de estudiantes y contextos educativos. Asimismo, el uso de enfoques metodológicos complementarios podría aportar evidencia adicional sobre su funcionamiento psicométrico. Estas líneas de trabajo permitirán fortalecer su uso como herramienta para el estudio de la inteligencia artificial en la educación superior. En definitiva, la EAEIA-GAD representa un paso concreto hacia una integración más reflexiva y evaluada de la IA-G en la docencia universitaria: un instrumento que pone en manos del profesorado evidencia real sobre las actitudes de sus estudiantes, con el fin de orientar decisiones pedagógicas más pertinentes, equitativas y eficaces.

## Referencias

- [1] D. B. McLeod, "Research on affect in mathematics education: A reconceptualization," in *Handbook of Research on Mathematics Teaching and Learning*, D. A. Grouws, Ed. New York: Macmillan, 1992, pp. 575–596.
- [2] T. Teo, "Students and teachers' intention to use technology: Assessing their measurement equivalence and structural invariance," *Journal of Educational Computing Research*, vol. 57, no. 1, pp. 201–225, 2019.
- [3] O. Zawacki-Richter, V. Marín, T. Bond, and F. Gouverneur, "Systematic review of research on artificial intelligence applications in higher education," *International Journal of Educational Technology in Higher Education*, vol. 16, no. 1, pp. 1–27, 2019.
- [4] J. C. Sánchez-Prieto, S. Olmos-Migueláñez, and F. J. García-Peñalvo, "Informal tools in formal contexts: Development of a model to assess the acceptance of mobile technologies among teachers," *Computers in Human Behavior*, vol. 55, pp. 519–528, 2017.
- [5] A. Katsantonis and I. G. Katsantonis, "University students' attitudes toward artificial intelligence: An exploratory study of the cognitive, emotional, and behavioural dimensions of AI attitudes," *Education Sciences*, vol. 14, no. 9, p. 988, 2024.
- [6] B. G. Acosta-Enríquez, M. A. Arbulú Ballesteros, O. Huamaní Jordan, C. López Roca, and K. Saavedra Tirado, "Analysis of college students' attitudes toward the use of ChatGPT in their academic activities: Effect of intent to use, verification of information and responsible use," *BMC Psychology*, vol. 12, p. 255, 2024.
- [7] C. K. Y. Chan and W. Hu, "Students' voices on generative AI: Perceptions, benefits, and challenges in higher education," *International Journal of Educational Technology in Higher Education*, vol. 20, no. 1, p. 42, 2023.
- [8] A. Almassaad, H. Alajlan, and R. Alebaikan, "Student perceptions of generative artificial intelligence," *Systems*, vol. 12, no. 10, p. 385, 2024.
- [9] B. Zhang and A. Dafoe, "Artificial intelligence: American attitudes and trends," *SSRN Electronic Journal*, 2019.
- [10] C. Stöhr, A. W. Ou, and H. Malmström, "Perceptions and usage of AI chatbots among students in higher education across genders, academic levels and fields of study," *Computers and Education: Artificial Intelligence*, vol. 7, p. 100259, 2024.
- [11] R. Huang, H. Kim, and C. Springer, "Artificial intelligence in higher education: Trends and challenges," *Computers and Education*, vol. 205, p. 104809, 2024.
- [12] W. C. Morales-García, L. Z. Sairitupa-Sanchez, S. B. Morales-García, and M. Morales-García, "Development and validation of a scale for dependence on artificial intelligence in university students," *Frontiers in Education*, vol. 9, p. 1323898, 2024.
- [13] A. Schepman and P. Rodway, "Initial validation of the General Attitudes towards Artificial Intelligence Scale (GAAIS)," *Computers in Human Behavior Reports*, vol. 1, p. 100014, 2020.
- [14] M. A. Robinson, "Using multi-item psychometric scales for research: A review and practical guide," *Human Resource Management*, vol. 57, no. 3, pp. 739–750, 2018.
- [15] A. F. Hayes and J. J. Coutts, "Use omega rather than Cronbach's alpha for estimating reliability," *Communication Methods and Measures*, vol. 14, no. 1, pp. 1–26, 2020.
- [16] T. J. Dunn, T. Baguley, and V. Brunnsden, "From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation," *British Journal of Psychology*, vol. 105, no. 3, pp. 399–412, 2014.
- [17] L. R. Nelson, "Computing alpha and omega reliability estimates (working paper)," Curtin University, 2024.

- [18] H. F. Kaiser, "An index of factorial simplicity," *Psychometrika*, vol. 39, no. 1, pp. 31–36, 1974.
- [19] M. S. Bartlett, "Tests of significance in factor analysis," *British Journal of Statistical Psychology*, vol. 3, no. 2, pp. 77–85, 1950.
- [20] J. L. Horn, "A rationale and test for the number of factors in factor analysis," *Psychometrika*, vol. 30, no. 2, pp. 179–185, 1965.
- [21] IBM Corp., *IBM SPSS Statistics for Windows, Version 29.0*. Armonk, NY: IBM Corp., 2022.
- [22] S. Grassini, "Development and validation of the AI attitude scale (AIAS-4): A brief measure of general attitude toward artificial intelligence," *Frontiers in Psychology*, vol. 14, p. 1191628, 2023.
- [23] R. F. DeVellis, *Scale Development: Theory and Applications*, 4th ed. Thousand Oaks, CA: SAGE Publications, 2016.
- [24] M. J. Allen and W. M. Yen, *Introduction to Measurement Theory*. Long Grove, IL: Waveland Press, 2002.
- [25] S. E. Embretson and S. P. Reise, *Item Response Theory for Psychologists*. Mahwah, NJ: Lawrence Erlbaum Associates, 2000.
- [26] N. J. Salkind, *Encyclopedia of Research Design*. Thousand Oaks, CA: SAGE, 2010.

## Biografías de los autores

### **Luis Gerardo Meza-Cascante**

Doctor en Educación con énfasis en Investigación Educativa y profesor catedrático de la Escuela de Matemática del Instituto Tecnológico de Costa Rica (ITCR). Ha dedicado su carrera académica a la investigación en didáctica de la matemática, evaluación educativa y, más recientemente, a la integración pedagógica de herramientas de inteligencia artificial en la educación superior. Es autor de diversas publicaciones nacionales e internacionales en estas áreas. Correo: [gemeza@itcr.ac.cr](mailto:gemeza@itcr.ac.cr)

### **Melvin Ramírez-Bogantes**

Docente e investigador de la Escuela de Matemática del Instituto Tecnológico de Costa Rica (ITCR) y doctorando en Educación. Actualmente su labor académica e investigativa se centra en la educación matemática, la modelización matemática y la integración de la inteligencia artificial en la enseñanza universitaria. Ha participado en congresos nacionales e internacionales y es autor de artículos académicos en temas vinculados con la enseñanza y el aprendizaje de la matemática. Correo: [meramirez@itcr.ac.cr](mailto:meramirez@itcr.ac.cr)

### **Leonel Chaves-Salas**

Docente e investigador de la Escuela de Matemática del Instituto Tecnológico de Costa Rica (ITCR) y doctorando en Educación. Actualmente su labor académica e investigativa se centra en la educación matemática, la integración de la inteligencia artificial en la enseñanza universitaria y la matemática recreativa como metodología de enseñanza. Ha participado en congresos nacionales e internacionales y es autor de artículos académicos en temas vinculados con la enseñanza y el aprendizaje de la matemática. Correo: [leochaves@itcr.ac.cr](mailto:leochaves@itcr.ac.cr)