Artículo de revisión

Gerardo Antonio Hernández-Torres[1]
https://orcid.org/0000-0001-5498-8358

Silvia Patricia Aquino Zúñiga[2*]
https://orcid.org/0000-0002-7223-8582

Manuel Jesús Izquierdo Sandoval[3]
https://orcid.org/0000-0001-5605-2318

[1] Universidad Juárez Autónoma de Tabasco, Dirección del Sistema Bibliotecario; Orcid: 0000-0001-54988358; gerardo.hernandez@ujat.mx

[2] Universidad Juárez Autónoma de Tabasco, División Académica de Educación y Artes, Orcid: 0000-0002-7223-8582; saquinozuniga@gmail.com

[3] Universidad Juárez Autónoma de Tabasco, División Académica de Educación y Artes; Orcid: 0000-00015605-2318 (QEPD).

*** Autor de correspondencia:**
Silvia Patricia Aquino Zúñiga, Universidad Juárez Autónoma de Tabasco, División Académica de Educación y Artes; Orcid: 0000-00015498-8358; Calle Revolución 1350 Edif. A Depto. 4, Col 18 de marzo, Villahermosa, México; saquinozuniga@gmail.com

Para citar este artículo:
Hernández-Torres, G., Aquino Zúñiga, S., e Izquierdo Sandoval M. (2026). Medición de las percepciones sobre el uso de la inteligencia artificial en la educación superior: propuesta de marco de adaptación para el contexto latinoamericano. *Papeles, 18*(35), e2317. https://doi.org/10.54104/papeles.v18n35.2317

Medición de las percepciones sobre el uso de la inteligencia artificial en la educación superior: propuesta de marco de adaptación para el contextolatinoamericano

Measuring Perceptions of Artificial Intelligence use in Higher Education: A Critical Review and Adaptation Framework for the Latin American Context

https://doi.org/10.54104/papeles.v18n35.2317

Recibido: 25 de octubre de 2025
Aprobado: 23 de enero de 2026
Publicado: 15 de mayo de 2026



Resumen

Introducción: la adopción de herramientas de inteligencia artificial en la educación superior ha motivado un creciente interés por comprender los factores que influyen en su aceptación y por evaluar la calidad de los instrumentos utilizados para su medición. En este contexto, resulta necesario analizar de manera crítica los modelos teóricos y las evidencias psicométricas que sustentan dichos instrumentos. **Metodología:** se realizó una revisión sistemática siguiendo las directrices Preferred Reporting Items for Systematic reviews and Meta-Analyses (Prisma), a partir de la cual se analizaron 26 estudios empíricos publicados entre 2022 y 2025 e indexados en Scopus, Web of Science (WoS), Scientific Electronic Library Online (SciELO) y Red de Revistas Científicas de América Latina y el Caribe, España y Portugal (Redalyc). Los estudios seleccionados emplean mayoritariamente cuestionarios con escalas tipo Likert para la medición de los constructos propuestos por el Technology Acceptance Model (TAM) y la Unified Theory of Acceptance and Use of Technology 2 (UTAUT2) que reportan información sobre procesos de validación y confiabilidad. **Resultados y discusión:** los resultados muestran una utilización consistente de validez de constructo y confiabilidad interna como criterios mínimos de evaluación instrumental, aunque con variaciones relevantes en los procedimientos metodológicos y en la adaptación contextual de los modelos. Asimismo, se identifican desafíos asociados a la equidad digital y a la confianza institucional que inciden en la aceptación de la inteligencia artificial (IA), en particular, en contextos latinoamericanos. A partir de estos hallazgos, se proponen lineamientos orientados a mejorar la evaluación de la aceptación de la IA en la educación superior latinoamericana. **Conclusiones:** la revisión sugiere que los cuestionarios basados en escalas Likert constituyen instrumentos ampliamente utilizados y potencialmente replicables; sin embargo, su aplicación en la educación superior latinoamericana requiere adaptaciones contextuales y normativas que garanticen la validez de los constructos medidos. En este sentido, se destaca la necesidad de avanzar hacia diseños instrumentales más sensibles a las condiciones estructurales e institucionales de la región.

Palabras clave:
Inteligencia artificial; percepción; enseñanza superior; cuestionario; medición

Abstract

Introduction: The adoption of artificial intelligence tools in higher education has generated growing interest in understanding the factors that influence their acceptance, as well as in assessing the quality of the instruments used for their measurement. In this context, it is necessary to critically examine the theoretical models and the psychometric evidence supporting these instruments. **Methodology:**

Keywords:
Artificial intelligence; perception; higher education; questionnaires; measurement

A systematic review was conducted following the Prisma guidelines, through which 26 empirical studies published between 2022 and 2025 and indexed in Scopus, Web of Science, SciELO and Redalyc were analyzed. The selected studies predominantly employ questionnaires with Likert-type scales to measure the constructs proposed by the TAM and UTAUT2 models and report information on instrument validation and reliability procedures. **Results and Discussion:** The findings reveal a consistent use of construct validity and internal reliability as minimum criteria for instrumental evaluation, although relevant variations are observed in methodological procedures and in the contextual adaptation of the models. In addition, challenges related to digital equity and institutional trust are identified as factors influencing the acceptance of artificial intelligence, particularly in Latin American contexts. Based on these findings, guidelines are proposed to improve the assessment of artificial intelligence acceptance in Latin American higher education. **Conclusions:** The review suggests that Likert-scale–based questionnaires are widely used and potentially replicable instruments; however, their application in Latin American higher education requires contextual and regulatory adaptations to ensure the validity of the measured constructs. In this regard, there is a clear need to advance toward instrumental designs that are more sensitive to the structural and institutional conditions of the region.

1. Introducción

A nivel global, las instituciones de educación superior han adaptado sus modelos para responder a las demandas sociales y del mercado laboral, integrando herramientas como la inteligencia artificial (IA) en la enseñanza y en los procesos de investigación. La integración de la IA en la educación superior ha demostrado ser un catalizador para transformar las dinámicas de enseñanza y aprendizaje. Los estudios recientes destacan el potencial de herramientas como ChatGPT para ofrecer experiencias educativas personalizadas y mejorar las competencias académicas de los estudiantes, en especial ,en escritura y aprendizaje adaptativo (Ayoubi, 2024; Mahmud et al., 2024). Sin embargo, esta adopción enfrenta desafíos, como preocupaciones sobre la privacidad, calidad de los contenidos generados y falta de integración curricular, en particular, en áreas como la educación médica (Li y Qin, 2023). A pesar de estos retos, la IA se posiciona como una herramienta clave para fomentar métodos educativos más efectivos en la era pospandemia, alineándose con tendencias globales hacia la innovación pedagógica (Alhwaiti, 2023; Saihi et al., 2024). Estas tecnologías emergentes abrirán nuevas oportunidades de estudio y empleo, e impactarán no solo a los profesionales en tecnología, sino también todas las disciplinas del conocimiento (Universidad Juárez Autónoma de Tabasco, 2024).

Es fundamental entender el uso y las percepciones de los actores educativos sobre la IA en la educación superior, para comprender su efecto e identificar fortalezas y áreas de oportunidad que permitan sugerir estrategias de implementación o lineamientos de uso adecuado, y así optimizar los beneficios de

esta tecnología y lograr una adecuada implementación en la educación superior.

La investigación de fenómenos sociales es abordada desde el enfoque cuantitativo a través de aplicación de principios de las ciencias naturales (Bryman, 1984; Shaker y Ruitenberg, 2007), que, por medio de la observación empírica y la verificación, comprueban o generan leyes y teorías que permiten la comprensión general del fenómeno (Cohen et al., 2018; Lodico et al., 2006). Hoy y Adams (2016) mencionan que una investigación en educación debe ser teórica, empírica, controlada y replicable; para ello, es necesario conocer lo que piensan los individuos sobre aspectos concretos de la educación, así como la expectativa, que se refiere a las ideas que los estudiantes preconciben sobre eventos futuros y que influyen en sus actitudes (Merhi, 2011).

Por medio de la percepción, se construye una creencia u opinión basada en las experiencias (Persson et al., 2022), de ahí su importancia para comprender los comportamientos hacia los cambios en el entorno educativo. Una herramienta que posibilita la obtención de este tipo de subjetividad del estudiante son los cuestionarios de escalas sociométricas, las cuales, según Fabila Echauri et al. (2014), permiten la medición de preferencias, actitudes, opiniones, etc., asignando a las respuestas un valor numérico que permite codificar la información recabada y cuantificar las respuestas obtenidas. En este sentido, la escala de Likert permite la representación de posición favorable o desfavorable del estudiante con respecto al fenómeno que se estudia.

El objetivo principal de este artículo es analizar la construcción de instrumentos que miden las percepciones sobre el uso de la IA, por medio de un cuestionario de escala Likert, de los estudiantes y profesores en educación superior, para comprender el proceso de elaboración y validación de estos, con la finalidad de generar recomendaciones que permitan la preparación de instrumentos adaptados al contexto latinoamericano, retomando las teorías y los modelos utilizados internacionalmente.

En este contexto, comprender las percepciones del uso de la IA en el nivel terciario se vuelve fundamental para orientar decisiones pedagógicas, normativas institucionales y de política educativa. No obstante, aunque hay evidencia de estudios del contexto internacional en el que se han diseñado y validado instrumentos cuantitativos para medir estas percepciones, su aplicación directa en entornos latinoamericanos plantea limitaciones relacionadas con la validez ecológica, las diferencias culturales y las brechas lingüísticas. De ahí surge la necesidad de realizar un análisis crítico de instrumentos que miden la percepción sobre el uso de IA en la educación superior que permita identificar sus fundamentos teóricos, tener claridad sobre las dimensiones y características de la población y su contexto, además de localizar limitaciones de los modelos más utilizados, sobre todo, al adaptarlos a la educación superior latinoamericana. En las siguientes secciones, se presentan los fundamentos teóricos, el método de la revisión sistemática, la discusión y el análisis de resultados, para la propuesta de lineamientos de diseño y conclusiones.

Por medio de la percepción, se construye una creencia u opinión basada en las experiencias, de ahí su importancia para comprender los comportamientos hacia los cambios en el entorno educativo.

1.1 Referentes teóricos sobre el uso de la IA

En América Latina, se han desarrollado estudios orientados a examinar el uso y la percepción de la IA en la educación superior, en particular, en países como México, Perú, Ecuador, Bolivia y Brasil. Por ejemplo, investigaciones como la de Acosta-Enríquez et al. (2024), desarrollada en Perú, analizan las actitudes de estudiantes universitarios hacia el uso de ChatGPT mediante un enfoque cuantitativo, integrando constructos provenientes de modelos consolidados, como Unified Theory of Acceptance and Use of Technology 2 (UTAUT2) y General Attitudes towards Artificial Intelligence Scale (GAAIS), además de dimensiones emergentes, como la verificación de información y el uso responsable de la tecnología. De manera similar, Cabero-Almenara et al. (2024) examinan la aceptación de la IA en la educación por parte de los docentes ecuatorianos y su relación con dimensiones, como expectativas de esfuerzo, motivación hedónica y creencias pedagógicas, tomando como base el modelo UTAUT2.

Otros estudios realizados en la región, sin embargo, adoptan enfoques de carácter exploratorio o aplicado, que, si bien aportan información valiosa desde una perspectiva contextual, no explicitan un marco teórico integrado para la construcción de sus instrumentos de medición o presentan limitaciones en la descripción de los procesos de validación psicométrica. Tal es el caso de investigaciones orientadas a la personalización del aprendizaje mediante modelos de aprendizaje automático en Ecuador, donde los cuestionarios se basan en modelos descriptivos, como Visual, Aural, Read/write, and Kinesthetic (VARK), sin una vinculación clara con teorías de aceptación tecnológica (Villegas-Ch et al., 2024). De igual forma, algunos estudios desarrollados en México emplean cuestionarios *ad hoc* para explorar percepciones generales sobre la IA, sin declarar modelos teóricos que orienten su diseño (Ruiz Mendoza et al., 2024).



Esta diversidad metodológica no evidencia una ausencia de producción científica en la región, sino un vacío recurrente en la estandarización teórico-metodológica de los instrumentos de medición, lo que refuerza la pertinencia de analizar críticamente aquellos cuestionarios fundamentados en modelos consolidados (como Technology Acceptance Model [TAM] y UTAUT2) y de proponer lineamientos para su adaptación contextual al ámbito latinoamericano.

Actualmente, existe un gran desarrollo de investigaciones relacionadas con la percepción sobre el uso de la IA en la educación superior, utilizando mayormente los cuestionarios de escala como instrumento para medir el grado de aceptación de esta tecnología (tabla 1), cuya elaboración es fundamentada en teorías establecidas o en el desarrollo de modelos unificados, lo que permite una implementación rápida y con la flexibilidad de integración en distintos contextos. Aunque la literatura sobre el uso de estos instrumentos esté en aumento, existe poca literatura orientada al análisis crítico de la validación y adaptabilidad a contextos regionales, lo que posibilita la generación de sesgos o carencias en la recopilación de datos.

Tabla 1. *Tipos de instrumentos que abordan la percepción de IA en la educación superior*

Instrumento	Variables / dimensiones	Número de ítems/ preguntas	Contexto	Participantes
Encuesta de escala Likert	Especialmente en torno a la aceptación tecnológica, actitudes, utilidad percibida, riesgo y confianza.	En promedio, se tienen instrumentos con 31 ítems, entre ellos, instrumentos con una estructura de 11 ítems como mínimo y 51 ítems como máximo.	Diversidad geográfica amplia, que abarca países de América, Europa, Asia, África y Oceanía, además de estudios multinacionales.	Principalmente estudiantes y profesores universitarios, además de consumidores y usuarios en línea, editores y profesionales de la salud.
Entrevista	Factores de adopción y uso de IA/ChatGPT, percepciones éticas, impacto en enseñanza, regulación cognitiva y motivacional.	En general, los estudios no describen las preguntas, pero se localizan investigaciones que se guían desde 3 hasta 11 preguntas.	India, Omán, Malasia, Corea del Sur, China, Sudáfrica, Pakistán, Arabia Saudita y Bangladés.	Profesores y estudiantes universitarios, investigadores y profesionales de la salud.
Grupo focal	Percepción sobre chatbots, videos con IA, motivación, rendimiento y experiencias colectivas	Las investigaciones mencionan el uso de grupos focales, con tres preguntas iniciales.	Reino Unido, Ghana, Turquía y Ucrania.	Profesores y estudiantes universitarios.
Otros (encuestas abiertas, debates, análisis de contenido, informes escritos, etc.)	Percepciones y preocupaciones sobre IA, adopción de herramientas de IA, grado de aceptación y brechas de uso.	Se manejan distintos tipos de procedimientos metodológicos, según el tipo de instrumento.	España, Austria, Ecuador, India, Arabia Saudita, Alemania y Polonia.	Estudiantes universitarios.

Fuente: elaboración propia.

Para conocer las percepciones de los profesores y estudiantes, los estudios utilizan teorías que explican la relación entre factores que determinan la aceptación y el uso de tecnologías, algunas enfocadas en tecnologías educativas, otras específicas para el uso de la IA o herramientas particulares, como ChatGPT. Cada una de estas permite el acercamiento a los constructos que determinan cómo los actores educativos perciben el proceso de integración de la IA en la educación superior, por ejemplo, el factor hedónico, la facilidad de uso, la sensación de utilidad, etc. Estas variables relacionan aspectos que pueden influir en el uso de las personas al implementarse alguna nueva tecnología en sus entornos, lo que determina el nivel de aceptación o rechazo que puede tener la implementación.

Estas variables han sido utilizadas en diversas teorías y modelos de aceptación tecnológica, como la UTAUT2, el TAM o el Technology Acceptance Model Edited to Assess ChatGPT Adoption (TAME-ChatGPT), que han servido de base para numerosos estudios relacionados con aceptación de IA.

La UTAUT2 explica la aceptación y el uso de tecnologías a partir de un conjunto de constructos que articulan dimensiones instrumentales, contextuales y motivacionales. Entre ellos, se destacan la expectativa de desempeño y de esfuerzo, que remiten a la percepción de utilidad y facilidad de uso de la tecnología en relación con el logro de tareas académicas. Asimismo, las condiciones facilitadoras se refieren a la disponibilidad de recursos técnicos, infraestructura y apoyo

institucional que hacen viable su implementación. A estas dimensiones se suman factores de carácter experiencial, como la motivación hedónica, vinculada al disfrute derivado del uso tecnológico, y el hábito, entendido como la internalización del uso a partir de experiencias previas y patrones recurrentes de interacción (Venkatesh et al., 2012).

El TAM aborda la adopción de tecnologías a partir de la relación entre la utilidad percibida y la facilidad de uso percibida, entendidas como creencias centrales que influyen en las actitudes y en la intención de uso. En el ámbito educativo, estos constructos permiten analizar hasta qué punto una tecnología es valorada por su contribución al desempeño académico y por la reducción del esfuerzo cognitivo requerido para su uso (Davis et al., 1989). Su amplia validación empírica ha favorecido su aplicación en diversos contextos educativos y su extensión mediante la incorporación de variables adicionales que fortalecen su capacidad explicativa frente a tecnologías emergentes, como la IA.

El TAME-ChatGPT constituye una extensión del TAM orientada a comprender la aceptación de modelos lingüísticos de gran escala en contextos educativos. Además de los constructos clásicos, incorpora dimensiones asociadas al riesgo percibido y a la ansiedad tecnológica, que reflejan preocupaciones vinculadas a la fiabilidad de la información, la integridad académica y la dependencia cognitiva (Sallam et al., 2023). Esta ampliación resulta relevante para el análisis de herramientas como ChatGPT, cuyo uso plantea tensiones específicas entre innovación pedagógica, ética académica y regulación institucional.

Cada una de estas teorías con sus respectivos constructos permite una generalización local, regional o global de los factores que limitan o potencian las innovaciones tecnológicas, así como determinan si la tecnología logrará su integración en el entorno estudiado o



mostrará un panorama de posibles adaptaciones contextuales para la implementación adecuada de tecnología.

La percepción y los factores que determinan la aceptación de la IA en la educación superior se han estudiado desde objetivos diversos, la mayoría tratando el uso de herramientas IA (de manera global), seguido de estudios que abordan específicamente la herramienta ChatGPT, y algunos más que se enfocan en otros chatbots, además de estudios que se centran en revisar el aprendizaje con IA, rendimiento académico, malas prácticas, evaluación, retroalimentación y solución de exámenes con IA. Todos estos buscan comprender cómo los estudiantes y profesores asimilan la integración de herramientas de IA en sus actividades académicas, contrastando los factores que determinan su uso y rechazo.

Ejemplo de ello lo describe el estudio de Cabero-Almenara et al. (2024) en el que exploran la aceptación de la IA en la educación por parte de los profesores y encuentran que la intención de uso es un determinante para el comportamiento de uso. Además, la edad y el área de conocimiento puede influenciarla, mostrando en los profesores más jóvenes un menor esfuerzo y una mayor

satisfacción en el empleo de la IA, o en profesores del área de ciencias sociales una mayor disposición de utilización.

En la medicina, también se está integrando la IA. Al respecto, Li y Qin (2023) exploran la percepción de estudiantes de pregrado y posgrado en áreas médicas y encuentran que los estudiantes de posgrado tienen mayor uso ético de la IA, en que factores como la expectativa de rendimiento, el hábito, la motivación hedónica y la confianza influyen positivamente en la intensión de uso. En cambio, Segarra Ciprés et al. (2024) analizan la integración de ChatGPT en asignaturas de ingeniería informática, en las que muestran cómo profesores y estudiantes la aceptan y utilizan como recurso formativo, así como evidencian que la mayoría de los estudiantes consideran la herramienta como útil y fácil de usar, no así los profesores, en quienes su aceptación y uso es en menor proporción, sobre todo, por la baja percepción de seguridad y el miedo a los errores.

Para medir y comprender estos factores, es necesario la construcción de instrumentos adecuados de recolección de datos. En el caso de la integración de IA en la educación superior, el cuestionario de escala Likert ha permitido obtener las percepciones de los estudiantes y profesores, pero es necesario conocer los elementos y marcos teóricos que fundamentan su elaboración. En la tabla 2, se identifican los criterios que deben considerarse para un completo abordaje teóricometodológico en el desarrollo de cuestionarios de escala.

Tabla 2. *Elementos teórico-metodológicos de los cuestionarios de escala Likert*

Criterio/elemento	Explicación y fundamentación
Naturaleza de la escala	Debe considerar la propiedad ordinal de las variables, en que cada nivel de opinión se representa numéricamente y permite un posicionamiento graduado de lo negativo a lo positivo (Fabila Echauri et al., 2014; Field, 2018; Hodge y Gillespie, 2007).
Cantidad de niveles de opinión	Puede usarse una escala con valores pares (4, 6) que obliga a tomar una posición positiva o negativa, o nones (5, 7) con punto neutral. En el caso de percepciones docentes, se recomienda escala par para evitar neutralidad (Cohen et al., 2018; Fabila Echauri et al., 2014; Hodge y Gillespie, 2007).
Unidimensionalidad	Cada ítem debe medir un solo aspecto de la dimensión, evitando mezclar elementos distintos para asegurar claridad y precisión (Cohen et al., 2018).
Univocidad	Cada ítem debe incitar la opinión sobre un aspecto específico y no sobre dos o más elementos, y evitar ambigüedad (Fabila Echauri et al., 2014; Hodge y Gillespie, 2007).
Validación del instrumento	Es necesario verificar que mida constructos y dimensiones planteados. Incluye validación de constructo mediante análisis factorial, o validación de contenido y ecológica si los ítems fueron fundamentados *a priori* (Cohen et al., 2018; Cortina, 1993; Fabila Echauri et al., 2014; Field, 2018).
Validación de contenido	Consiste en la revisión por jueces expertos para confirmar la pertinencia de los ítems respecto de las dimensiones del constructo (Boyle y Fisher, 2008; Cohen et al., 2018; Fabila Echauri et al., 2014).
Validación ecológica	Se prueba con sujetos de perfiles similares a la población meta, quienes confirman la pertinencia de los ítems en el contexto real (Cohen et al., 2018).
Confiabilidad del instrumento	Verifica estabilidad y consistencia interna de los datos recolectados. El procedimiento más usado es el alfa de Cronbach, en consideración a valores superiores a 0,80 como aceptables (Boyle y Fisher, 2008; Cohen et al., 2018; Cortina, 1993; Field, 2018).
Análisis por dimensión	Cuando los ítems fueron construidos *a priori* según unidimensionalidad y univocidad, el alfa de Cronbach debe calcularse por dimensión para asegurar consistencia interna (Cortina, 1993; Field, 2018).

Fuente: elaboración propia.

Para la investigación educativa, es fundamental abordar vacíos de conocimiento, ya que esto determina la importancia de su contribución (Scates y Hoban, 1937), por lo que una revisión adecuada de literatura es la base para todas las etapas del proceso de investigación, estableciendo un marco teórico, identificando temas críticos, ajustando la metodología, etc., para darle credibilidad y validez (Cohen et al., 2018).

En general, los estudios elegidos surgen de una revisión de literatura que guía su proceso investigativo; en algunos, la revisión compara diferentes teorías que buscan obtener esas percepciones sobre el uso de la IA (Cabero-Almenara et al., 2024); en otros, la revisión se enfoca en conocer investigaciones que han utilizado el modelo teórico para comprender cómo fue implementado y qué resultados obtuvieron (Li y Qin, 2023); finalmente, algunos estudian los factores asociados a los modelos teóricos, en busca de similitudes o diferencias en cuanto a forma de estudiarse, asociaciones y resultados (Falebita y Kok, 2024). Esto permite construir un marco teórico conceptual que les ayuda a realizar su proceso de investigación, en el que surgen las preguntas de investigación, los objetivos del estudio, las hipótesis y las variables.

Las preguntas de investigación derivadas del conocimiento de teorías, estudios y avances actuales del problema a investigar guían el diseño de investigación, clarificando el tema, los objetivos, el alcance y sus límites (Barroga y Matanguihan, 2022; Köhler et al., 2017; White, 2013). De la elección de artículos, solo siete generan preguntas de investigación para sus estudios, lo que podría sugerir una debilidad en su diseño de investigación. Fundamentalmente, las preguntas de investigación de los estudios están enfocadas en conocer la relación o influencia que tiene una variable con respecto a la intención de usar IA, o bien de manera específica, por ejemplo, Kanont et al. (2024) propone la pregunta "¿Cuáles son las relaciones entre beneficio esperado, utilidad percibida y facilidad de uso percibida de la IA generativa y la intención conductual, y cómo afectan estas relaciones al uso real de la IA generativa en la educación superior?" (p. 19), o bien en general, englobando diversos factores, como Strzelecki et al. (2024), que sugieren la pregunta "¿Qué factores afectan la intención de los académicos de utilizar una herramienta de inteligencia artificial como ChatGPT?" (p. 3). Cabe mencionar que estos artículos también contemplan objetivos generales en su diseño de investigación.

Para la investigación educativa, es fundamental abordar vacíos de conocimiento, ya que esto determina la importancia de su contribución

Los estudios analizados buscan comprobar hipótesis que desarrollan en su diseño, a la espera de encontrar relaciones lógicas entre las variables que permitan comprender el problema investigado y obtener resultados fiables (Barroga y Matanguihan, 2022). Estas hipótesis se alinean con los modelos teóricos presentados en sus estudios; en relación con los trabajos que utilizan la UTAUT2 como base para su diseño de investigación, la mencionan como un marco integral de diversas teorías de aceptación de tecnología, que permite una comprensión holística de los factores que influyen en la adopción de tecnología (Amer Jid Almahri et al., 2024); para el TAM, Cambra-Fierro et al. (2024) y Falebita y Kok (2024) aclaran que el modelo es un marco ampliamente reconocido para comprender y predecir la aceptación y el uso de la tecnología por parte de los usuarios, ya que ha sido utilizado y validado en diferentes investigaciones.

Estos modelos teóricos buscan identificar los factores que motivan u obstruyen la aceptación de tecnologías en diferentes contextos, lo que representa un fenómeno social. Como mencionan Scott (2007) y Zyphur y Pierides (2020), es una construcción de diferentes realidades o propiedades distintas de los objetos sociales, que, a través del tratamiento teórico y empírico, puede representarse en una explicación matemática o estadística. Al respecto, Bryman (1984) señala que la metodología cuantitativa representa un enfoque para las investigaciones sociales, para lo cual una encuesta social es considerada el instrumento de investigación más apropiado, ya que mantiene la objetividad, la replicabilidad y la causalidad. De acuerdo con esto, se observa cómo la mayoría de las investigaciones emplean teorías y modelos que utilizan como base metodológica los cuestionarios de escala Likert, ya que permiten, por su fundamento epistemológico, la representación estadística de la posición favorable o desfavorable de la aceptación tecnológica (Fabila Echauri et al., 2014).

2. Metodología

A través de una revisión sistemática, que, según Page et al. (2021), permite sintetizar el conocimiento existente de un área para señalar problemas y evaluar teorías sobre fenómenos de interés, se identificaron 26 estudios relevantes sobre el impacto del uso de la IA en el rendimiento académico y las percepciones de estudiantes y profesores sobre su implementación en nivel de pregrado o posgrado. Para asegurar la fiabilidad y aplicabilidad, esta revisión sistemática se rigió por las directrices de elementos preferidos para informes de revisiones sistemáticas y metaanálisis (Preferred Reporting Items for Systematic reviews and Meta-Analyses [Prisma]), mediante su lista de verificación para limitar el sesgo en la selección de artículos hacia la comprensión de tendencias en la investigación sobre IA y educación superior.

Las bases de datos consultadas para la recopilación de los documentos fueron Scopus, Web of Science (WoS), Scientific Electronic Library Online (SciELO) y Red de Revistas Científicas de América Latina y el Caribe, España y Portugal (Redalyc). Como mencionan Zhu y Liu (2020), WoS y Scopus son dos bases de datos líderes a nivel mundial, donde se localizan revistas académicas de alto impacto en diversas áreas del conocimiento; las consultas en estas bases se realizaron en julio de 2025. Además, para reducir el sesgo de cobertura y fortalecer la pertinencia latinoamericana, se incorporaron SciELO y Redalyc como bases de datos complementarias, aplicando los mismos criterios de elegibilidad y manteniendo la trazabilidad del proceso mediante Prisma.

La selección de estas bases de datos respondió a la necesidad de equilibrar cobertura internacional y regional, manteniendo estándares comparables de reporte metodológico y psicométrico para facilitar la comparación teórica y la evaluación de los procesos de construcción, validación y confiabilidad de los instrumentos analizados conforme a las directrices Prisma. Dado que los motores de búsqueda de SciELO y Redalyc presentan diferencias en el tratamiento de operadores booleanos frente a WoS y Scopus, la estrategia de búsqueda se adaptó al formato de cada plataforma sin alterar los bloques conceptuales de la pregunta de investigación.

La estrategia de búsqueda se elaboró con la combinación de operadores booleanos (AND, OR) que permitieron la integración de todas las palabras clave definidas y obtener la siguiente ecuación de búsqueda:

> (undergraduate OR postgraduate OR "higher education" OR university) AND ("artificial intelligence" OR AI

> OR ChatGPT) AND ("academic performance" OR adoption) AND (perception OR acceptance OR challenges OR benefits OR impact OR uses)

Para SciELO, se utilizó la ecuación:

> ("inteligencia artificial" OR "artificial intelligence" OR ChatGPT)
>
> AND ("educación superior" OR universidad OR posgrado OR pregrado OR "higher education" OR university
>
> OR postgraduate OR undergraduate) AND (adopción OR adoption OR percepción OR perception OR aceptación
>
> OR acceptance OR uso OR uses OR impacto OR impact OR beneficios OR benefits OR desafíos OR challenges)

En Redalyc, se empleó la ecuación:

> "inteligencia artificial" AND "educación superior" AND (adopción OR aceptación OR percepción OR uso OR "intención")

Con esta estrategia, se optimizó la exploración de documentos en las bases de datos seleccionadas, aplicándose la búsqueda en títulos, resumen y palabras clave de los artículos, tras lo cual se obtuvo un total de 1314 documentos, de los cuales 519 son de WoS y 795 de Scopus. En la búsqueda complementaria en bases de datos regionales, en SciELO se recuperaron 300 registros, mientras en Redalyc se localizaron 875 documentos. En conjunto de todas las bases de datos, se obtuvieron 2489 registros.

Para la selección efectiva de los artículos encontrados, se consideraron diferentes criterios de exclusión e inclusión, con la finalidad de mantener la calidad y relevancia de los documentos en relación con el tema. Los criterios de inclusión y exclusión utilizados se detallan en la tabla 3.

Los artículos se tomaron de 2022 a 2025 en consideración a que, a partir de 2022, la IA tuvo un aumento en su exposición, sobre todo, con la IA generativa, específicamente con ChatGPT, creada en noviembre de 2022 (OpenAI, 2022), que impactó las diferentes áreas del conocimiento, sobre todo, la educación. Además, se considera la revisión solo de artículos académicos que tengan una relación directa con las percepciones, el uso, el impacto o las barreras de la integración de la IA en la educación superior, ya que nos interesa conocer cómo se está abordando la integración y qué piensan los actores educativos al respecto. La figura 1 detalla el procedimiento de sistematización de documentos con la metodología Prisma.

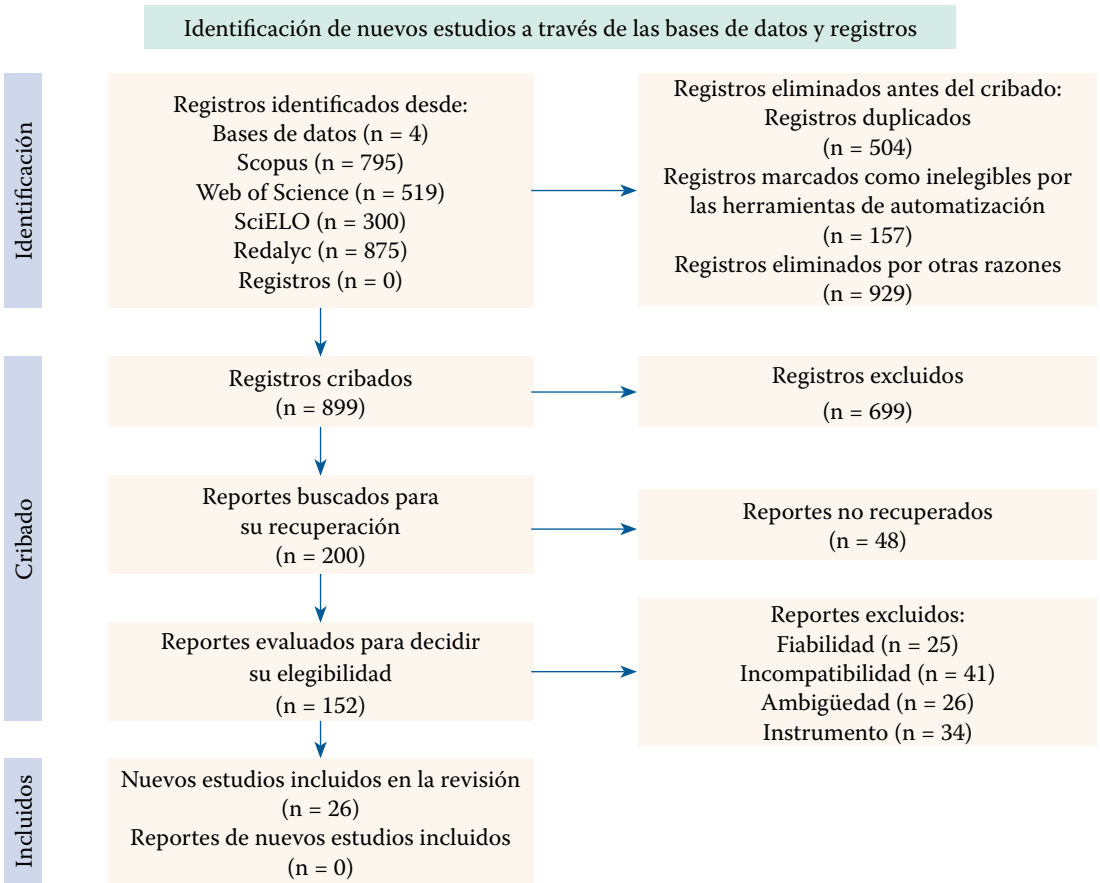
Se analizaron los documentos para determinar si cumplían con los requisitos del estudio y verificar si contenían el instrumento y su proceso de construcción. En la

Tabla 3. *Criterios de inclusión y exclusión utilizados para la selección de los artículos*

Criterio de inclusión	Criterio de exclusión
Artículos publicados entre 2022 y 2025	Artículos fuera de los años comprendidos en el estudio
Artículos académicos	Revisiones sistemáticas, libros, capítulos de libros o tesis
Disponibilidad en inglés o español	Idiomas diferentes del inglés o español
Acceso a texto completo	Sin acceso a texto completo
Artículos relacionados con la percepción sobre el uso de la IA en educación superior	Artículos no relevantes al tema de investigación
Artículos que utilicen cuestionarios de escala como instrumento	Artículos con otro tipo de instrumento Artículos duplicados Artículos retractados

Fuente: elaboración propia.

Figura 1. *Diagrama del proceso de revisión basado en Prisma.*



Fuente: adaptado de Haddaway et al. (2022).

búsqueda principal (Scopus y WoS), 107 reportes pasaron a lectura a texto completo y 17 documentos se incluyeron; en la búsqueda complementaria (SciELO y Redalyc), 45 reportes pasaron a lectura de contenido, de ellos se incluyeron 9 documentos. Durante la búsqueda en SciELO y Redalyc, se observó una recuperación considerable de registros no pertinentes, aun utilizando operadores booleanos y filtros por periodo e idioma. Esto puede relacionarse con diferencias en la forma en que estas plataformas indexan contenidos y procesan consultas (p. ej., búsqueda en texto completo, metadatos heterogéneos y soporte limitado de anidación booleana), lo que reduce la precisión y exige un cribado más estricto. Por ello, se aplicaron de manera rigurosa los criterios de elegibilidad del estudio en el cribado por título/resumen y en la evaluación a texto completo, documentando el flujo mediante Prisma. Finalmente, 26 documentos cumplieron con todos los criterios.

Los documentos revisados analizan principalmente a estudiantes y profesores universitarios; en algunos casos, combinan a usuarios en línea, editores o profesionales de la salud. Estas poblaciones son de una diversidad geográfica amplia, que abarca países de América, Europa, Asia, África y Oceanía, además de algunos estudios multinacionales. Se pudo observar que los documentos consolidan marcos teóricos, como UTAUT2, TAM, TAME-ChatGPT, entre otros, con el objetivo de analizar los factores determinantes de la aceptación y el uso de la IA, en consideración a variables individuales, sociales y

contextuales, para comprender cómo influyen en las actitudes, las intenciones y los comportamientos de adopción en distintos entornos académicos, profesionales y de consumo.

Los estudios confirman que la aceptación de la IA en educación depende de una combinación de factores:

- Percepción de utilidad y facilidad de uso.
- Motivación, emociones y confianza.
- Factores sociales: influencia de pares, normas y condiciones de apoyo.
- Dimensiones culturales y de difusión: compatibilidad, observabilidad, riesgo, valores subjetivos.

En conjunto, muestran que la aceptación tecnológica de la IA es multidimensional e integra factores técnicos, psicológicos, sociales y culturales.

3. Resultados y discusión

El cuestionario de escala Likert comprende diversas características que deben cumplirse en su diseño metodológico. Primero, es indispensable identificar la variable que desea medirse (Fabila Echauri et al., 2014); para que una variable pueda ser tratada estadísticamente, tiene que ser medida, ya sea asignando un valor numérico a la observación o clasificándola en categorías (Glass y Hopkins, 2008; Gravetter et al., 2021).

Los estudios que utilizan el TAM como modelo teórico emplean constructos base, como utilidad percibida y facilidad de uso percibida, agregando variables adicionales para comprender mejor la integración de la IA en la educación superior. Por ejemplo, Almulla (2024) integra aprendizaje colaborativo, uso de ChatGPT, interacción con ChatGPT, calidad de la interacción, calidad de la información, motivación del aprendizaje y satisfacción del aprendizaje, mientras Rahman et al. (2023) incluyen la informatividad percibida, el disfrute percibido y la confianza.

Los estudios basados en la UTAUT2 utilizan constructos, como expectativa de rendimiento y de esfuerzo, influencia social, condiciones facilitadoras, motivación hedónica, valor del precio y hábito. Otros estudios, como los de Romero-Rodríguez et al. (2023), mantienen solo estos constructos; pero, algunos, como los de Grassini et al. (2024), reducen los constructos originales, eliminando el valor del precio al ser irrelevante para herramientas gratuitas. Finalmente, casos como G. C. et al. (2024) agregan constructos nuevos, como valor del aprendizaje y precisión de la información. Todos estos ajustes se fundamentan en la revisión de literatura y en la búsqueda de la comprensión del problema que se investiga, sobre todo, como mencionan Walsh et al. (2021), al no contemplar o controlar las variables de confusión, puede ocasionar que se tenga un sesgo por variable omitida.

Posterior a la identificación de las variables a medir, se continúa con la etapa de elaboración de ítems (Hernández Hernández y Izquierdo Sandoval, 2020), como frases que expresan una idea positiva o negativa respecto del problema investigado, para estimular una opinión. Para lograrlo, se define una escala de estimación, que se refiere a una graduación que suele ser de "totalmente en desacuerdo" a "totalmente de acuerdo" (Fabila Echauri et al., 2014). Cada grado de la escala recibe un valor numérico acorde con el investigador, con el cual pueden sumarse los ítems y analizarse estadísticamente para posibles generalizaciones y replicaciones (Dauzón-Ledesma y Izquierdo, 2023); estas escalas de estimación pueden tener entre dos y siete alternativas de respuesta. Los estudios analizados se dividen en escalas de 5 y 7 alternativas, lo que incluye un punto neutral que permite al encuestado evitar adoptar una posición, lo cual es adecuado y necesario para este instrumento,

puesto que se desea obtener la percepción real, sin forzar al encuestado a elegir una opinión.

Según Hodge y Gillespie (2007), los ítems deben ser unidimensionales, esto es, que solo pueden presentar información relacionada con una dimensión del constructo central, además deben tener univocidad, es decir, un ítem debe alentar la opinión de un aspecto específico de la dimensión. Los estudios analizados, en general, no presentan información sobre la búsqueda de la univocidad y la unidimensionalidad, lo que puede representar una debilidad metodológica; sin embargo, buscan obtener la confiabilidad y la validez del instrumento elaborado o adaptado para su estudio.

Con el fin de sistematizar los modelos teóricos y las evidencias psicométricas reportadas en los estudios incluidos, la tabla 4 presenta una síntesis comparativa de los instrumentos utilizados para analizar la aceptación de la IA en educación superior.

En los estudios recuperados en SciELO y Redalyc, se identificó una tendencia a diseños cuantitativos transversales basados en autoinforme. Aunque varios trabajos reportan indicadores psicométricos (p. ej., alfa de Cronbach y, en algunos casos, validez convergente/discriminante), se observaron limitaciones recurrentes: a) uso frecuente de muestreo no probabilístico o por conveniencia, b) concentración en una o pocas instituciones, c) tamaños muestrales reducidos en estudios de validación/instrumento en algunos casos y d) heterogeneidad en el nivel de detalle con que se describen los procesos de construcción y validación del cuestionario. Estas características metodológicas deben considerarse al interpretar el alcance y la generalización de los hallazgos regionales.

A partir de esta síntesis, se observan aspectos relevantes en la validez y confiabilidad como criterios para la construcción de los instrumentos de medición; sin embargo, se presentan variaciones entre estos y el enfoque metodológico utilizado. En este contexto, el análisis que sigue se centra en examinar comparativamente las estrategias de validación y confiabilidad reportadas, para identificar patrones recurrentes y vacíos metodológicos en la evaluación de instrumentos de aceptación de la IA en la educación superior.

Tabla 4. *Síntesis de modelos teóricos y evidencias psicométricas de los estudios incluidos*

Autor/año	Modelo teórico	Tipo de validez reportada	Confiabilidad reportada	Contexto geográfico
Amer Jid Almahri et al. (2024)	UTAUT2	Validez de constructo (PLSSEM)	Alfa de Cronbach, CR	Reino Unido
Strzelecki et al. (2024)	UTAUT2	Validez de constructo (PLSSEM)	Alfa de Cronbach, CR	Polonia
Cabero-Almenara et al. (2024)	UTAUT2	Validez de contenido, validez de constructo	Alfa de Cronbach	Ecuador
Mohd Rahim et al. (2022)	UTAUT2	Validez convergente y discriminante (PLS-SEM)	Alfa de Cronbach, CR	Malasia
Li y Qin (2023)	UTAUT2	Validez de constructo	Alfa de Cronbach	China
Strzelecki (2024)	UTAUT2	Validez de constructo (PLSSEM)	Alfa de Cronbach, CR	Polonia
Grassini et al. (2024)	UTAUT2	Validez de constructo (PLSSEM)	Alfa de Cronbach, CR	Noruega
Romero-Rodríguez et al. (2023)	UTAUT2	Validez de constructo	Alfa de Cronbach	España
G. C. et al. (2024)	UTAUT2	Validez de constructo (PLSSEM)	Alfa de Cronbach, CR	Nepal
Alzahrani y Alzahrani (2024)	UTAUT2	Validez convergente y discriminante	Alfa de Cronbach, CR	Arabia Saudita

Autor/año	Modelo teórico	Tipo de validez reportada	Confiabilidad reportada	Contexto geográfico
Falebita y Kok (2024)	TAM	Validez de constructo (PLSSEM)	Alfa de Cronbach, CR	Nigeria
Cambra-Fierro et al. (2024)	TAM	Validez de constructo (CB-SEM)	Alfa de Cronbach	España
Shahzad et al. (2024)	TAM	Validez de constructo (PLSSEM)	Alfa de Cronbach, CR	China
Rahman et al. (2023)	TAM	Validez de constructo (SEM)	Alfa de Cronbach	Bangladés
Almogren et al. (2024)	TAM	Validez de constructo (SEM)	Alfa de Cronbach	Arabia Saudita
Kanont et al. (2024)	TAM	Validez de constructo (SEM)	Alfa de Cronbach	Tailandia
Almulla (2024)	TAM	Validez de constructo (CFA)	Alfa de Cronbach	Arabia Saudita
Salazar-Altamirano et al. (2025)	TAM	Validez de constructo (PLSSEM)	Alfa de Cronbach, CR	México
Marimon et al. (2025)	TAM	Validez de contenido, validez de constructo	Alfa de Cronbach, CR	Argentina
Alshammari et al. (2025)	TAM	Validez de contenido, validez convergente y discriminante	CR	Arabia Saudita
Sallam et al. (2023)	TAME-ChatGPT	Validez de contenido, validez de constructo (EFA)	Alfa de Cronbach	Jordania
Romero Calle et al. (2025)	TPB	Validez convergente y discriminante (PLS-SEM)	Alfa de Cronbach, CR	Ecuador
Robles Morales (2025)	No especificado	Validez de contenido, validez convergente y discriminante	Alfa de Cronbach, CR	República Dominicana
Gil-Vera (2024)	No especificado	Validez de constructo (CB-SEM)	Alfa de Cronbach	Colombia
Sigüenza Orellana et al. (2024)	No especificado	Validez de contenido, validez de constructo (EFA)	Alfa de Cronbach	Ecuador
Gómez-García et al. (2025)	No especificado	Validez de constructo, validez convergente y discriminante	Alfa de Cronbach, CR	España

CR: confiabilidad compuesta.
Nota: los estudios se presentan agrupados según el modelo teórico empleado para facilitar la comparación entre enfoques.
Fuente: elaboración propia.

En relación con la confiabilidad de consistencia interna, que es la medida para saber si cada pregunta en una prueba mide lo mismo (Boyle y Fisher, 2008), la mayoría de los estudios la obtienen a través del alfa de Cronbach, aunque en algunos casos la combinan con otras estadísticas, como Mohd Rahim et al. (2022) y Shahzad et al. (2024), que usan la fiabilidad compuesta, o en el estudio de Strzelecki (2024), que utiliza el coeficiente de fiabilidad de Dijkstra y Henseler; otros más también buscan la confiabilidad entre evaluadores, que se refiere al grado de acuerdo entre los jueces (Duckett, 2021), como en los estudios de Almogren et al. (2024) y Sallam et al. (2023). Estos estudios encontraron una gran consistencia interna, eliminaron, en ciertos casos, solo uno o dos ítems con un coeficiente de correlación bajo y demostraron que el instrumento tiene un diseño original (UTAUT2 o TAM) bien definido, con validez y una alta adaptabilidad.

Verificar si un instrumento mide lo que se propone medir es comprobar su validez (Field, 2018); en la mayoría de los estudios analizados, se buscó la validez de constructo, en la que un instrumento de medición debe corresponder y representar con precisión el constructo teórico que el investigador pretende medir (Duckett, 2021). Esto se logra

a través de la validez convergente, cuando los ítems o las escalas que lo componen están de acuerdo entre sí y reflejan un constructo común; y la validez discriminante, cuando los ítems de un constructo son únicos en relación con otros constructos (Cohen et al., 2018).

Todos mostraron una alta validez de constructo, lo que indica una solidez en los constructos y en los ítems que los representan, además de tener flexibilidad para incorporar otros constructos que permitan robustecer la teoría, lo que hace posible una mejor comprensión de los problemas abordados. Algunos estudios obtienen la validez transcultural (Cohen et al., 2018), realizando traducciones hacia delante y hacia atrás (Sallam et al., 2023), esto que demuestra la adaptabilidad del instrumento a diferentes contextos geográficos y culturales, aunque también puede ser una debilidad metodológica de los estudios.

Gravetter et al. (2021) definen la población como el conjunto de individuos de interés en un estudio y la muestra como aquellos seleccionados para representarla. Además, Cohen et al. (2018) destacan la importancia de definir correctamente la población y elegir una muestra representativa para garantizar la validez de los resultados. Los estudios analizados incluyen poblaciones diversas en términos geográficos, académicos y disciplinarios, que abarca países como Polonia, Ecuador, Malasia, España, etc., midiendo percepciones de profesores o estudiantes de pregrado o posgrado, de áreas de conocimiento, como medicina, matemáticas, etc., lo que sugiere una fortaleza del instrumento, al medir las percepciones de manera adecuada a pesar de estas discrepancias poblacionales. Sin embargo, una proporción considerable de los estudios presenta una debilidad metodológica al no definir completamente la población y generalizar sus características.

Predomina el muestreo no probabilístico, siendo el muestreo por conveniencia el más común, además de un estudio que utiliza el muestreo bola de nieve (Rahman et al., 2023). El muestreo por conveniencia facilita la recolección de datos de manera rápida y económica, pero presenta limitaciones desde el punto de vista de la representatividad, lo que impide generalizar los resultados a una población mayor, que es más adecuado para estudios exploratorios (Cohen et al., 2018), como lo es en este caso, en el que se busca tener un primer acercamiento a las percepciones y los usos de la IA en la educación superior.

McKinley y Rose (2019) resaltan la importancia de definir atributos específicos para el muestreo intencional, como en Mohd Rahim et al. (2022), quienes definen criterios claros de elección, lo que fortalece la muestra elegida y selecciona a los participantes más adecuados para la investigación. Para la mayoría de los estudios, se precisan los tamaños de muestras a partir de distintas técnicas o enfoques, como la fórmula de Yamane (Strzelecki et al., 2024), el enfoque de Roscoe (Shahzad et al., 2024), la calculadora de tamaño de muestra *a priori* para modelos de ecuaciones estructurales (Cambra-Fierro et al., 2024), entre otras, lo que sugiere una debilidad, al no establecerse una consideración homogénea para las muestras.

3.1 Equidad digital y confianza institucional en América Latina

Aunque la mayoría de los estudios incluidos en la revisión no se realizaron en América Latina, los constructos evaluados mediante instrumentos basados en los modelos TAM y UTAUT2 resultan útiles para comprender los retos vinculados a la adopción de la IA en la educación superior de la región. Variables como las condiciones facilitadoras, la expectativa de esfuerzo y la utilidad percibida adquieren matices particulares en escenarios caracterizados por brechas persistentes de conectividad, desigualdades en el acceso



a infraestructura tecnológica y niveles heterogéneos de competencias digitales entre estudiantes y docentes. En este contexto, la aceptación de herramientas de IA no se explica solo por los beneficios pedagógicos que se les atribuyen, sino también por la posibilidad real de integrarlas y sostener su uso en instituciones públicas con recursos limitados.

De forma complementaria, la confianza institucional se configura como un factor central en la adopción de la IA en las universidades latinoamericanas. A diferencia de aquellos contextos en los que existen marcos regulatorios consolidados y políticas institucionales explícitas sobre el uso de estas tecnologías, en numerosas instituciones de la región la falta de lineamientos claros en materia de ética, protección de datos e integridad académica puede afectar la percepción de legitimidad y seguridad asociada a la IA. En este escenario, los estudios que incorporan la confianza como variable explicativa aportan elementos clave para comprender los procesos de aceptación de la IA en contextos en los que la articulación entre tecnología y gobernanza institucional aún se encuentra en desarrollo.

Desde una perspectiva crítica, estos elementos refuerzan la necesidad de adaptar los instrumentos de medición no solo a nivel lingüístico, sino también a nivel contextual y normativo. La incorporación explícita de indicadores vinculados a la equidad digital y la confianza institucional permitiría capturar con mayor precisión las condiciones reales de adopción de la IA en la educación superior latinoamericana, así como evitar interpretaciones descontextualizadas sustentadas exclusivamente en modelos concebidos en entornos con mayores niveles de desarrollo y madurez tecnológica.

3.2 Propuesta de lineamientos para el diseño en América Latina

La revisión de los estudios analizados muestra que la aplicación de cuestionarios de escala Likert se realiza principalmente en formato digital, a través de medios, como correo electrónico, Facebook, WhatsApp, LinkedIn, y otras plataformas de distribución en línea. Este formato ofrece ventajas logísticas, pero también plantea retos desde el punto de vista de la validez y honestidad de las respuestas (Cohen et al., 2018). En América Latina, donde existen brechas tecnológicas y de acceso a internet, estas limitaciones pueden ser más significativas y afectar la representatividad de los resultados. Por ello, se recomienda reforzar el diseño metodológico mediante pruebas piloto regionales que permitan ajustar el lenguaje, la interfaz y el modo de aplicación a los contextos locales (G. C. et al., 2024). Asimismo, resulta esencial garantizar el anonimato y la no trazabilidad de las respuestas, lo cual ha demostrado mejorar la precisión y la honestidad de los participantes (Kanont et al., 2024; Romero-Rodríguez et al., 2023).

La revisión y el análisis de datos también constituyen una etapa crítica para evitar sesgos y preservar la validez de resultados (O'Mara-Eves y Thomas, 2016). En este sentido, los estudios revisados aplican estrategias que pueden adaptarse a la región, como la eliminación de respuestas con varianza cero (Strzelecki, 2024), la detección de valores atípicos y pruebas de normalidad,

homogeneidad y multicolinealidad (Amer Jid Almahri et al., 2024), o el uso de imputación de medias para manejar datos faltantes (Almulla, 2024). En contextos en los que las muestras suelen ser pequeñas o diversas, estos procedimientos ayudan a mantener el poder estadístico y la estabilidad de los modelos.

En cuanto al análisis, los métodos no paramétricos se confirman como los más adecuados cuando los datos no presentan una distribución normal (Field, 2018), ya que son más robustos ante violaciones de supuestos estadísticos (Norman, 2010). En los estudios revisados, se emplean pruebas como Mann-Whitney, Kruskal-Wallis y rho de Spearman (Cabero-Almenara et al., 2024). Sin embargo, para alcanzar una mayor capacidad explicativa y predictiva, varios autores han implementado modelos de ecuaciones estructurales por mínimos cuadrados parciales (PLS-SEM, por sus siglas en inglés) (Falebita y Kok, 2024; Shahzad et al., 2024), que, en algunos casos, combinan con redes neuronales artificiales (PLS-SEM-ANN, por sus siglas en inglés) para mejorar el poder de predicción (Mohd Rahim et al., 2022). También se reporta el uso de modelos basados en covarianza (CB-SEM, por sus siglas en inglés) en análisis paramétricos (Cambra-Fierro et al., 2024), cuya adaptación en la región podría fortalecer la rigurosidad de los estudios futuros.

Los resultados de los estudios tienen una tendencia hacia constructos, como el hábito, la motivación hedónica y la utilidad percibida como factores de influencia clave en la adopción de herramientas basadas en IA, como ChatGPT. Por ejemplo, Strzelecki et al. (2024) indican que el hábito es el mayor predictor de intención de uso, resultado similar en Strzelecki (2024) y Grassini et al. (2024). Igualmente, la utilidad percibida muestra una relación sólida con la intención de uso (Kanont et al., 2024; Shahzad et al., 2024). Aunque la facilidad de uso percibida también es significativa, su impacto varía según el contexto; mientras estudios, como el de Almogren et al. (2024), confirman su influencia, otros como el de Li y Qin (2023) no encontraron una relación significativa.

Finalmente, resulta relevante mencionar que algunas variables moderadoras (edad, género, nivel educativo, experiencia, etc.) en estudios, como el de Amer Jid Almahri et al. (2024), presentan efectos en la intención de uso, mientras en otros no se encuentran impactos significativos. A partir de estas evidencias, se proponen los siguientes lineamientos de diseño y validación adaptados a América Latina:

- Contextualización tecnológica y sociotécnica: Emplear plataformas de bajo consumo de datos, accesibles desde dispositivos distintos y compatibles con redes inestables, que garanticen equidad en la participación.
- Validación cultural y lingüística: Realizar pruebas piloto interinstitucionales que consideren variantes idiomáticas, conceptuales y contextos institucionales.
- Garantías éticas y de confidencialidad: Reforzar el anonimato y la transparencia en el tratamiento de datos, siguiendo las recomendaciones de Cohen et al. (2018).
- Control de calidad de los ítems: Integrar reactivos invertidos o de verificación interna para reducir el sesgo de respuesta y evaluar la atención del participante.
- Combinación de enfoques analíticos: Utilizar modelos no paramétricos en muestras pequeñas y métodos estructurales o híbridos en muestras mayores, y promover la triangulación metodológica.
- Creación de bancos de ítems regionales: Desarrollar repositorios latinoamericanos de ítems validados y comparables entre países o instituciones, con información sobre confiabilidad y validez contextual.
- Inclusión de dimensiones emergentes: Agregar constructos relevantes para la

región como equidad tecnológica, ética académica, confianza institucional y bienestar educativo.

En conjunto, estos lineamientos se orientan hacia una metodología regional de diseño y validación de cuestionarios de escala, que mantenga los estándares internacionales de confiabilidad y validez, pero incorporando la diversidad lingüística, cultural y tecnológica que caracteriza a América Latina. De esta forma, las escalas podrán medir no solo percepciones universales sobre la IA, sino también los matices locales de su aceptación, resistencia y apropiación educativa.

4. Conclusiones

El estudio permitió constatar que los cuestionarios de escala Likert empleados en estudios internacionales muestran una alta consistencia interna, validez de constructo y flexibilidad para integrar nuevos factores asociados a la aceptación de la IA. Sin embargo, los resultados también evidencian vacíos significativos cuando se pretende aplicar estos instrumentos en América Latina, donde las realidades educativas, tecnológicas y culturales difieren de los entornos en los cuales fueron originalmente diseñados.

En América Latina, la diversidad institucional, la desigualdad en el acceso a recursos tecnológicos y las diferencias en la formación tecnológica de profesores y estudiantes requiere que la adaptación cultural y lingüística de los cuestionarios tengan una rigurosidad mayor. De igual manera, en los procesos de validación ecológica, se consideran aspectos contextuales como la infraestructura tecnológica, la confianza institucional en la IA, las políticas o los lineamientos de uso y las actitudes éticas relacionadas con el uso académico de estas herramientas.

Se evidencian vacíos significativos cuando se pretende aplicar estos instrumentos en América Latina, donde las realidades educativas, tecnológicas y culturales difieren de los entornos en los cuales fueron originalmente diseñados.

Por tanto, se recomienda que los investigadores latinoamericanos no se limiten a traducir y aplicar instrumentos desarrollados en otros países, sino que realicen procesos de reconstrucción teóricometodológica que aseguren la pertinencia contextual, cultural y conceptual de cada ítem. Esto incluye realizar pruebas piloto regionales o locales, aplicar técnicas compuestas de validación (contenido, constructo y ecológica) y fomentar colaboraciones interinstitucionales para adaptar bancos de ítems a la realidad educativa latinoamericana.

Además, los resultados sugieren que el diseño de futuros instrumentos incorpore dimensiones emergentes relevantes para la región, como la equidad tecnológica, la percepción ética y la confianza institucional en la IA, que pueden enriquecer los modelos de aceptación existentes (TAM, UTAUT2, TAME-ChatGPT). De este modo, los cuestionarios de escala Likert no solo contribuirán a medir las percepciones o actitudes, sino que también orientarán la formulación de políticas universitarias y estrategias pedagógicas adecuadas al contexto, convirtiéndose en herramientas clave para una integración responsable de la IA en la educación superior latinoamericana.

Financiación

Esta investigación no tiene financiación externa.

Conflicto de intereses

Los autores declaran que no tienen conflicto de intereses.

Implicaciones éticas

No aplica.

Contribución de autores

Diseño de la investigación (Gerardo Antonio Hernández-Torres y Manuel Jesús Izquierdo Sandoval), análisis de datos (Gerardo Antonio Hernández-Torres y Silvia Patricia Aquino Zúñiga), metodología (Gerardo Antonio Hernández-Torres y Silvia Patricia Aquino Zúñiga) y revisión del manuscrito (Gerardo Antonio Hernández-Torres y Silvia Patricia Aquino Zúñiga). Todos los autores han leído y aprobado la versión enviada a la revista.

Declaración de las tecnologías generativas asistidas por inteligencia artificial (IA) en el proceso de escritura

Ninguna tecnología asistida por la IA fue usada durante la escritura del artículo.

Referencias

Acosta-Enríquez, B. G., Arbulú Ballesteros, M. A., Huamaní Jordan, O., López Roca, C. y Saavedra Tirado, K. (2024). 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*, *12*(1). https://doi.org/10.1186/s40359-024-01764-z

Alhwaiti, M. (2023). Acceptance of artificial intelligence application in the post-covid ERA and its impact on faculty members' occupational well-being and teaching self efficacy: A path analysis using the utaut 2 model. *Applied Artificial Intelligence*, *37*(1), 2175110. https://doi.org/10.1080/08839514.2023.2175110

Almogren, A. S., Al-Rahmi, W. M. y Dahri, N. A. (2024). Exploring factors influencing the acceptance of ChatGPT in higher education: A smart education perspective. *Heliyon*, *10*(11), e31887. https://doi.org/10.1016/j.heliyon.2024.e31887

Almulla, M. A. (2024). Investigating influencing factors of learning satisfaction in AI ChatGPT for research: University students perspective. *Heliyon*, *10*(11), e32220. https://doi.org/10.1016/j.heliyon.2024.e32220

Alshammari, S. H., Almankory, A. Z. y Alrashidi, M. E. (2025). Efectos de la conciencia y confianza en la disposición estudiantil para usar ChatGPT: Modelo TAM-ECM integrado. *RIED: Revista Iberoamericana de Educación a Distancia*, *28*(2), 155-180. https://doi.org/10.5944/ried.28.2.43476

Alzahrani, A. y Alzahrani, A. (2024). Comprendiendo la adopción de ChatGPT en universidades: El impacto del TPACK y UTAUT2 en los docentes. *RIED: Revista Iberoamericana de Educación a Distancia*, *28*(1). https://doi.org/10.5944/ried.28.1.41498

Amer Jid Almahri, F. A., Bell, D. y Gulzar, Z. (2024). Chatbot technology use and acceptance using educational personas. *Informatics*, *11*(2). https://doi.org/10.3390/informatics11020038

Ayoubi, K. (2024). Adopting ChatGPT: Pioneering a new era in learning platforms. *International Journal of Data and Network Science*, *8*(2), 1341-1348. https://doi.org/10.5267/j.ijdns.2023.11.001

Barroga, E. y Matanguihan, G. J. (2022). A practical guide to writing quantitative and qualitative research questions and hypotheses in scholarly articles. *Journal of Korean Medical Science*, *37*(16), e121. https://doi.org/10.3346/jkms.2022.37.e121

Boyle, J. y Fisher, S. (2008). *Educational testing: A competence-based approach*. John Wiley & Sons.

Bryman, A. (1984). The debate about quantitative and qualitative research: A question of method or epistemology? *The British Journal of Sociology*, *35*(1), 75-92. https://doi.org/10.2307/590553

Cabero-Almenara, J., Palacios-Rodríguez, A., Loaiza-Aguirre, M. I. y Rivas-Manzano, M.ª del R. de. (2024). Acceptance of educational artificial intelligence by teachers and its relationship with some variables and pedagogical beliefs. *Education Sciences*, *14*(7). https://doi.org/10.3390/educsci14070740

Cambra-Fierro, J. J., Fuentes Blasco, M.ª, López-Pérez, M.ª E. y Trifu, A. (2025). ChatGPT adoption and its influence on faculty well-being: An empirical research in higher education. *Education and Information Technologies*, *30*(2), 1517-1538. https://doi.org/10.1007/s10639-024-12871-0

Cohen, L., Manion, L. y Morrison, K. (2018). *Research methods in education* (8.ª ed.). Routledge.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, *78*(1), 98-104. https://doi.org/10.1037/0021-9010.78.1.98

Dauzón-Ledesma, L. e Izquierdo, J. (2023). Language learning investment in Higher Education: Validation and implementation of a Likert-scale questionnaire in the context of compulsory EFL Learning. *Education Sciences*, *13*(4). https://doi.org/10.3390/educsci13040370

Davis, F. D., Bagozzi, R. P. y Warshaw, P. R. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management Science*, *35*(8), 982-1003. https://doi.org/10.1287/mnsc.35.8.982

Duckett, L. J. (2021). Quantitative research excellence: Study design and reliable and valid measurement of variables. *Journal of Human Lactation*, *37*(3), 456-463. https://doi.org/10.1177/08903344211019285

Fabila Echauri, A. M., Minami, H. y Izquierdo Sandoval, M. J. (2014). La escala de Likert en la evaluación docente: Acercamiento a sus características y principios metodológicos. *Perspectivas docentes*, 50. https://doi.org/10.19136/pd.a0n50.589

Falebita, O. S. y Kok, P. J. (2025). Artificial intelligence tools usage: A structural equation modeling of undergraduates' technological readiness, self-efficacy and attitudes. *Journal for STEM*

Education Research, *8*(2), 257-282. https://doi.org/10.1007/s41979-024-00132-1 Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (5.ª ed.). Sage.

G. C, S. B., Bhandari, P., Gurung, S. K., Srivastava, E., Ojha, D. y Dhungana, B. R. (2024). Examining the role of social influence, learning value and habit on students' intention to use ChatGPT: The moderating effect of information accuracy in the UTAUT2 model. *Cogent Education*, *11*(1), 2403287. https://doi.org/10.1080/2331186X.2024.2403287

Gil-Vera, V. D. (2024). Uso de ChatGPT por estudiantes universitarios: Un análisis relacional. *Formación Universitaria*, *17*(5), 129-138. https://doi.org/10.4067/s0718-50062024000400129

Glass, G. V. y Hopkins, K. D. (2008). *Statistical methods in education and psychology* (3.ª ed.). Allyn & Bacon.

Gómez-García, M., Ruiz-Palmero, J., Boumadan-Hamed, M. y Soto-Varela, R. (2025). Perceptions of future teachers and pedagogues on responsible AI: A measurement instrument. *RIED: Revista Iberoamericana de Educación a Distancia*, 28(2), 105-130. https://doi.org/10.5944/ried.28.2.43288

Grassini, S., Aasen, M. L. y Møgelvang, A. (2024). Understanding university students' acceptance of ChatGPT: Insights from the UTAUT2 model. *Applied Artificial Intelli-*

gence, *38*(1), 2371168. https://doi.org/10.1080/08839514.2024.2371168

Gravetter, F. J., Wallnau, L. B., Forzano, L. B. y Witnauer, J. E. (2021). *Essentials of statistics for the behavioral sciences* (10.ª ed.). Cengage.

Haddaway, N. R., Page, M. J., Pritchard, C. C. y McGuinness, L. A. (2022). PRISMA2020: An R package and Shiny app for producing Prisma 2020-compliant flow diagrams, with interactivity for optimised digital transparency and Open Synthesis. *Campbell Systematic Reviews*, *18*(2), e1230. https://doi.org/10.1002/cl2.1230

Hernández Hernández, M. e Izquierdo Sandoval, M. J. (2020). Cambios curriculares y enseñanza del inglés: Cuestionario de percepción docente. *Sinéctica: Revista Electrónica de Educación*, *54*, 1-22. https://doi.org/10.31391/S2007-7033(2020)0054-012

Hodge, D. R. y Gillespie, D. F. (2007). Phrase completion scales: A better measurement approach than Likert scales? *Journal of Social Service Research*, *33*(4), 1-12. http://dx.doi.org/10.1300/J079v33n04_01

Hoy, W. K. y Adams, C. M. (2016). *Quantitative research in education: A primer* (2.ª ed.). Sage.

Kanont, K., Pingmuang, P., Simasathien, T., Wisnuwong, S., Wiwatsiripong, B., Poonpirome, K., Songkram, N. y Khlaisang, J. (2024). Generative-AI, a learning assistant? Factors influencing higher-ed students' technology acceptance. *Electronic Journal of E-Learning*, *22*(6), 18-33. https://doi.org/10.34190/ejel.22.6.3196

Köhler, T., Landis, R. S. y Cortina, J. M. (2017). From the editors: Establishing methodological rigor in quantitative management learning and education research: The role of design, statistical methods, and reporting standards. *Academy of Management Learning & Education*, *16*(2), 173-192. https://doi.org/10.5465/amle.2017.0079

Li, Q. y Qin, Y. (2023). AI in medical education: Medical student perception, curriculum recommendations and design suggestions. *BMC Medical Education*, *23*(1). https://doi.org/10.1186/s12909-023-04700-8

Lodico, M. G., Spaulding, D. T. y Voegtle, K. H. (2006). *Methods in educational research: From theory to practice*. Jossey-Bass.

Mahmud, A., Sarower, A. H., Sohel, A., Assaduzzaman, M. y Bhuiyan, T. (2024). Adoption of ChatGPT by university students for academic purposes: Partial least square, artificial neural network, deep neural network and classification algorithms approach. *Array*, *21*, 100339. https://doi.org/10.1016/j.array.2024.100339

Marimon, F., Arias Valle, M. B., Coria Augusto, C. J. y Larrea Arnau, C. M. (2025). Del optimismo a la confianza: El impacto de ChatGPT en la confianza de los estudiantes en el aprendizaje asistido por IA. *RIED: Revista Iberoamericana de Educación a Distancia*, *28*(2), 131-153. https://doi.org/10.5944/ried.28.2.43238

McKinley, J. y Rose, H. (eds.) (2019). *The Routledge handbook of research methods in applied linguistics*. Routledge. https://doi.org/10.4324/9780367824471

Merhi, R. (2011). Expectativas del estudiantado en la universidad del nuevo milenio. *La Cuestión Universitaria*, *7*, 23-31. https://polired.upm.es/index.php/lacuestionuniversitaria/article/view/3353/3418

Mohd Rahim, N. I., Iahad, N. A., Yusof, A. F. y Al-Sharafi, M. A. (2022). AI-based chatbots adoption model for higher-education institutions: A hybrid PLS-SEM-neural network modelling approach. *Sustainability*, *14*(19), 12726. https://doi.org/10.3390/su141912726

Norman, G. (2010). Likert scales, levels of measurement and the "laws" of statistics. *Advances in Health Sciences Education*, *15*(5), 625-632. https://doi.org/10.1007/s10459-010-9222-y

O'Mara-Eves, A. y Thomas, J. (2016). Ongoing developments in meta-analytic and quantitative synthesis methods: Broadening the types of research questions that can be addressed. *Review of Education*, *4*(1), 5-27. https://doi.org/10.1002/rev3.3062

OpenAI. (2022, 30 de noviembre). *Presentamos ChatGPT*. https://openai.com/es419/index/chatgpt/

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S. … AlonsoFernández, S. (2021). Declaración Prisma 2020: Una guía actualizada para la publicación de revisiones sistemáticas. *Revista Española de Cardiología*, *74*(9), 790-799. https://doi.org/10.1016/j.recesp.2021.06.016

Persson, P. B., Hillmeister, P. y Persson, A. B. (2022). Perception. *Acta Physiologica*, *235*(3), e13842. https://doi.org/10.1111/apha.13842

Rahman, M. S., Sabbir, M. M., Zhang, J., Moral, I. H. y Hossain, G. M. S. (2023). Examining students' intention to use ChatGPT: Does trust matter? *Australasian Journal of Educational Technology*, *39*(6), 51-71. https://doi.org/10.14742/ajet.8956

Robles Morales, R. E. (2025). Factores determinantes en la adopción de inteligencia artificial en la educación superior dominicana. *Cuaderno de Pedagogía Universitaria*, *22*(43), 79-103. https://doi.org/10.29197/cpu.v22i43.647

Romero Calle, G. P., Tivillin-Gutama, D. M. y Bonisoli, L. (2025). La inteligencia artificial y su influencia en el comportamiento de los estudiantes. *Kairós: Revista de Ciencias Económicas, Jurídicas y Administrativas*, *8*(14), 67-87. https://doi.org/10.37135/kai.03.14.04

Romero-Rodríguez, J. M., Ramírez-Montoya, M. S., Buenestado-Fernández, M. y Lara-Lara, F. (2023). Use of ChatGPT at university as a tool for complex thinking: Students' perceived usefulness. *Journal of New Approaches in Educational Research*, *12*(2), 323-339. https://doi.org/10.7821/naer.2023.7.1458

Ruiz Mendoza, K. K., Miramontes Arteaga, M.ª A. y Reyna García, C. (2024). Percepciones y expectativas de estudiantes universitarios sobre la IAG. *European Public & Social Innovation Review*, *9*, 1-21. https://doi.org/10.31637/epsir-2024-357

Saihi, A., Ben-Daya, M., Hariga, M. y As'ad, R. (2024). A structural equation modeling analysis of generative AI chatbots adoption among students and educators in higher education. *Computers and Education: Artificial Intelligence*, *7*, 100274. https://doi.org/10.1016/j.caeai.2024.100274

Salazar-Altamirano, M. A., Martínez-Arvizu, O. J., Galván-Vela, E., Ravina-Ripoll, R., HernándezArteaga, L. G. y Gómez Sánchez, D. (2025). AI as a facilitator of creativity and wellbeing in business students: A multigroup approach between public and private universities. *Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ci*ência da Informação, *30*, 1-30. https://doi.org/10.5007/1518-2924.2025.e103485

Sallam, M., Salim, N. A., Barakat, M., Al-Mahzoum, K., Al-Tammemi, A. B., Malaeb, D., Hallit, R. y Hallit, S. (2023). Assessing health students' attitudes and usage of ChatGPT in Jordan: Validation study. *JMIR Medical Education*, *9*, e48254. https://doi.org/10.2196/48254

Scates, D. E. y Hoban, C. F. (1937). Critical questions for the evaluation of research. *The Journal of Educational Research*, *31*(4), 241-254. https://doi.org/10.1080/00220671.1937.10880747

Scott, D. (2007). Resolving the quantitative-qualitative dilemma: A critical realist approach. *International Journal of Research & Method in Education*, *30*(1), 3-17. https://doi.org/10.1080/17437270701207694

Segarra Ciprés, M., Grangel Seguer, R. y Belmonte Fernández, Ó. (2024). ChatGPT como herramienta de apoyo al aprendizaje en la educación superior: Una experiencia docente. *Revista Tecnología, Ciencia y Educación*, *28*, 7-44. https://doi.org/10.51302/tce.2024.19083

Shahzad, M. F., Xu, S. y Javed, I. (2024). ChatGPT awareness, acceptance, and adoption in higher education: The role of trust as a cornerstone. *International Journal of Educational Technology in Higher Education*, *21*(1), 46. https://doi.org/10.1186/s41239-024-00478-x

Shaker, P. y Ruitenberg, C. (2007). 'Scientifically-based research': The art of poli-

tics and the distortion of science. *International Journal of Research & Method in Education*, *30*(2), 207-219. https://doi.org/10.1080/17437270701383545

Sigüenza Orellana, J., Andrade Cordero, C. y Chitacapa Espinoza, J. (2024). Validación del cuestionario para docentes: Percepción sobre el uso de ChatGPT en la educación superior. *Revista Andina de Educación*, *8*(1), 000816. https://doi.org/10.32719/26312816.2024.8.1.6

Strzelecki, A. (2024). Students' acceptance of ChatGPT in higher education: An extended unified theory of acceptance and use of technology. *Innovative Higher Education*, *49*(2), 223-245. https://doi.org/10.1007/s10755-023-09686-1

Strzelecki, A., Cicha, K., Rizun, M. y Rutecka, P. (2024). Acceptance and use of ChatGPT in the academic community. *Education and Information Technologies*, *29*(17), 22943-22968. https://doi.org/10.1007/s10639-024-12765-1

Universidad Juárez Autónoma de Tabasco. (2024). *Plan de Desarrollo Institucional 2024-2028*. https://archivos.ujat.mx/2024/rectoria/PDI-24-28.pdf

Venkatesh, V., Thong, J. Y. y Xu, X. (2012). Consumer acceptance and use of information technology: Extending the unified theory of acceptance and use of technology. *MIS Quarterly*, *36*(1), 157-178. https://doi.org/10.2307/41410412

Villegas-Ch, W., García-Ortiz, J. y Sánchez-Viteri, S. (2024). Personalization of learning: Machine learning models for adapting educational content to individual learning styles. *IEEE Access*, *12*, 121114-121130. https://doi.org/10.1109/ACCESS.2024.3452592

Walsh, C., Stein, M. M., Tapping, R., Smith, E. M. y Holmes, N. G. (2021). Exploring the effects of omitted variable bias in physics education research. *Physical Review Physics Education Research*, *17*(1), 010119. https://doi.org/10.1103/PhysRevPhysEducRes.17.010119

White, P. (2013). Who's afraid of research questions? The neglect of research questions in the methods literature and a call for question-led methods teaching. *International Journal of Research & Method in Education*, *36*(3), 213-227. https://doi.org/10.1080/1743727X.2013.809413

Zhu, J. y Liu, W. (2020). A tale of two databases: The use of Web of Science and Scopus in academic papers. *Scientometrics*, *123*(1), 321-335. https://doi.org/10.1007/s11192-020-03387-8

Zyphur, M. J. y Pierides, D. C. (2020). Making quantitative research work: From positivist dogma to actual social scientific inquiry. *Journal of Business Ethics*, *167*(1), 49-62. https://doi.org/10.1007/s10551-019-04189-6