



AIESAD

Ried

**Revista Iberoamericana de
Educación a Distancia**

La Revista Iberoamericana de la Educación Digital

VOL. 28 N° 2 JULIO, 2025
ISSN: 1138-2783

RIED

Esta publicación de periodicidad semestral está dirigida a los estudiosos e investigadores del ámbito educativo, docentes universitarios y público interesado en su objeto de estudio. La RIED centra su atención en la difusión de ensayos, trabajos de carácter científico y experiencias innovadoras dentro del ámbito de la educación a distancia en cualesquiera de sus formulaciones y de las tecnologías aplicadas a la educación.

La RIED se gestiona íntegramente a través del Open Journal System (OJS), tanto para la edición como para la relación con los autores y revisores, así como para la difusión electrónica en abierto.

La RIED, además de su formato impreso, se publica en formato electrónico en dos sedes: OJS en UNED de España: <http://revistas.uned.es/index.php/ried>

INTERCAMBIOS y SUSCRIPCIONES:

RIED. Revista Iberoamericana de Educación a Distancia.

UTPL – SAN CAYETANO ALTO, s/n

Loja (Ecuador)

ried@utpl.edu.ec

Consejo Directivo de AIESAD (Asesor en RIED)

- **Presidente:** Jaime Leal Afanador, Rector Magfco. Universidad Nacional Abierta y a Distancia UNAD, Colombia
- **Directora Ejecutiva:** Constanza Abadía García, Vicerrectora. Universidad Nacional Abierta y a Distancia – UNAD. Colombia.
- **Vicepresidente Primero:** Ricardo Mairal Usón, Rector Magfco. Universidad Nacional de Educación a Distancia UNED. España
- **Vicepresidente Segundo:** Rodrigo Arias Camacho, Rector Magfco. Universidad Estatal a Distancia UNED. Costa Rica
- **Vicepresidente Tercero:** Santiago Acosta Aide, Rector Magfco. Universidad Técnica Particular de Loja. Ecuador
- **Vicepresidente Cuarto:** David Mejía Rodríguez, Rector Magnífico. Universidad Internacional de La Rioja – UNIR. México.
- **Vocales:**
 - Alfredo Alonso, Rector Magfco. Universidad Nacional de Quilmes UNQ. Argentina
 - Carla Padrel de Oliveira, Rectora Magfca. Universidad Aberta UAB. Portugal
 - Alejandra Hernández Acosta, Rectora Magnífica. Universidad Abierta para Adultos – UAPA. República Dominicana
 - Fray Álvaro José Arango Restrepo, O.P., Rector Magnífico. Universidad Santo Tomás – USTA. Colombia
- **Secretaría permanente y Tesorería:** Laura Alba Juez, Vicerrectora de Internacionalización. Universidad Nacional de Educación a Distancia – UNED. España.

Director/Editor (Director/Editor-in-Chief)

- Dr. Lorenzo García Aretio, Universidad Nacional de Educación a Distancia (UNED), España

Directora/Editora Adjunta (Deputy Director/Editor)

- Dra. Cristina Díaz de la Cruz, UTPL (Ecuador)

Consejo Editorial y Científico (Editorial and Scientific Board)

- Jordi Adell Segura. Univ. Jaime I (España).
- José Ignacio Aguaded Gómez. Univ. de Huelva (España). Editor “Comunicar”
- Luisa Aires. Univ. Aberta (Portugal).
- Terry Anderson. Athabasca Univ. (Canadá). Editor Emeritus IRRODL.
- Manuel Area Moreira. Univ. de La Laguna (España).
- Elena Barberá Gregori. UOC (España).
- Elena Bárcena Madera. UNED (España)
- Antonio Bartolomé Pina. Univ. de Barcelona (España).
- Julio Cabero Almenara. Univ. de Sevilla (España). Editor Pixelbit.
- Isabel Cantón Mayo. Univ. de León (España).
- Linda Castañeda. Univ. de Murcia (España).
- Manuel Castro Gil. UNED (España).
- M. Elena Chan Núñez. Univ. de Guadalajara (México).
- Cristóbal Cobo. Univ. of Oxford (R. Unido). The World Bank.
- César Coll Salvador. Univ. de Barcelona (España).
- Grainne Conole. e4innovation (R.Unido).
- Laura Czerniewicz. Univ. of Cape Town (Sudáfrica). CILT.
- Carlos Delgado Kloos. Univ. Carlos III (España). Cátedra UNESCO Educación Digital.
- Frida Díaz Barriga. UNAM (México).
- M. Esther del Moral Pérez. Univ. de Oviedo (España).
- Pierre Dillenbourg. EPFL (Suiza). Coordinator CHILI Lab.
- Josep M. Duarte. UOC (España). Editor ETHE Journal. EDEN President.
- Rubén Edel Navarro. Univ. Veracruzana (México).
- María Jesús Gallego-Arrufat. Univ. de Granada (España).
- Francisco J. García Peñalvo. Univ. de Salamanca (España). Editor EKS.

- Ana García-Valcárcel Muñoz-Repiso. Univ. de Salamanca (España).
- Mercè Gisbert Cervera. Univ. Rovira i Virgili (España).
- Carina Soledad González González. Univ. de La Laguna (España).
- Mercedes González Sanmamed. Univ. de A Coruña (España).
- Jaime Leal Afanador. UNAD (Colombia). Presidente AIESAD.
- M. del Carmen Llorente-Cejudo. Univ. de Sevilla (España).
- Ricardo Mairal Usón. Rector UNED (España). EADTU Ex-President.
- Carlos Marcelo García. Univ. de Sevilla (España).
- João Mattar. Pontificia Univ. Católica de São Paulo (Brasil).
- Rory McGreal. Athabasca Univ. Editor de IRRODL (Canadá). UNESCO Chair in OER.
- Daniel Mill. Univ. Federal de São Carlos (Brasil). Grupo Horizonte.
- António Moreira Teixeira. Univ. Aberta (Portugal). EDEN Ex-President.
- Lina Morgado. Univ. Aberta (Portugal).
- Jaime MuñozArteaga. Univ. Autónoma Aguascalientes (México).
- Sara Osuna-Acedo. UNED (España).
- Adolfin Pérez Garcías. Univ. Islas Baleares (España).
- Mar Pérez-Sanagustín. Univ. Paul Sabatier (Francia).
- Teresa Pessoa. Univ. de Coimbra (Portugal).
- M. Paz Prendes Espinosa. Univ. de Murcia (España).
- Claudio Rama. IESAL/UNESCO (Venezuela).
- M. Soledad Ramírez Montoya. TEC Monterrey (México). Cátedra UNESCO-ICDE.
- Rosabel Roig. Univ. de Alicante (España). Editora NAER.
- Jesús Salinas Ibáñez. Univ. Islas Baleares (España). Editor EDUTEC.
- Ángeles Sánchez-Elvira. UNED (España) Cátedra UNESCO EaD.
- Juana Sancho Gil. Univ. de Barcelona (España).
- Albert Sangrà. UOC (España). Cátedra UNESCO E-learning.
- Alan W. Tait. The OU (R.Unido). EDEN Ex-President. Ex-Editor Open Learning.
- Hernán Thomas. Univ. Nacional de Quilmes (Argentina).
- Javier Tourón Figueroa. UNIR (España).
- Martin James Weller. The OU (R. Unido). Director OER Hub. Editor JIME.
- Miguel Zapata Ros. Univ. de Murcia (España). Editor RED.

Delegación de AIESAD en RIED (AIESAD's Delegate in RIED)

- Santiago Acosta Aide, Univ. Técnica Particular de Loja (UTPL), Ecuador
- Cristina Díaz de la Cruz. UTPL, Ecuador

Secretaría Técnica (Technical Secretariat)

- José Luis García Boyé, AIESAD, España

Consejo de Redacción (Editing Board)

- Laura T. Alonso Díaz, Univ. Extremadura (España).
- Julia Ángel Osorio, UNAD (Colombia).
- Analia Chiecher, Univ. Río Cuarto, (Argentina).
- Paz Díez Arcón, UPM (España)
- Anna Engel, Univ. de Barcelona (España).
- Francesc Esteve Mon, Univ. Jaime I (España).
- Alicia García Holgado, Univ. Salamanca (España)
- Montse Guitert Catasús, UOC (España).
- M. José Hernández Serrano, Univ. Salamanca (España)
- Agustín Lagunes Domínguez, Univ. Veracruzana (México).
- Verónica Marín Díaz, Univ. Córdoba (España).
- Victoria I. Marín, Univ. Lleida (España)
- Elena M. Martín Monje, UNED (España).
- Rafael Morales Gamboa, Univ. Guadalajara (México).
- João Paz, Univ. Aberta (Portugal).
- Neuza Pedro, Univ. Lisboa (Portugal)
- Iliana Ramírez Asanza, UTPL (Ecuador).
- Enrique Ruiz Velasco, UNAM, (México).
- José M. Sáez López, UNED (España).
- Carolina Schmitt Nunes, Univ. Sta. Catarina (Brasil).
- Cassio Santos, Univ. Lisboa (Portugal)
- Juan José Sosa Alonso, Univ. La Laguna (España).
- Romero Tori, Univ. Sao Paulo (Brasil).
- Gemma Tur Ferrer, Univ. Islas Baleares (España).
- José Carlos Vázquez Parra, TEC Monterrey (México)

Redes Sociales (Social Networks)

- José Luis García Boyé
- Carla Netto
- Carolina Schmitt

Apoyo Técnico (Technical Assistance)

- Alexis Moreno-Pulido, Responsable de Biblioteca (UNED)

SopORTE OJS y Publicación digital

- Servicio Publicación y Difusión Digital - BIBLIOTECA, UNED

LA REVISTA IBEROAMERICANA DE EDUCACIÓN A DISTANCIA (RIED) SE ENCUENTRA INDIZADA ACTUALMENTE EN LAS SIGUIENTES BASES DE DATOS Y CATÁLOGOS:

BASES DE DATOS Y PLATAFORMAS DE EVALUACIÓN

- BASE. Bielefeld Academic Search Engine
- CAPES
- CARHUS Plus+
- CCHS-CSIC
- CEDAL (Instituto Latinoamericano de Comunicación Educativa (ILCE) de México)
- CIRC (Clasificación Integrada de Revistas Científicas)
- CiteFactor – Academic Scientific Journals
- CREDI- OEI (Centro de Recursos de la OEI)
- Crossref (Metadata Search)
- Dialnet (Alertas de Literatura Científica Hispana)
- DICE (Difusión y Calidad Editorial de Revistas)
- EI Compendex
- EBSCO. Fuente Académica Premier
- ERA. Educational Research Abstracts
- ERIH-Plus. European Reference Index for the Humanities and Social Sciences.
- EZB-Electronic Journals Library Genamics JournalSeek
- HEDBIB (International Bibliographic Database on Higher Education)
- IN-RECS (Índice de Impacto de Revistas Españolas de Ciencias Sociales)
- IRESIE (Índice de Revistas de Educación Superior e Investigación Educativa)
- ISOC (CSIC/CINDOC)
- JournalTOCs
- MIAR (Matriz para Evaluación de Revistas)
- ProQuest-CSA
- Psycodoc
- REDIB. Red Iberoamericana de Innovación y Conocimiento Científico
- REDALYC. Red de Revistas Científicas de América Latina y el Caribe, España y Portugal
- REDINED. Red de Información Educativa
- RESH - Revistas Españolas de Ciencias Sociales (CSIC/CINDOC)
- ResearchBib. Academic Resource Index
- Scopus
- Web of Science (SSCI)
- WEBQUALIS

DIRECTORIOS Y BUSCADORES

- DOAJ
- Dulcinea
- Google Scholar
- LATINDEX (Publicaciones Científicas Seriadadas de América, España y Portugal)
- Recolecta
- Sherpa/Romeo
- Scirus
- Ulrich's Periodicals (CSA)

PORTALES Y REPOSITARIOS ESPECIALIZADOS

- Actualidad Iberoamericana
- Asociación Internacional de Estudios en comunicación social
- CLARISE - Comunidad Latinoamericana Abierta Regional de Investigación Social y Educativa
- Educ.ar
- Enlaces educativos en español de la Universitat de València
- e-sPacio-UNED. Repositorio institucional de la UNED
- Institut Français de l'éducation
- Plataforma de revistas 360º
- Red Iberoamericana de Revistas de Comunicación y Cultura
- REDIAL & CEISAL
- Universia. Biblioteca de recursos

CATÁLOGOS DESTACADOS DE BIBLIOTECA

- 360grados
- British Library
- Buz
- Catàleg Col·lectiu de les Universitats de Catalunya
- Catálogo Colectivo de Publicaciones
- Periódicas Español CCPP
- Catálogo de la Biblioteca de Educación (Ministerio de Educación, Cultura y Deporte)
- Catálogo del CSIC (CIRBIC)
- CENDOC
- CIDE
- CISNE
- COMPLUDOC
- COPAC (Reino Unido)
- ICDL
- INRP
- IOE (Institute of Education. University of London)
- Library of Congress (LC)
- KINGS
- MIGUEL DE CERVANTES
- Observatorio de revistas científicas de Ciencias Sociales
- REBIUN
- SUDOC (Francia)
- UBUCAT
- UIB
- WORDLCAT (OCLC)
- ZDB (Alemania)

La Asociación Iberoamericana de Educación Superior a Distancia (AIESAD) es una entidad sin ánimo de lucro, constituida por universidades o instituciones de educación superior que imparten sus ofertas educativas en esta modalidad de enseñanza y promueve el estudio e investigación del modelo de enseñanza superior a distancia. *RIED. Revista Iberoamericana de Educación a Distancia* es el instrumento de la AIESAD para la difusión internacional de los avances en la investigación e innovación dentro del ámbito de la enseñanza y aprendizaje abiertos y a distancia.



RIED. Revista Iberoamericana de Educación a Distancia es una publicación científica que se edita semestralmente los meses de enero y julio. Promueve el intercambio institucional con otras revistas de carácter científico. La *RIED* no se hará responsable de las ideas y opiniones expresadas en los trabajos publicados. La responsabilidad plena será de los autores de los mismos.



"Los textos publicados en esta revista están sujetos a una licencia "Reconocimiento-No comercial 3.0" de Creative Commons. Puede copiarlos, distribuirlos, comunicarlos públicamente, siempre que reconozca los créditos de la obra (autor, nombre de la revista, instituciones editoras) de la manera especificada en la revista."



AIESAD

Ried

Revista Iberoamericana de
Educación a Distancia

VOL. 28 N° 2

Julio, 2025

Índice

ChatGPT, IA generativa y tecnología educativa: aportes y desafíos en la educación virtual

Estudios e investigaciones

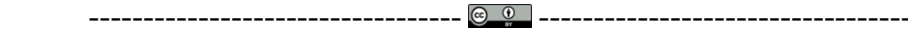
Asistentes ChatGPT en educación superior en línea y satisfacción del alumnado: un caso de estudio Miguel-Ángel Cabeza-Rodríguez	9
Agente de Inteligencia Artificial Generativa en investigación científica. Un análisis explicativo del aprendizaje en el aula Roberto Berrios Zepeda; Lorgia Márquez Mora	39
Diseño y análisis de narrativas enriquecidas con inteligencia artificial. Proyecto de transferencia UNED - UTMACH Javier Gil Quintana; Eduardo García Blázquez; Carmen Cantillo Valero; José Javier Hueso Romero	57
Inteligencia Artificial y chatbots para una educación superior sostenible: una revisión sistemática Eva Jiménez-García; Judit Ruiz-Lázaro; Sonia Martínez-Requejo; Sara Redondo-Duarte	81
Percepciones de futuros docentes y pedagogos sobre uso responsable de la IA. Un instrumento de medida Melchor Gómez-García; Julio Ruiz-Palmero; Moussa Boumadan-Hamed; Roberto Soto-Varela .	105
Del optimismo a la confianza: el impacto de ChatGPT en la confianza de los estudiantes en el aprendizaje asistido por IA Frederic Marimon; María Belén Arias-Valle; César Javier Coria-Augusto; Claudio Marcelo Larrea Arnau	131
Efectos de la conciencia y confianza en la disposición estudiantil para usar ChatGPT: modelo TAM-ECM integrado Sultan Hammad Alshammari; Abdullah Zaid Almankory; Muna Eid Alrashidi	155
Marcos normativos para una IA ética y confiable en la educación superior: estado de la cuestión María Obdulia González-Fernández; María Asunción Romero-López; Natalia Fátima Sgreccia; María José Latorre Medina	181
Desafíos de la Inteligencia Artificial generativa en educación superior: fomentando su uso crítico en el estudiantado Teresa Romeu Fontanillas; Marc Romero Carbonell; Montse Guitert Catasús; Pablo Baztán Quemada.....	209

IA generativa y pensamiento crítico en la educación universitaria a distancia: desafíos y oportunidades César Muñoz Martínez; Vanessa Roger-Monzo; Fernando Castelló-Sirvent	233
El estudiantado de secundaria ante la RV en materias STEM. Efecto de la variable de género David García-Marín; Ricardo Roncero Palomar ; Marina Santín; Victoria– Mora de la Torre	275
Impacto de variables sociodemográficas en la percepción del alumnado universitario sobre la educación en línea post-COVID-19 Laura Alexandra Arteaga-Briones; Luis Ángel Guamán-Lazo; Andrea Alexandra Azúa-Arteaga; Alexander López-Padrón	299
Adopción del aprendizaje en línea en la educación superior de China: la perspectiva de los docentes Chongya Ma; Farrukh Ijaz; Muhammad Fareed; Hood Laeeq	325
Propiedades psicométricas de una escala para medir estilos de enseñanza en línea en docentes universitarios Estanislao Casanova Sánchez; Armando Lozano Rodríguez; Beatriz Yesenia Ahumada Flores; Fernanda Inéz García Vázquez	353
Despertando el compromiso estudiantil: el poder transformador de H5P en la educación superior Fridel Julio Ramos-Azcuy; María Rodríguez-Gámez; Jeovanny Moisés Benavides-Bailón; María Margoth Bonilla-Jiménez; Ángel Enrique Arroba-Cárdenas	379
Aula invertida gamificada: aplicación de una estrategia didáctica para trabajar el pensamiento computacional en futuros docentes Jesús Carpena Arias; Francesc M. Esteve Mon	403

*Estudios e
investigaciones*

Asistentes ChatGPT en educación superior en línea y satisfacción del alumnado: un caso de estudio

ChatGPT assistants in online higher education and student satisfaction: a case study



Miguel-Ángel Cabeza-Rodríguez - Universidad Francisco de Vitoria, UFV (España)

RESUMEN

En este estudio se han analizado la percepción y satisfacción de los estudiantes sobre un asistente virtual basado en OpenAI ChatGPT 3.5, integrado en 21 asignaturas diferentes del campus virtual de una universidad en línea. Utilizando un enfoque metodológico mixto se recopiló información sobre una muestra de 391 estudiantes mediante el cuestionario validado COMUNICA, que incluyó cuatro constructos: eficiencia del asistente virtual, impacto en el aprendizaje, desarrollo de habilidades, y aspectos técnicos y de accesibilidad. El análisis incluyó estadística descriptiva, pruebas estadísticas inferenciales, análisis factorial exploratorio (AFE) y confirmatorio (AFC), complementado con un análisis cualitativo de comentarios del alumnado y docentes. Los resultados cuantitativos destacan que el género femenino valora más la eficacia del asistente que el masculino. El AFC confirmó que los factores se pueden agrupar bajo una única variable latente: la satisfacción del alumnado. Además, la eficiencia del asistente virtual resultó ser el factor más significativo en la percepción de satisfacción del alumnado, seguido por el impacto en el aprendizaje, desarrollo de habilidades y aspectos técnicos. El análisis cualitativo reveló percepciones mayoritariamente positivas, resaltando la utilidad del asistente en el aprendizaje, un interés en extender su uso a otras asignaturas y sugerencias de mejora en la precisión de las respuestas y la funcionalidad. Se concluye que los asistentes virtuales tienen un impacto positivo en la educación superior, optimizando el aprendizaje autónomo y la interacción educativa, aunque persisten desafíos técnicos y de diseño que limitan su potencial completo.

Palabras clave: inteligencia artificial; chatbot; ChatGPT; educación superior; enseñanza en línea; tecnología.

ABSTRACT

The perception and satisfaction of students about a virtual assistant based on OpenAI ChatGPT 3.5, integrated in 21 different subjects of the virtual campus of an online university, have been analyzed in this study. Using a mixed methodological approach, information was collected on a sample of 391 students using the validated COMUNICA questionnaire, which included four constructs: Virtual Assistant Efficiency, Learning Impact, Skill Development, and Technical and Accessibility Aspects. The analysis included descriptive statistics, inferential statistical tests, Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA), complemented by a qualitative analysis of student and teacher comments. The quantitative results highlight that the female gender values the effectiveness of the assistant more than the male gender. The CFA confirmed that the factors can be grouped under a single latent variable: student satisfaction. In addition, the efficiency of the virtual assistant was found to be the most significant factor in the perception of student satisfaction, followed by the impact on learning, skill development and technical aspects. The qualitative analysis revealed mostly positive perceptions, highlighting the usefulness of the assistant in learning, an interest in extending its use to other subjects and suggestions for improvement in the accuracy of answers and functionality. It is concluded that virtual assistants have a positive impact on higher education, optimizing autonomous learning and educational interaction, although technical and design challenges persist that limit their full potential.

Keywords: artificial intelligence; chatbot; ChatGPT; higher education; online teaching; technology.

INTRODUCCIÓN

Contexto general

En los últimos años la educación en línea ha experimentado un crecimiento significativo, impulsada por avances tecnológicos y la creciente demanda de modalidades de aprendizaje flexibles y accesibles (Pokhrel y Chhetri, 2021). La pandemia de COVID-19 aceleró la adopción de tecnologías digitales en la educación, destacando la necesidad de soluciones innovadoras para mantener la calidad y continuidad del aprendizaje a distancia mediante la movilidad virtual (Ruiz-Corbella y García-Aretio, 2023).

Los *chatbots* son agentes conversacionales, o “asistentes virtuales”, creados mediante programas informáticos que utilizan inteligencia artificial para simular conversaciones humanas; estos han emergido como herramientas prometedoras en el contexto educativo (Labadze et al., 2023). La capacidad de estos asistentes virtuales para brindar asistencia inmediata y personalizada los convierte en herramientas valiosas para mejorar la experiencia educativa en línea.

En este escenario, en el que existe una clara relación entre el aprendizaje autorregulado y el rendimiento académico (Cheng et al., 2023), los asistentes virtuales han sido utilizados para atender consultas frecuentes, ofrecer soporte técnico y proporcionar orientación académica básica, lo que contribuye a aliviar la carga de trabajo del personal docente y administrativo (Dwivedi et al., 2021; Thottoli et al., 2024)

Estado del arte

La aplicación de la inteligencia artificial en educación ha sido ampliamente analizada (Chen, Chen y Lin, 2020), así como más específicamente en educación superior (Wang et al., 2023). La implementación de estos asistentes virtuales en la educación superior ha generado interés y debate académico (Hwang y Chang, 2023; Pérez et al., 2020; Peters et al., 2024) debido a su potencial para mejorar la interacción entre estudiantes y sus instituciones educativas. Estas herramientas pueden ofrecer soporte en tiempo real, responder preguntas frecuentes, proporcionar recursos de aprendizaje personalizados y facilitar la comunicación, especialmente en cursos en línea donde la interacción humana puede ser limitada (Okonkwo y Ade-Ibijola, 2021).

La implementación de asistentes virtuales en entornos educativos ha demostrado múltiples beneficios según la bibliografía académica reciente (Kuhail et al., 2023; Pérez et al., 2020). En cuanto al acceso y disponibilidad, los asistentes virtuales proporcionan apoyo educativo sin limitaciones de horario, permitiendo a los estudiantes acceder a recursos y asistencia en cualquier momento. Esto resulta

particularmente valioso para estudiantes con horarios no convencionales, o en diferente zona horaria con respecto al docente (Winkler y Soellner, 2018).

La personalización del aprendizaje es otro aspecto que ha sido analizado (Pérez-Marín, 2021). Los asistentes virtuales pueden adaptar su interacción según el nivel de conocimiento y ritmo de aprendizaje de cada estudiante, con resultados que indican mejoras significativas en la retención de información cuando se emplea un enfoque personalizado. Permiten además una mayor eficiencia en la gestión docente, al reducir la carga laboral de los profesores, permitiéndoles enfocarse en tareas pedagógicas más complejas (Chen, Xie et al., 2020; Onal y Kulavuz-Onal, 2024).

Los asistentes virtuales educativos presentan una serie de desafíos en su implementación, que abarcan desde los aspectos técnicos hasta los pedagógicos y éticos (Hwang y Chang, 2023). Entre los desafíos técnicos se encuentra una funcionalidad técnica limitada para manejar variaciones en las preguntas, y la necesidad de entrenamiento continuo para mejorar su rendimiento (Okonkwo y Ade-Ibijola, 2021).

Otra dificultad técnica es la de imitar el lenguaje humano de manera auténtica y natural. Los asistentes virtuales deben ser capaces de construir conversaciones coherentes basadas en mensajes previos, lo cual es complejo; sin embargo, desde el lanzamiento de OpenAI ChatGPT (Adiguzel et al., 2023), se ha producido no solo un salto cualitativo, sino también cuantitativo en cuanto a la comercialización y popularización de los modelos de lenguaje extenso (*Large Language Models*, LLMs) (Han et al., 2021; Tamkin et al., 2021) que forman parte del aprendizaje profundo (*deep learning*) en redes neuronales (Perrotta y Selwyn, 2020), y este a su vez en el campo del aprendizaje automático (*machine learning*) que pertenece a la inteligencia artificial.

La efectividad de los asistentes virtuales en entornos educativos (Al-Emran et al., 2024) depende de varios factores clave: La eficiencia del asistente virtual, su impacto real en el proceso de aprendizaje, su contribución al desarrollo de habilidades específicas y los desafíos técnicos y de accesibilidad que presentan son determinantes en la satisfacción del alumnado. Comprender cómo estos factores influyen en la experiencia educativa es crucial para optimizar el diseño e implementación de estas herramientas en la educación en línea.

Objetivo

Este estudio tiene por objetivo analizar la satisfacción de los estudiantes de una universidad en línea con un asistente virtual implementado en el campus mediante una interfaz a OpenAI ChatGPT (Adiguzel et al., 2023; Tlili et al., 2023) en 21 aulas virtuales diferentes en 13 estudios universitarios diferentes. Se eligió ChatGPT por ser considerado el LLM con mayor facilidad para desarrollar la integración en el campus virtual y por su reconocimiento y aceptación por el público en general (Zhao et al., 2023).

A través de un cuestionario validado y confiable, y un análisis estadístico inferencial, se busca encontrar diferencias por género, grupo de edad o centro académico con relación a los cuatro constructos estudiados: eficiencia del asistente virtual, impacto en el aprendizaje, desarrollo de habilidades y aspectos técnicos y de accesibilidad.

También se busca explorar las relaciones de estos constructos con la satisfacción del alumnado mediante la confirmación de un modelo factorial propuesto.

Adicionalmente, se pretende analizar cualitativamente los comentarios proporcionados por el alumnado.

METODOLOGÍA

Fuente

La presente investigación utiliza una metodología mixta (mayoritariamente cuantitativa) no experimental, transversal, descriptiva e inferencial. La muestra de estudiantes ($n=391$) empleada en este estudio fue obtenida mediante un muestreo no probabilístico por conveniencia, facilitado por la colaboración de 21 docentes de 5 centros académicos de una universidad española en línea que dio acceso a una población de $N=3.419$ estudiantes. Un primer docente facilitó el acceso a su alumnado como piloto en un primer Proyecto de Innovación Docente Aplicada (PIDA) para, mediante un cuestionario estructurado, obtener información sobre el uso de un asistente virtual con inteligencia artificial integrado en el campus en línea de su asignatura. Este primer PIDA se titulaba “Primus: primer asistente virtual de IA en el grupo PROEDUCA”.

Tras el piloto, se contactó con los coordinadores académicos de todos los estudios universitarios para que distribuyesen en sus claustros una solicitud de colaboración en un PIDA titulado “Evaluación del impacto de un asistente virtual de inteligencia artificial en la experiencia de aprendizaje de estudiantes universitarios en un entorno educativo en línea”. De los docentes que respondieron, se seleccionaron los 20 con mayor número de alumnado, quienes proporcionaron acceso directo a su aula y sus estudiantes.

En cuanto a la metodología pedagógica y los medios utilizados, además de la implantación del asistente virtual en el campus, no se realizó ninguna gestión, control o coordinación adicional en las diferentes aulas, manteniendo cada docente su libertad de cátedra.

Instrumento

Cuestionario para el alumnado

Como instrumento de investigación para recabar información del alumnado se diseñó el cuestionario estructurado COMUNICA (Cuestionario de Opinión sobre el Manejo y Uso de Nuevas Interfaces Conversacionales Asistidas). Este instrumento, tal como se muestra en la Tabla 1, consta de cuatro ítems (Hair et al., 2019) por cada uno de sus cuatro constructos: eficiencia del asistente virtual, impacto en el aprendizaje, desarrollo de habilidades, y aspectos técnicos y de accesibilidad. Cada ítem incluye una escala Likert de 5 valores, siendo 1 equivalente a totalmente en desacuerdo y 5 a totalmente de acuerdo. Todos los ítems están formulados en afirmativo (Haladyna et al., 2002) pero, para evitar sesgos, se incluyeron la mitad de los ítems (los que tienen numeración 3 y 4) con un sentido opuesto al constructo, por lo que estos ítems requieren inversión de escala. Al final del cuestionario se incluyó un campo opcional de comentarios. El instrumento recibió una positiva evaluación de idoneidad por el Comité de Ética de Investigación de la Universidad en línea, con referencia UNIR CEI 101/2023.

Tabla 1
Ítems y constructos del cuestionario COMUNICA

Constructo	Identificador ítem	Ítem
Eficiencia del Asistente Virtual	E1	El asistente virtual de IA me ayudó a comprender mejor los conceptos de la asignatura.
	E2	El asistente virtual de IA proporcionó información y ejemplos que eran relevantes y aplicables a los temas de la asignatura.
	E3	Las respuestas proporcionadas por el asistente virtual de IA fueron difíciles de comprender.
	E4	El asistente virtual de IA proporcionó respuestas imprecisas o no relevantes a mis consultas relacionadas con la asignatura.
Impacto en el Aprendizaje	IM1	El asistente virtual de IA me animó a explorar y aprender de manera autónoma más allá de los contenidos de la asignatura.
	IM2	Utilizar el asistente virtual de IA contribuyó a que comprendiera mejor los conceptos de la asignatura.
	IM3	El asistente virtual de IA falló en aclarar mis dudas y preguntas relacionadas con la asignatura.
	IM4	El asistente virtual de IA me complicó la comprensión de conceptos de la asignatura.

Constructo	Identificador ítem	Ítem
Desarrollo de Habilidades	D1	El uso del asistente virtual de IA me impulsó a desarrollar y aplicar el pensamiento crítico sobre los temas de la asignatura.
	D2	Utilizar el asistente virtual de IA me ayudó a mejorar mis habilidades de investigación y búsqueda de información.
	D3	Interactuar con el asistente virtual de IA dificultó el desarrollo de mi habilidad para resolver problemas relacionados con la asignatura.
	D4	El asistente virtual de IA me dificultó la aplicación práctica de los conocimientos adquiridos en la asignatura.
Aspectos Técnicos y de Accesibilidad	A1	Acceder al asistente virtual de IA a través del campus fue sencillo y agradable.
	A2*	Cuando surgieron problemas técnicos con el asistente virtual de IA, se resolvieron rápidamente.
	A3	La interfaz del asistente virtual de IA era compleja y difícil de usar.
	A4	Encontré problemas técnicos en el uso del asistente virtual de IA.

*Nota**: El ítem A2 fue finalmente eliminado por criterio de optimización de confiabilidad.

Fuente: elaboración propia.

Dentro del primer PIDA el instrumento fue validado por 6 expertos, y posteriormente se realizó un piloto con 66 alumnos de la asignatura Física I del Grado en Física de la Universidad en primer cuatrimestre del curso 2023/2024, en el que el único investigador de ese primer PIDA era también docente. En el cuestionario de este piloto se obtuvo respuesta de 17 estudiantes.

Al comienzo del segundo cuatrimestre de ese mismo curso, se desarrolló el segundo PIDA, en el que su investigador principal contó con la colaboración de 20 docentes de 5 de los 6 centros de la Universidad (ver Tabla 2): la Facultad de Ciencias de la Salud, la Facultad de Derecho, la Facultad de Empresa y Comunicación, la Facultad de Educación y la Escuela Superior de Ingeniería y Tecnología. El único centro académico no presente fue la Facultad de Ciencias Sociales y Humanidades, de la que no surgieron docentes con un número suficiente de alumnado para ser seleccionados en el PIDA. En el cuestionario de este segundo PIDA se obtuvo respuesta de un total de 374 estudiantes.

Tabla 2
Centros académicos, estudios y asignaturas involucradas en la investigación del segundo PIDA

Nº Docente	Centro académico	Estudio universitario	Curso	Asignatura
1	Facultad de Educación	Grado en Maestro en Educación Primaria	3º	Conocimiento de las Ciencias Sociales en Educación Primaria
2	Facultad de Educación	Grado en Maestro en Educación Primaria	3º	Conocimiento de las Ciencias Sociales en Educación Primaria
3	Facultad de Educación	Grado en Maestro en Educación Primaria	3º	Didáctica de la Educación Física en Educación Primaria
4	Facultad de Educación	Grado en Magisterio de Educación Infantil	2º	Educación musical y su didáctica
5	Facultad de Educación	Grado en Maestro en Educación Primaria	3º	Morphosyntax and Semantics of the English Language
6	Facultad de Educación	Grado en Maestro en Educación Primaria	2º	Educación Musical y su Didáctica
7	Facultad de Educación	Grado en Maestro en Educación Primaria	2º	Didáctica de las Matemáticas en Educación Primaria
8	Facultad de Educación	Grado en Maestro en Educación Primaria	1º	Tecnologías de la Información Aplicadas a la Educación II
9	Facultad de Educación	Grado en Magisterio de Educación Infantil	2º	Tecnologías de la Información Aplicadas a la Educación I
10	Facultad de Educación	Grado en Magisterio de Educación Infantil	1º	Educación y Convivencia Dentro y Fuera del Aula
11	Facultad de Educación	Máster en formación del Profesorado de ESO, Bachillerato y FP	1º	Aprendizaje y desarrollo de la personalidad del adolescente
12	Facultad de Educación	Máster en formación del Profesorado de ESO, Bachillerato y FP	1º	Aprendizaje y desarrollo de la personalidad del adolescente

Nº Docente	Centro académico	Estudio universitario	Curso	Asignatura
13	Facultad de Empresa y Comunicación	Grado en Finanzas y Contabilidad	1º	Microeconomía
14	Facultad de Empresa y Comunicación	Grado en Administración y Dirección de Empresas	1º	Comunicación personal y organizacional
15	Facultad de Empresa y Comunicación	Máster en Marketing Digital	1º	Mobile Marketing
16	Escuela Superior de Ingeniería y Tecnología	Máster en Energías Renovables	1º	Energía Hidráulica y Geotérmica
17	Escuela Superior de Ingeniería y Tecnología	Máster en Energías Renovables	1º	Energía del Mar
18	Facultad de Derecho	Grado en Derecho	1º	Fundamentos de Derecho Público
19	Facultad de Derecho	Máster de acceso a la abogacía y la procura	1º	Práctica procesal contencioso administrativa
20	Facultad de Ciencias de la Salud	Máster en Neuropsicología Clínica	1º	Rehabilitación Neuropsicológica

Fuente: elaboración propia.

Los 20 docentes del segundo PIDA recibieron una formación del investigador principal sobre el asistente virtual, quien además se responsabilizó de la configuración del asistente en cada asignatura. Estos docentes explicaron el PIDA a su alumnado, informando sobre las limitaciones del asistente, incluyendo la posibilidad de sesgos y errores, y se les alentó a usar su juicio crítico y a buscar confirmación con el docente, o fuentes adicionales, cuando lo considerasen necesario. Al término de su docencia, facilitaron a su alumnado el cuestionario diseñado por el investigador principal, recogándose una muestra total de 374 respuestas en el segundo PIDA, que sumadas a las 17 del primer PIDA constituyeron 391 respuestas totales.

Los datos recabados del cuestionario se estudiaron mediante estadística descriptiva. Tras comprobar la consistencia interna de cada constructo (Cronbach, 1951), se analizó mediante estadística inferencial las posibles diferencias debidas al género, los grupos de edad, y los diferentes centros académicos en cada uno de los cuatro constructos.

Tras verificar la adecuación del muestreo para el análisis factorial, se realizó un análisis factorial exploratorio (AFE) como un paso previo y clave para identificar las estructuras subyacentes, ya que permite detectar variables latentes que reflejen constructos teóricos relevantes al tema de estudio (Hair et al., 2019). Este enfoque exploratorio fue necesario debido a la ausencia inicial de una hipótesis inicial clara sobre la estructura factorial, lo que requería una exploración de los datos antes de pretender validar un modelo teórico propuesto. La herramienta empleada para desarrollar el análisis estadístico descriptivo, el inferencial y el AFE fue IBM SPSS 29.0.2.

Los resultados del AFE se emplearon para determinar el número de dimensiones, lo que proporcionó una base sólida para el posterior análisis factorial confirmatorio (AFC) (Brown, 2015; Kline, 2023). La herramienta empleada para el desarrollo del AFC fue el lenguaje de programación R 4.4.2 con la biblioteca lavaan 0.6-19 (Rosseel, 2012).

Con el fin de minimizar el impacto de la falta de normalidad en los datos (Curran et al., 2003) se empleó en el AFC una estimación asintótica libre de distribución (*Asymptotic Distribution-Free*, ADF) con un *bootstrapping* de 5000 remuestreos (Yung y Bentler, 1994). Este estimador también es conocido como *Weighted Least Squares* (WLS) en algunas herramientas, como la biblioteca lavaan de lenguaje R.

Los cuatro índices de bondad de ajuste empleados en el AFC y sus respectivos valores de corte fueron: el índice de Tucker-Lewis $TLI > 0,95$ (Hu y Bentler, 1999), el *Standardized Root Mean Square Residual* SRMR $< ,1$ (Kline, 2023), el *Comparative Fit Index* CFI $> ,95$ (Hu y Bentler, 1999), y el *Root Mean Square Error of Approximation* RMSEA $< ,05$ (Browne y Cudeck, 1992).

Finalmente, se analizaron las respuestas al campo opcional de comentarios disponible al final del cuestionario, con el fin de estudiar de una forma cualitativa la opinión del alumnado. En un primer lugar, se evaluó el sentimiento general de los comentarios con técnicas de análisis de sentimientos de aprendizaje automático

(Nandwani y Verma, 2021) mediante el lenguaje de programación R 4.4.2 y la biblioteca *SentimentAnalysis*.

En segundo lugar, para capturar de forma contextualizada la complejidad de las percepciones, desde aspectos técnicos hasta implicaciones éticas, se empleó una metodología cualitativa con enfoque inductivo basado en la codificación abierta y la categorización temática. Los datos provinieron del corpus de 185 comentarios del alumnado, y se siguieron tres fases:

Codificación inicial: se realizó una lectura iterativa de los comentarios para identificar conceptos emergentes, utilizando la técnica de *grounded theory* (Glaser et al., 1968) para evitar sesgos predefinidos.

Agrupación temática: los códigos se organizaron en categorías y subcategorías mediante un proceso de comparación constante, priorizando la recurrencia y la relevancia pedagógica.

Validación cruzada: se verificó la coherencia interna de las categorías mediante la triangulación entre dos investigadores, asegurando que cada comentario se asignara a la subcategoría más precisa.

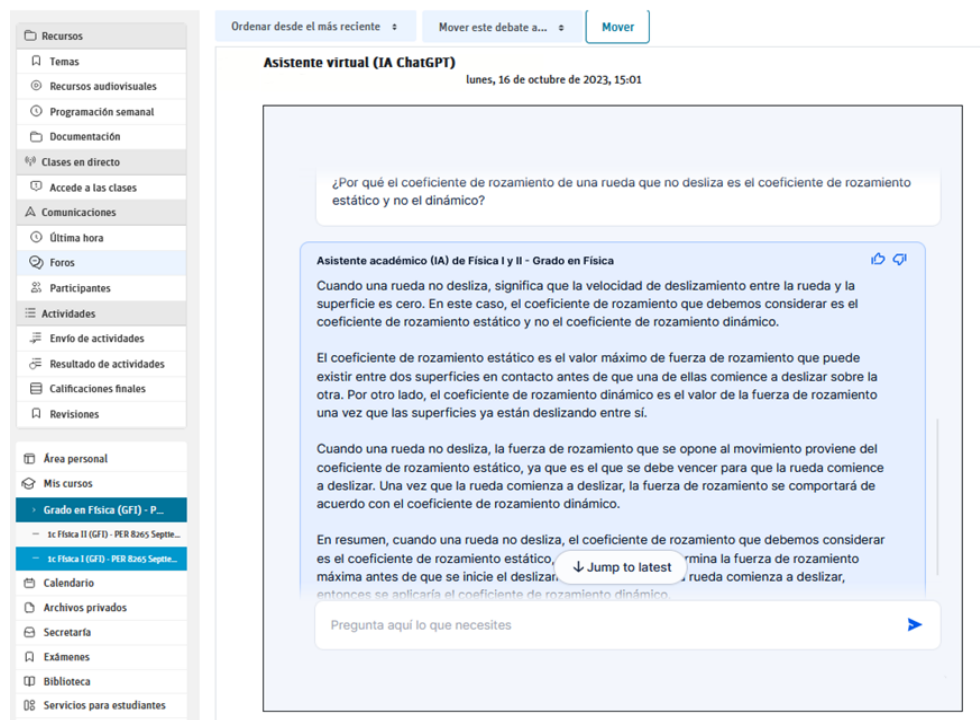
Asistente virtual

Pedagógicamente, el asistente virtual toma el papel de un ayudante y guía (*coach*), con una interacción elegida e iniciada por el alumnado (Pérez-Marín, 2021), que debe seleccionar un enlace en el campus para que le aparezca una ventana (un *iframe* en lenguaje HTML) integrada en la asignatura del campus en línea.

Se empleó la interfaz de programación de aplicaciones (*Application Programming Interface*, API) de un LLM, concretamente OpenAI ChatGPT 3.5. La interfaz para el alumnado quedó integrada en el campus en línea de sus respectivas asignaturas, ver Figura 1. Se ajustó la plantilla del asistente virtual para las asignaturas mediante la API del LLM, de forma que el hiperparámetro de temperatura del LLM, que controla su nivel de creatividad, fuese el mínimo, con el fin de reducir la posibilidad de alucinaciones en las respuestas al alumnado (OpenAI, s. f.).

Figura 1

Ejemplo de integración del asistente virtual en el campus en línea para una asignatura



Fuente: elaboración propia.

Se emplearon ajustes de ingeniería de petición (*prompt engineering*) (White et al., 2023) para configurar el LLM del asistente virtual en sus funciones hacia el campo específico de conocimiento de la correspondiente asignatura, así como también para minimizar la posibilidad de alucinaciones y fomentar al aprendizaje autodirigido (Chang et al., 2023). La redacción del texto en ingeniería de petición resulta clave en la calidad de las respuestas (Liu et al., 2024), y en este estudio se implantó la siguiente plantilla para todos los 21 asistentes virtuales.

“Eres un asistente virtual para el alumnado de la asignatura _____ en el grado/máster universitario _____.

Tu objetivo es enriquecer el aprendizaje del estudiante mediante respuestas precisas aplicables y correctamente fundamentadas, promoviendo al mismo tiempo la curiosidad intelectual y la investigación autodirigida.

Al responder, asegúrate de seguir las siguientes directrices:

1. *Basa tus respuestas en principios y teorías académicas establecidas, citando ejemplos relevantes de aplicaciones en la vida real cuando sea posible.*
2. *Prioriza la precisión y la relevancia académica, evitando especulaciones no fundamentadas.*
3. *Cuando sea posible, incluye estudios de caso relacionados con el tema en cuestión, para facilitar la comprensión y la aplicación del conocimiento que se presenta.*
4. *Cuando proporciones información estadística o datos, aclara el contexto y la fuente de dicha información, y sé transparente sobre la fiabilidad y las limitaciones de los datos.*
5. *Si una pregunta está fuera del alcance de tu entrenamiento o es demasiado específica para ser respondida con precisión, guía al usuario a que consulte con su docente.*
6. *Para preguntas que impliquen múltiples interpretaciones o enfoques, presenta un espectro de perspectivas académicas, resaltando la diversidad de opiniones y conclusiones.*
7. *Utiliza un lenguaje claro y formal, adecuado para un contexto académico, explicando cualquier jerga específica de la asignatura”.*

RESULTADOS

Análisis estadístico descriptivo

De los alumnos totales matriculados $N=3.419$ en las 21 asignaturas, se consiguieron respuestas de un 11 %, con una muestra ($n=391$) en la que la distribución de género fue 68 % género femenino, 32 % género masculino, y 0 % otros géneros.

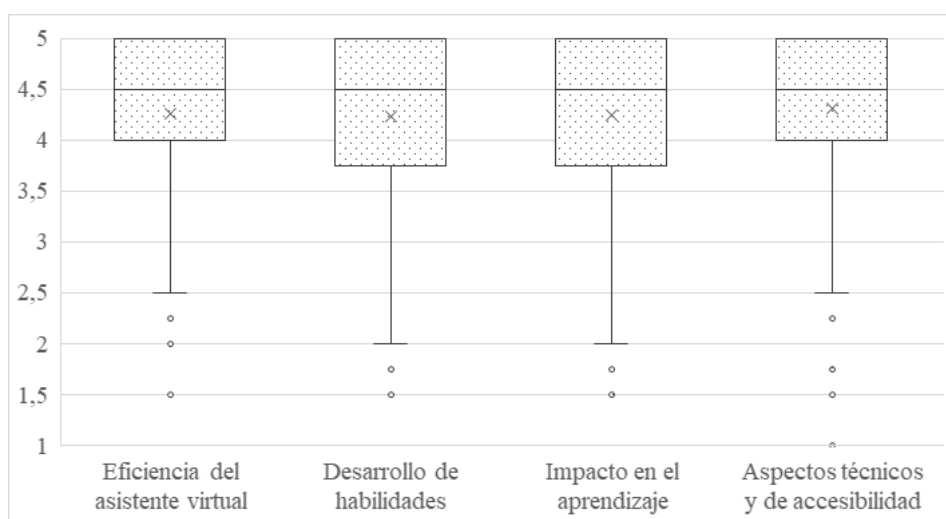
En cuanto a las diferentes generaciones de edad, la distribución fue la siguiente: nacidos antes de 1964 (*Baby Boomers*) un 1 %, entre 1965 y 1980 (generación X) un 25 %, entre 1981 y 1996 (generación Y) un 52 %, y después de 1997 (generación Z) un 23 %.

El alumnado de la muestra se encontró distribuido en 5 de los 6 centros académicos de la Universidad. En la Facultad de Educación un 57 %, en la Facultad de Empresa y Comunicación un 20 %, en la Facultad de Ciencias de la Salud un 10 %, en la Escuela Superior de Ingeniería y Tecnología un 7 %, y en la Facultad de Derecho un 6 %.

En la Figura 2 puede apreciarse que existe una clara falta de normalidad en la distribución de los resultados. Este último aspecto se corrobora mediante una prueba de normalidad (Anderson y Darling, 1952) en los cuatro constructos: AE=17,35, AD=15,08, AI=15,90, AA=18,20, $p<,001$. En un diagrama de cajas de los resultados del cuestionario para los cuatro constructos se aprecian distribuciones con cierta similitud; además, puede verse que las cuatro medianas resultan encontrarse en el valor 4,5.

Figura 2

Diagrama de cajas de los constructos



Fuente: elaboración propia.

Análisis estadístico inferencial

Análisis de constructos según género, grupos de edad y centro académico

Para el constructo de la Eficiencia del asistente virtual, existe una mayor valoración en el grupo femenino, $MD_f=4,50$, frente al masculino, $MD_m=4,25$, que mediante una prueba de la mediana para muestras independientes resulta ser significativa $p=,07$ ($\alpha=,05$). Además, de acuerdo con una prueba (no paramétrica de comparación de distribuciones independientes) de Mann-Whitney $U(n_f=271, n_m=120, df=1, p=,03)=14.142,5$ se puede rechazar la hipótesis nula de que ambas distribuciones sean iguales entre las dos categorías de género. El tamaño del efecto

resultó ser reducido (Cohen, 2013; Hattie, 2023) con $\delta=,22$ como medida de tamaño del efecto que cuantifica la diferencia entre las medias de dos grupos en unidades de desviación estándar, permitiendo evaluar la magnitud práctica de dicha diferencia.

Para este mismo constructo, y mediante pruebas de Kruskal-Wallis (no paramétrica de comparación de distribuciones independientes), no se encontraron diferencias estadísticamente significativas entre los grupos de edad $H(n_{BB}=3, n_X=97, n_Y=202, n_Z=89, df=3, p=,42)=2,69$, o entre centros académicos $H(n_{Educación}=224, n_{Empresa}=77, n_{Salud}=40, n_{Derecho}=22, n_{Ingeniería}=28, df=4, p=,21)=6,73$.

Con relación al constructo Desarrollo de habilidades y el género, no se encontraron diferencias estadísticamente significativas entre géneros según una prueba de Mann-Whitney $U(n_f=271, n_m=120, p=,89)=13,58$.

Para este mismo constructo, y mediante pruebas de Kruskal-Wallis, no se encontraron diferencias estadísticamente significativas entre los grupos de edad $H(n_{BB}=3, n_X=97, n_Y=202, n_Z=89, df=3, p=,16)=5,23$, o entre centros académicos $H(n_{Educación}=224, n_{Empresa}=777, n_{Salud}=40, n_{Derecho}=22, n_{Ingeniería}=28, df=4, p=,23)=6,42$.

Analizando el constructo del Impacto en el aprendizaje, no se encontraron diferencias estadísticamente significativas entre géneros según una prueba de Mann-Whitney $U(n_f=271, n_m=120, p=,28)=15,12$.

Para este mismo constructo, y mediante pruebas de Kruskal-Wallis, no se encontraron diferencias estadísticamente significativas entre los grupos de edad $H(n_{BB}=3, n_X=97, n_Y=202, n_Z=89, df=3, p=,16)=5,83$, o entre centros académicos $H(n_{Educación}=224, n_{Empresa}=777, n_{Salud}=40, n_{Derecho}=22, n_{Ingeniería}=28, df=4, p=,19)=7,59$.

Finalmente, para el constructo de Aspectos técnicos y accesibilidad, no se encontraron diferencias estadísticamente significativas entre géneros según una prueba de Mann-Whitney $U(n_f=271, n_m=120, p=,10)=15,72$.

Para este mismo constructo, y mediante pruebas de Kruskal-Wallis, no se encontraron diferencias estadísticamente significativas entre los grupos de edad $H(n_{BB}=3, n_X=97, n_Y=202, n_Z=89, df=3, p=,79)=1,52$, o entre centros académicos $H(n_{Educación}=224, n_{Empresa}=777, n_{Salud}=40, n_{Derecho}=22, n_{Ingeniería}=28, df=4, p=,07)=9,89$.

Validación, adecuación del muestreo y confiabilidad

Para la validación del instrumento se aplicó el Índice de Validez del Contenido de la Escala (S-CVI) para evaluar la pertinencia y representatividad de los ítems del cuestionario de acuerdo con los 6 expertos, alcanzándose un alto promedio S-CVI/Ave=99,31. Además, un S-CVI/UA del 95,83 % muestra que una proporción considerable de ítems recibió el acuerdo unánime de los expertos. Estos valores de índices superan los niveles comúnmente aceptados para confirmar la validez de contenido (Haynes et al., 1995; Polit et al., 2007).

La adecuación del muestreo para el análisis factorial ha sido verificada desde tres perspectivas. En primer lugar una prueba de Kaiser-Meyer-Olkin, obteniendo

un resultado $KMO=.946$ que puede considerarse excelente (Kaiser, 1974). En segundo lugar, se superó una prueba de esfericidad (Bartlett, 1954) $\chi^2=4543,25$, $gl=105$, $p<.001$. Y, en tercer lugar, se realizó un análisis de la matriz de correlaciones, resultando $r_{12}=.59$.

En este último análisis, la elección de una rotación Promax resultó consistente con los resultados de la matriz de correlación, donde la correlación es moderada, por lo que el uso de una rotación oblicua es adecuado para este análisis, dado que los factores no son ortogonales: la correlación moderada entre los componentes 1 y 2 (.59) puede reflejar una relación significativa entre ellos, y comparten aproximadamente el 35 % de su varianza, pero no son perfectamente redundantes.

La viabilidad del análisis de fiabilidad por alfa de Cronbach se soporta en dos aspectos. Primero, por la existencia de un autovalor $8,80>3$ (Yurdugül, 2008), y en segundo lugar, por un tamaño de muestra que supera el mínimo $141<n=391$ (Bonett, 2002). Este mínimo ha sido obtenido partiendo de un valor de alfa de Cronbach en la hipótesis nula $CA_0=.7$, un valor esperado $CA_1=.8$, una potencia de 90 %, una probabilidad de error tipo I $\alpha=.05$, y cuatro ítems por constructo.

Analizando la confiabilidad de cada uno de los cuatro constructos (ver Tabla 3) se aprecia que han obtenido alfas de Cronbach iguales o superiores a .8, indicando un nivel de consistencia interna adecuado para investigación básica (Nunnally, 1978). Adicionalmente, ninguno de ellos supera un alfa de Cronbach de .9, lo cual podría llegar a indicar un nivel innecesario de redundancia entre ítems (Streiner, 2003).

Tabla 3

Alfas de Cronbach de los constructos del cuestionario COMUNICA

Constructo	Ítems por constructo (inicial)	Alfa de Cronbach inicial	Ítem a eliminar para mejorar Alfa	Alfa de Cronbach final	Ítems por constructo (final)
Eficiencia del Asistente Virtual	4	,82	Ninguno	,82	4
Desarrollo de Habilidades	4	,82	Ninguno	,82	4
Impacto en el Aprendizaje	4	,80	Ninguno	,80	4
Aspectos Técnicos y de Accesibilidad	4	,79	A2	,80	3

Fuente: elaboración propia.

En el caso del constructo Aspectos Técnicos y de Accesibilidad, mejora su alfa de Cronbach de $CA_A=.79$ a $CA_A=.80$ cuando el ítem A2 es eliminado, por lo que se

prescinde de este ítem en el resto del análisis y se deja este constructo con los 3 ítems restantes. Este valor de 3 ítems por constructo es suficiente para la investigación (Hair et al., 2019) y compatible con el tamaño mínimo de muestra para 3 ítems por constructo, $159 < n = 391$ (Bonett, 2002).

Análisis Factorial Exploratorio

En un nuevo análisis de factores exploratorio de 15 ítems del instrumento, al no haber establecido previamente un marco teórico que pueda explicar correlaciones entre factores, se estudian las componentes principales con una rotación Promax, para obtener una estructura de factores simple e interpretable. De estos 15 ítems analizados se obtienen dos autovalores mayores que 1 (8,80 y 2,09) que permiten explicar una varianza total acumulada de 58,13 % y 71,27 % respectivamente.

Tabla 4
Matriz patrón del AFE

Ítem	Componente	
	1	2
A3	,85	-,04
A4	,83	-,03
D4	,83	,06
I4	,84	,04
E3	,82	-,02
D3	,80	,02
E4	,77	,02
I3	,74	-,02
D1	-,12	,92
E1	,02	,85
I1	-,02	,84
I2	,04	,85
D2	,01	,86
E2	,04	,85
A1	,23	,63

Fuente: elaboración propia.

La Tabla 4 muestra la matriz patrón, obtenida en 3 iteraciones, que ajusta las cargas factoriales considerando la correlación entre componentes por haber empleado una rotación oblicua Promax con normalización Kaiser. Se aprecian en ella

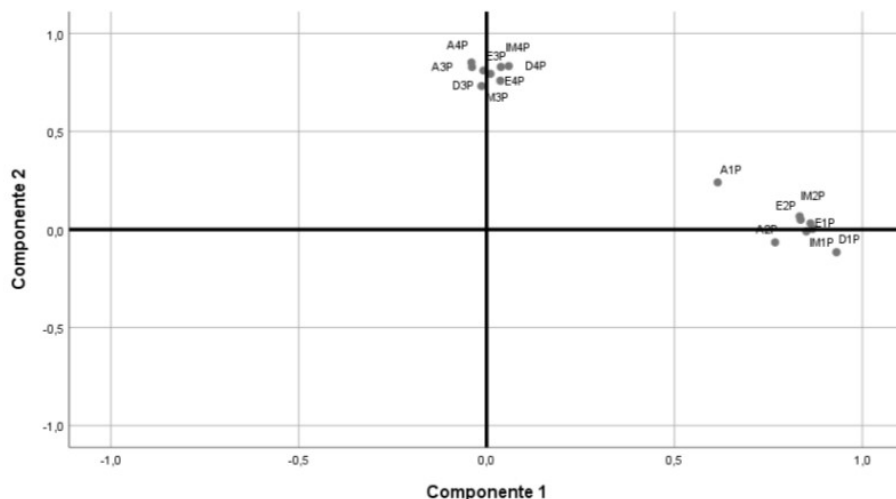
muy claramente dos componentes. En la primera componente el autovalor es igual a 8,80. En esta componente se encuentran todos los 8 ítems que originalmente se formularon con una asociación negativa a su respectivo constructo. Son fácilmente reconocibles por su numeración en el constructo respectivo, con identificador numérico 3 o 4. Los datos originales en escala Likert necesitaron una inversión para el tratamiento de datos.

En cuanto a la segunda componente, cuenta con autovalor igual a 2,09. En esta componente se encuentran los 7 ítems que se formularon con una asociación positiva a su respectivo constructo. Son fácilmente reconocibles por su numeración en el constructo con identificador numérico 1 o 2.

Se considera que esta estructura de doble componente separando preguntas con asociación positiva y otras con negativa, no es sino un efecto del método, es decir, ha sido producida por la tendencia de los ítems formulados de manera positiva o negativa a agruparse en factores separados no por su contenido real, sino por el formato en que se encuentran formulados. La Figura 3 muestra el gráfico de componentes en espacio rotado. Las respuestas a ítems negativos y positivos tienden a cargar componentes distintos, incluso cuando ambos tipos de ítems pretendan medir iguales constructos (Marsh, 1996; Roszkowski y Soven, 2010). Debido a este efecto, se estima que se ha obtenido una estructura factorial con un reflejo especular de un mismo componente, sin más que una dimensión real diferenciada que permite explicar un total de 71,27 % de la varianza.

Figura 3

Gráfico de componente en espacio rotado

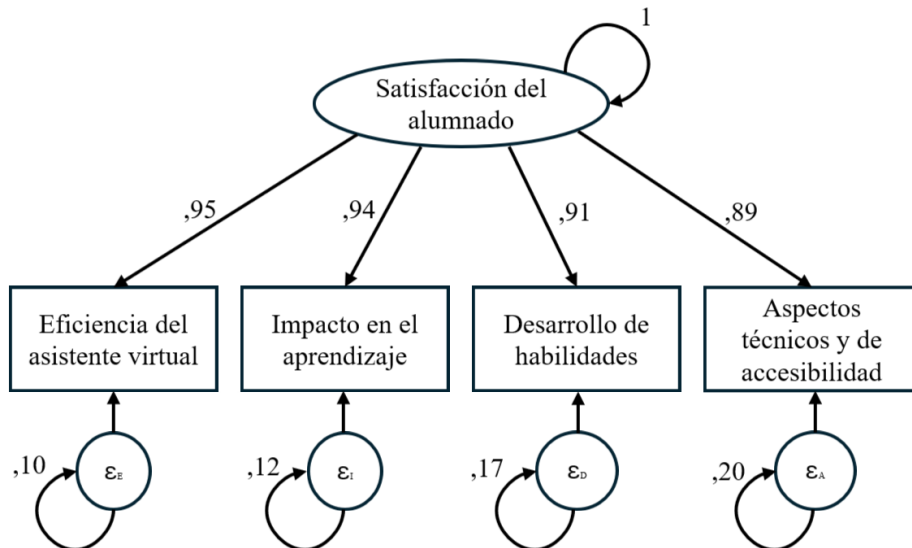


Fuente: elaboración propia.

Análisis Factorial Confirmatorio

En la Figura 4 se muestra el modelo desarrollado en el AFC, en el que se ha definido como variable latente principal la “Satisfacción del alumnado”, que se refleja a través de las cuatro variables observadas: eficiencia del asistente virtual, impacto en el aprendizaje, desarrollo de habilidades y aspectos técnicos y de accesibilidad. Cada uno de estos factores se ha encontrado que está asociado con una carga factorial (.95, .94, .91 y .89 respectivamente), lo que muestra la fuerza de la relación entre la variable latente y los indicadores observados. Además, cada indicador incluye un término de error asociado (ε_E , ε_I , ε_D y ε_A), con varianzas de error específicas (.10, .12, .17 y .20 respectivamente).

Figura 4
Modelo del Análisis Factorial Confirmatorio



Fuente: elaboración propia.

Por otro lado, se supera el tamaño mínimo de muestra para realizar este AFC, $30 < n = 391$ (Wolf et al., 2013), con una variable latente y cuatro factores con cargas $\lambda \geq ,80$.

En cuanto a los índices de bondad de ajuste, el índice de Tucker-Lewis obtenido es mayor que 0,95 (Hu y Bentler, 1999) con $TLI = 0,97$. El valor obtenido del $SRMR = ,05$, es inferior a ,09 (Hu y Bentler, 1999) y ,1 (Kline, 2023). En el *Comparative Fit Index*, resultó un valor $CFI = ,98 > ,95$ (Hu y Bentler, 1999). Finalmente, se obtuvo un índice $RMSEA = 0,04 < 0,05$ (Browne y Cudeck, 1992). Todos los índices indicados confirman la corrección del ajuste.

Aunque el valor de $\chi^2/gl = 1,63 < 2$ con $p = ,16 > ,05$ puede parecer correcto, solo se indica a modo meramente informativo, ya que su uso como índice de bondad de ajuste es desaconsejable (Brown, 2015; Wheaton, 1987).

Análisis cualitativo

Del alumnado que participó respondiendo el cuestionario ($n = 391$) un 47 % ($n_{com} = 185$) incluyó sus comentarios y sugerencias en un campo opcional reservado para el efecto al final de este. Se analizaron los sentimientos (Nandwani y Verma, 2021) de los 185 comentarios con aprendizaje automático mediante la biblioteca

SentimentAnalysis (Feuerriegel y Pröllochs, 2023) en el lenguaje de programación R 4.4.2, empleando el índice *SentimentQDAP*. El resultado fue mayoritariamente positivo con un índice *SentimentQDAP*=66 %, donde se ha convertido su rango original de salida [-1,1] a [0,100], siendo el 0 % totalmente negativo, y el 100 % totalmente positivo.

Mediante un enfoque inductivo basado en la codificación abierta y la categorización temática, se obtuvieron las categorías y subcategorías que se muestran en la Tabla 5, que incluye un ejemplo representativo de cada una de ellas. Exceptuando la categoría de “Otros comentarios” que incluye agradecimientos sin evaluación crítica o comentarios genéricos neutros, las tres subcategorías y categorías más frecuentes fueron: “Comentarios positivos” de “Percepción general” (17 %), el “Apoyo académico” en el “Impacto en el aprendizaje” (10 %), y las peticiones de “Integración en otras áreas” como “Sugerencias de mejora” (8 %).

Tabla 5
Categorías y subcategorías de los comentarios

Categoría	Subcategoría	Número de comentarios	Porcentaje	Ejemplo representativo
1. Percepción General	1.1. Comentarios positivos	32	17 %	“Excelente herramienta, felicidades”.
	1.2. Comentarios negativos	7	4 %	“El chatbot es completamente inútil”.
	1.3. Comentarios neutros	5	3 %	“Es útil, pero no sustituye al profesor”.
2. Funcionalidad y usabilidad	2.1. Facilidad de uso	12	6 %	“Intuitivo y sencillo de manejar”.
	2.2. Precisión de las respuestas	9	5 %	“A veces da información errónea o confusa”.
	2.3. Errores técnicos	4	2 %	“No permite guardar un historial de sesiones”.
3. Impacto en el aprendizaje	3.1. Apoyo académico	18	10 %	“Clarifica conceptos científicos complejos”.
	3.2. Limitaciones	8	4 %	“No resuelve problemas prácticos, solo teóricos”.

Categoría	Subcategoría	Número de comentarios	Porcentaje	Ejemplo representativo
4. Comparación con otras herramientas	4.1. Asistente virtual frente a ChatGPT4	6	3 %	"ChatGPT4 da respuestas más completas".
	4.2. Ventajas competitivas	3	2 %	"Está adaptada al temario específico".
5. Sugerencias de mejora	5.1. Ampliación de funcionalidades	10	5 %	"Incluir gráficos o enlaces a recursos externos".
	5.2. Capacitación y soporte	5	3 %	"Ofrecer tutoriales para usar la herramienta".
	5.3. Integración en otras áreas	14	8 %	"Implementarla en todas las asignaturas".
6. Innovación educativa	6.1. Aceptación de la IA	9	5 %	"Es el futuro de la educación".
	6.2. Resistencia o preocupaciones	3	2 %	"Fomenta la vaguería en algunos estudiantes".
7. Otros comentarios	-	40	22 %	"Gracias por la herramienta" (agradecimientos sin evaluación crítica) o "Todo bien" (comentarios genéricos neutros).

Fuente: elaboración propia.

DISCUSIÓN

Análisis de los resultados

Análisis estadístico inferencial

Se ha encontrado de forma estadísticamente significativa una mayor valoración de la eficiencia del asistente virtual por parte del género femenino. Las diferencias en expectativas y percepciones entre géneros pueden haber influido en la evaluación de la eficiencia del asistente virtual, sin embargo, el resultado es el opuesto al indicado por algunos autores (Cai et al., 2017), para quienes el género masculino muestra

una actitud más favorable hacia el uso de la tecnología que el femenino, aunque con diferencias menores.

Tras encontrar una única componente en el AFE, y proponer como variable latente la Satisfacción del alumnado, se puede afirmar que todos los índices de bondad de ajuste del AFC indican que el modelo propuesto es aceptable, existiendo por lo tanto una convergencia de este con los datos. El modelo del AFC sugiere que la “Eficiencia del asistente virtual” es el indicador más significativo de la satisfacción, mientras que los “Aspectos técnicos y de accesibilidad” tienen la menor relación. Este modelo permite evaluar la contribución relativa de la satisfacción en cada factor según los datos y el modelo propuesto.

Las cargas de factores encontradas se encuentran dentro del rango ,5 y ,95 (Bagozzi y Yi, 1989), en concreto $,89 < \lambda < ,95$ $p < ,001$, lo cual sugiere que los factores medidos están explicados en su mayor parte por la variable latente, la satisfacción del alumnado. Es recomendable (Hair et al., 2019) que las cargas factoriales estandarizadas sean mayores a ,7; esto implica que la varianza explicada por el factor (comunalidad) sea al menos 50 % de la varianza total de la variable, lo que a su vez significa que la unicidad (varianza de error) sea $\epsilon < ,5$, indicando que los factores observados están adecuadamente representados por la variable latente (Hair et al., 2019). Estos criterios aseguran que la variable latente explica una proporción significativa de la varianza en las variables observadas, confirmando la validez del modelo.

Análisis cualitativo

El análisis de sentimientos mediante aprendizaje de máquina arrojó un valor mayoritariamente positivo (66m%). Para obtener una visión más completa, el análisis inductivo muestra que, en cuanto a las sugerencias y opiniones expresadas por el alumnado en el campo de comentarios del cuestionario, entre los tres aspectos más comentados, se encontró en primer lugar la utilidad y contribución al aprendizaje. La mayoría de los estudiantes destacan que el asistente virtual de IA es una herramienta muy útil que les ha ayudado a comprender conceptos complejos, resolver dudas rápidamente y mejorar su proceso de aprendizaje en las asignaturas.

El segundo aspecto más comentado fue el deseo de implementación en otras asignaturas. Muchos alumnos expresaron su interés y la necesidad de que el asistente fuese incorporado en otras materias, considerando que su uso sería beneficioso y aumentaría la eficacia de su estudio en general.

El tercer aspecto más frecuente incluyó limitaciones observadas y sugerencias de mejora. Algunos estudiantes señalan áreas de mejora para el asistente virtual, como la mejora, precisión y profundidad de las respuestas, la incorporación de ejemplos más claros, una mejor orientación sobre cómo utilizar la herramienta efectivamente, y la capacidad de mantener el historial de conversaciones, una funcionalidad no existente en la implantación realizada.

En algunos casos también expresaron que la eficacia del asistente virtual se vio limitada, porque echaron en falta una mayor orientación sobre cómo utilizar la herramienta. Adicionalmente, hubo quienes opinaron que el asistente virtual no sustituyó el valor del pensamiento crítico ni la interacción con los docentes.

Los comentarios del alumnado reflejan en general una percepción mayoritariamente positiva del asistente virtual, como complemento en el proceso educativo, reconociendo sus beneficios, pero también señalando áreas donde puede ser perfeccionado para maximizar su potencial en el aprendizaje. Este sentimiento mayoritariamente positivo en los comentarios se ve reflejado en el análisis cuantitativo de sentimientos, y en el análisis cualitativo de los 185 comentarios recibidos en el cuestionario COMUNICA.

Aunque los comentarios netamente positivos (17 %) superan claramente a los negativos (4 %), aparecen aceptaciones con reservas por percibirse como un complemento, no un reemplazo, de la interacción humana. (“no sustituye al profesor”). Además, un 10 % de los comentarios estuvieron relacionados con el impacto en el aprendizaje en el apoyo académico que brindó el asistente virtual. Finalmente, la mayoría de las sugerencias (8 % de todos los comentarios) solicitaban la ampliación de la existencia del asistente a todas las asignaturas del estudio del alumnado. Globalmente, el análisis cualitativo subraya la necesidad de equilibrar avances tecnológicos con pedagogías centradas en el alumnado, asegurando que la IA en educación priorice la rigurosidad académica y la accesibilidad.

Limitaciones de la investigación

Este estudio emplea un muestreo no probabilístico por conveniencia, lo que limita la generalización de los resultados. Para futuros trabajos, se considera emplear muestreo probabilístico y diseñar estrategias para equilibrar la muestra por género, que, aunque representativa de los estudios, se encuentra desbalanceada con un 68 % de mujeres. Adicionalmente, un análisis longitudinal enriquecería la comprensión del efecto sostenido de la herramienta en el aprendizaje y la satisfacción del alumnado. Por ello, se plantea como línea de investigación futura el seguimiento en diversas etapas, lo que ciertamente ampliaría la robustez de los hallazgos, y ofrecería una visión dinámica.

No se pudieron implementar funciones que permitiesen un posterior análisis de aprendizaje (*Learning Analytics*) (Chang et al., 2023), por lo que en futuras implementaciones se debería intentar añadir esta característica con el fin de poder extraer más información sobre el empleo de la herramienta por parte del alumnado, como número de sesiones, duración de sesión, etc.

Durante el segundo cuatrimestre del curso 2023/2024, en la ejecución del segundo PIDA con los 20 docentes involucrados, ya se encontró disponible la versión 4 de ChatGPT (Peters et al., 2024), por lo que algunos alumnos realizaron comparaciones y encontraron que en la mayoría de las ocasiones la respuesta con

el acceso directo a esta versión resultaba más completa que con la versión 3.5 implantada en el campus virtual para el asistente.

No se incorporaron mecanismos específicos para controlar otras variables externas (por ejemplo, condiciones socioeconómicas, experiencia previa con tecnologías digitales u otros factores ambientales) que podrían influir en la percepción del asistente virtual. En futuras investigaciones se podría estudiar diseños que permitan el control o la inclusión de un mayor número de covariables, así como el uso de métodos estadísticos (por ejemplo, análisis multivariado o modelización estructural) para mitigar el impacto de posibles variables externas.

Los 20 docentes del segundo PIDA fueron contactados en entrevistas no estructuradas, pero no se desarrolló formalmente una investigación cualitativa o cuantitativa sobre sus opiniones debido a limitaciones temporales. En estas entrevistas principalmente se destacó la necesidad de una mayor personalización en la configuración del asistente virtual para adaptarlo a las características particulares de cada asignatura, con el fin de mejorar la precisión de sus respuestas.

Para una investigación futura, podría resultar aconsejable la creación de un cuestionario como instrumento estructurado adicional para recoger las percepciones de los docentes involucrados en el desarrollo del asistente en sus aulas, lo cual permitirá obtener una valiosa información adicional a la del alumnado (Kasneci et al., 2023).

Conclusiones

Este estudio confirma la relevancia y eficacia de los asistentes virtuales basados en inteligencia artificial en el ámbito de la educación superior en línea, destacando tanto sus beneficios como los retos que presenta su implementación, no solo en un determinado campo de conocimiento (Polverini y Gregorcic, 2024; Taani y Alabidi, 2024; Vierhauser et al., 2024; Wardat et al., 2023) sino para 21 diferentes titulaciones universitarias de los diferentes centros en una universidad en línea.

En primer lugar, los hallazgos cuantitativos demuestran que el género femenino ha valorado más la eficiencia del asistente virtual. El género masculino y el femenino difieren significativamente en cómo perciben la utilidad y la facilidad de uso de las herramientas tecnológicas, y en concreto, las mujeres tienden a ser más influenciadas por factores sociales y a tener expectativas diferentes sobre la utilidad de la tecnología en comparación con los hombres, lo cual puede ser la causa de esa mayor valoración de la eficiencia del asistente virtual por parte del género femenino. Mientras que el género masculino solo considera factores relacionados con la productividad, el femenino toma en cuenta diversos aspectos, incluida también la productividad, cuando se toman decisiones sobre la adopción y el uso de tecnología (Venkatesh y Morris, 2000).

Los resultados del AFC subrayan que la eficiencia del asistente virtual es el factor más significativo en la percepción de satisfacción del alumnado. Este aspecto está

estrechamente vinculado con la capacidad del asistente virtual para proporcionar respuestas precisas y relevantes, ello refuerza su papel como herramienta de apoyo en el aprendizaje autónomo, lo que concuerda con Følstad y Brandtzæg (2017).

Asimismo, los resultados del análisis factorial confirmatorio muestran que los factores asociados al impacto en el aprendizaje, el desarrollo de habilidades, y los aspectos técnicos y de accesibilidad también contribuyen de manera significativa a la percepción de utilidad del asistente virtual, aunque en menor medida. Estas dimensiones reflejan tanto el potencial pedagógico de los asistentes virtuales como la importancia de su diseño técnico para garantizar una experiencia de usuario óptima (Bahrini et al., 2023).

Sin embargo, además de analizar el potencial de los asistentes virtuales (Grassini, 2023) technological advancements, especially artificial intelligence (AI este trabajo también identifica novelmente limitaciones clave, como las diferencias significativas en la percepción de eficiencia del asistente entre estudiantes de diferentes géneros, así como la menor correlación significativa de los aspectos técnicos y de accesibilidad con la satisfacción general. Esto indica la necesidad de seguir explorando cómo optimizar estas herramientas para garantizar una experiencia equitativa y accesible para todos los usuarios.

Por otro lado, el análisis cualitativo es consistente con estudios previos que destacan la relación entre herramientas de inteligencia artificial y la mejora en el aprendizaje autónomo (Ait Baha et al., 2024).

Finalmente, este estudio contribuye a la bibliografía existente proporcionando evidencia empírica sobre la implementación y el impacto de asistentes virtuales en contextos educativos (Motlagh et al., 2023). Los hallazgos tienen implicaciones prácticas para instituciones académicas, desarrolladores de tecnología educativa y responsables de políticas, quienes podrán tomar decisiones más fundamentadas sobre cómo integrar estas herramientas de manera efectiva y sostenible en los procesos de enseñanza y aprendizaje (Dempere et al., 2023).

Conflicto de intereses

No se declara ningún potencial conflicto de intereses con respecto a la investigación, autoría y/o publicación de este artículo.

Financiación

Ambos Proyectos de Innovación Docente Aplicada asociados a esta investigación fueron financiados por la Universidad Internacional de la Rioja durante el curso 2023/2024, recibiendo por ellos el segundo premio en la VI edición de los “Premios a las mejores prácticas en el aula virtual” de la Universidad Internacional de la Rioja con fecha 10 de julio de 2024.

REFERENCIAS

- Adiguzel, T., Kaya, M. H. y Cansu, F. K. (2023). Revolutionizing education with AI: Exploring the transformative potential of ChatGPT. *Contemporary Educational Technology*, 15(3), ep429. <https://doi.org/10.30935/cedtech/13152>
- Ait Baha, T., El Hajji, M., Es-Saady, Y. y Fadili, H. (2024). The impact of educational chatbot on student learning experience. *Education and Information Technologies*, 29(8), 10153-10176. <https://doi.org/10.1007/s10639-023-12166-w>
- Al-Emran, M., AlQudah, A. A., Abbasi, G. A., Al-Sharafi, M. A. y Iranmanesh, M. (2024). Determinants of Using AI-Based Chatbots for Knowledge Sharing: Evidence From PLS-SEM and Fuzzy Sets (fsQCA). *IEEE Transactions on Engineering Management*, 71, 4985-4999. <https://doi.org/10.1109/TEM.2023.3237789>
- Anderson, T. W. y Darling, D. A. (1952). Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. *The Annals of Mathematical Statistics*, 23(2), 193-212. <https://doi.org/10.1214/aoms/1177729437>
- Bagozzi, R. P. y Yi, Y. (1989). The Degree of Intention Formation as a Moderator of the Attitude-Behavior Relationship. *Social Psychology Quarterly*, 52(4), 266. <https://doi.org/10.2307/2786991>
- Bahrini, A., Khamoshifar, M., Abbasimehr, H., Riggs, R. J., Esmaceli, M., Majdabadkohne, R. M. y Pasehvar, M. (2023). ChatGPT: Applications, Opportunities, and Threats. *2023 Systems and Information Engineering Design Symposium (SIEDS)*, 274-279. <https://doi.org/10.1109/SIEDS58326.2023.10137850>
- Bartlett, M. S. (1954). A Note on the Multiplying Factors for Various χ^2 Approximations. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 16(2), 296-298. <https://doi.org/10.1111/j.2517-6161.1954.tb00174.x>
- Bonett, D. G. (2002). Sample Size Requirements for Testing and Estimating Coefficient Alpha. *Journal of Educational and Behavioral Statistics*, 27(4), 335-340. <https://doi.org/10.3102/10769986027004335>
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research*. The Guilford Press.
- Browne, M. W. y Cudeck, R. (1992). Alternative Ways of Assessing Model Fit. *Sociological Methods & Research*, 21(2), 230-258. <https://doi.org/10.1177/0049124192021002005>
- Cai, Z., Fan, X. y Du, J. (2017). Gender and attitudes toward technology use: A meta-analysis. *Computers & Education*, 105, 1-13. <https://doi.org/10.1016/j.compedu.2016.11.003>
- Chang, D. H., Lin, M. P.-C., Hajian, S. y Wang, Q. Q. (2023). Educational Design Principles of Using AI Chatbot That Supports Self-Regulated Learning in Education: Goal Setting, Feedback, and Personalization. *Sustainability*, 15(17), 12921. <https://doi.org/10.3390/su151712921>
- Chen, L., Chen, P. y Lin, Z. (2020). Artificial Intelligence in Education: A Review. *IEEE Access*, 8, 75264-75278. <https://doi.org/10.1109/ACCESS.2020.2988510>
- Chen, X., Xie, H., Zou, D. y Hwang, G.-J. (2020). Application and theory gaps during the rise of Artificial Intelligence in Education. *Computers and Education: Artificial Intelligence*, 1, 100002. <https://doi.org/10.1016/j.caeai.2020.100002>
- Cheng, Z., Zhang, Z., Xu, Q., Maeda, Y. y Gu, P. (2023). A meta-analysis addressing the relationship between self-regulated learning strategies and academic performance in online higher education. *Journal of Computing in Higher Education*

- Education. <https://doi.org/10.1007/s12528-023-09390-1>
- Cohen, J. (2013). *Statistical Power Analysis for the Behavioral Sciences* (o ed.). Routledge. <https://doi.org/10.4324/9780203771587>
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297-334. <https://doi.org/10.1007/BF02310555>
- Curran, P. J., Bollen, K. A., Chen, F., Paxton, P. y Kirby, J. B. (2003). Finite Sampling Properties of the Point Estimates and Confidence Intervals of the RMSEA. *Sociological Methods & Research*, 32(2), 208-252. <https://doi.org/10.1177/0049124103256130>
- Dempere, J., Modugu, K., Hesham, A. y Ramasamy, L. K. (2023). The impact of ChatGPT on higher education. *Frontiers in Education*, 8, 1206936. <https://doi.org/10.3389/feduc.2023.1206936>
- Dwivedi, Y. K., Hughes, D. L., Coombs, C., Constantiou, I., Duan, Y., Edwards, J. S., Gupta, B., Lal, B., Misra, S., Prashant, P., Raman, R., Rana, N. P., Sharma, S. K. y Upadhyay, N. (2021). Impact of COVID-19 pandemic on information management research and practice: Transforming education, work and life. *International Journal of Information Management*, 55, 102211. <https://doi.org/10.1016/j.ijinfomgt.2020.102211>
- Feuerriegel, S. y Pröllochs, N. (2023). *Sentiment Analysis R Package* (Versión 1.3-5) [R]. <https://cran.r-project.org/web/packages/SentimentAnalysis/vignettes/SentimentAnalysis.html><https://doi.org/10.1016/j.ijinfomgt.2020.102211>
- Følstad, A. y Brandtzæg, P. B. (2017). Chatbots and the new world of HCI. *Interactions*, 24(4), 38-42. <https://doi.org/10.1145/3085558>
- Glaser, B. G., Strauss, A. L. y Strutzel, E. (1968). The Discovery of Grounded Theory; Strategies for Qualitative Research: *Nursing Research*, 17(4), 364. <https://doi.org/10.1097/00006199-196807000-00014>
- Grassini, S. (2023). Shaping the Future of Education: Exploring the Potential and Consequences of AI and ChatGPT in Educational Settings. *Education Sciences*, 13(7), 692. <https://doi.org/10.3390/educsci13070692>
- Hair, J. F., Black, W. C., Babin, B. J. y Anderson, R. E. (2019). *Multivariate data analysis* (Eighth edition). Cengage.
- Haladyna, T. M., Downing, S. M. y Rodriguez, M. C. (2002). A Review of Multiple-Choice Item-Writing Guidelines for Classroom Assessment. *Applied Measurement in Education*, 15(3), 309-333. https://doi.org/10.1207/S15324818AME1503_5
- Han, X., Zhang, Z., Ding, N., Gu, Y., Liu, X., Huo, Y., Qiu, J., Yao, Y., Zhang, A., Zhang, L., Han, W., Huang, M., Jin, Q., Lan, Y., Liu, Y., Liu, Z., Lu, Z., Qiu, X., Song, R., ... Zhu, J. (2021). Pre-trained models: Past, present and future. *AI Open*, 2, 225-250. <https://doi.org/10.1016/j.aiopen.2021.08.002>
- Hattie, J. (2023). *Visible Learning: The Sequel: A Synthesis of Over 2,100 Meta-Analyses Relating to Achievement* (1.a ed.). Routledge. <https://doi.org/10.4324/9781003380542>
- Haynes, S. N., Richard, D. y Kubany, E. S. (1995). Content validity in psychological assessment: A functional approach to concepts and methods. *Psychological assessment*, 7(3), 238. <https://doi.org/10.1037/1040-3590.7.3.238>
- Hu, L. y Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1-55. <https://doi.org/10.1080/10705519909540118>
- Hwang, G.-J. y Chang, C.-Y. (2023). A review of opportunities and challenges of chatbots in education. *Interactive*

- Learning Environments*, 31(7), 4099-4112. <https://doi.org/10.1080/10494820.2021.1952615>
- Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39(1), 31-36. <https://doi.org/10.1007/BF02291575>
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Kline, R. B. (2023). *Principles and Practice of Structural Equation Modeling* (Fifth edition). Guilford.
- Kuhail, M. A., Alturki, N., Alramlawi, S. y Alhejori, K. (2023). Interacting with educational chatbots: A systematic review. *Education and Information Technologies*, 28(1), 973-1018. <https://doi.org/10.1007/s10639-022-11177-3>
- Labadze, L., Grigolia, M. y Machaidze, L. (2023). Role of AI chatbots in education: Systematic literature review. *International Journal of Educational Technology in Higher Education*, 20(1), 56. <https://doi.org/10.1186/s41239-023-00426-1>
- Liu, X., Zheng, Y., Du, Z., Ding, M., Qian, Y., Yang, Z. y Tang, J. (2024). GPT understands, too. *AI Open*, 5, 208-215. <https://doi.org/10.1016/j.aiopen.2023.08.012>
- Marsh, H. W. (1996). Positive and negative global self-esteem: A substantively meaningful distinction or artifacts? *Journal of Personality and Social Psychology*, 70(4), 810-819. <https://doi.org/10.1037/0022-3514.70.4.810>
- Motlagh, N. Y., Khajavi, M., Sharifi, A. y Ahmadi, M. (2023). *The Impact of Artificial Intelligence on the Evolution of Digital Education: A Comparative Study of* <https://doi.org/10.3390/digital1010002> *OpenAI Text Generation Tools including ChatGPT, Bing Chat, Bard, and Ernie* (Versión 1). arXiv. <https://doi.org/10.48550/ARXIV.2309.02029>
- Nandwani, P. y Verma, R. (2021). A review on sentiment analysis and emotion detection from text. *Social Network Analysis and Mining*, 11(1), 81. <https://doi.org/10.1007/s13278-021-00776-6>
- Nunnally, J. C. (1978). *Psychometric theory* (2d ed). McGraw-Hill.
- Okonkwo, C. W. y Ade-Ibijola, A. (2021). Chatbots applications in education: A systematic review. *Computers and Education: Artificial Intelligence*, 2, 100033. <https://doi.org/10.1016/j.caeai.2021.100033>
- Onal, S. y Kulavuz-Onal, D. (2024). A Cross-Disciplinary Examination of the Instructional Uses of ChatGPT in Higher Education. *Journal of Educational Technology Systems*, 52(3), 301-324. <https://doi.org/10.1177/00472395231196532>
- OpenAI. (s. f.). *OpenAI API Reference*. Recuperado 25 de noviembre de 2024, de <https://platform.openai.com/docs/api-reference/chat/create#chat-create-temperature>
- Pérez, J. Q., Daradoumis, T. y Puig, J. M. M. (2020). Rediscovering the use of chatbots in education: A systematic literature review. *Computer Applications in Engineering Education*, 28(6), 1549-1565. <https://doi.org/10.1002/cae.22326>
- Pérez-Marín, D. (2021). A Review of the Practical Applications of Pedagogic Conversational Agents to Be Used in School and University Classrooms. *Digital*, 1(1), 18-33. <https://doi.org/10.3390/digital1010002>
- Perrotta, C. y Selwyn, N. (2020). Deep learning goes to school: Toward a relational understanding of AI in education.

- Learning, Media and Technology*, 45(3), 251-269. <https://doi.org/10.1080/17439884.2020.1686017>
- Peters, M. A., Jackson, L., Papastephanou, M., Jandrić, P., Lazaroiu, G., Evers, C. W., Cope, B., Kalantzis, M., Araya, D., Tesar, M., Mika, C., Chen, L., Wang, C., Sturm, S., Rider, S. y Fuller, S. (2024). AI and the future of humanity: ChatGPT-4, philosophy and education – Critical responses. *Educational Philosophy and Theory*, 56(9), 828-862. <https://doi.org/10.1080/00131857.2023.2213437>
- Pokhrel, S. y Chhetri, R. (2021). A Literature Review on Impact of COVID-19 Pandemic on Teaching and Learning. *Higher Education for the Future*, 8(1), 133-141. <https://doi.org/10.1177/2347631120983481>
- Polit, D. F., Beck, C. T., & Owen, S. V. (2007). Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Research in nursing & health*, 30(4), 459-467. <https://doi.org/10.1002/nur.20199>
- Polverini, G. y Gregorcic, B. (2024). How understanding large language models can inform the use of ChatGPT in physics education. *European Journal of Physics*, 45(2), 025701. <https://doi.org/10.1088/1361-6404/ad1420>
- Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*, 48(2). <https://doi.org/10.18637/jss.v048.i02>
- Roszkowski, M. J. y Soven, M. (2010). Shifting gears: Consequences of including two negatively worded items in the middle of a positively worded questionnaire. *Assessment & Evaluation in Higher Education*, 35(1), 113-130. <https://doi.org/10.1080/02602930802618344>
- Ruiz-Corbella, M. y García-Aretio, L. (2023). Virtual mobility in higher education, ¿chance or utopy? *Revista Española de Pedagogía*, 68(246). <https://doi.org/10.22550/2174-0909.3568>
- Streiner, D. L. (2003). Starting at the Beginning: An Introduction to Coefficient Alpha and Internal Consistency. *Journal of Personality Assessment*, 80(1), 99-103. https://doi.org/10.1207/S15327752JPA8001_18
- Taani, O. y Alabidi, S. (2024). ChatGPT in education: Benefits and challenges of ChatGPT for mathematics and science teaching practices. *International Journal of Mathematical Education in Science and Technology*, 1-30. <https://doi.org/10.1080/0020739X.2024.2357341>
- Tamkin, A., Brundage, M., Clark, J. y Ganguli, D. (2021). *Understanding the Capabilities, Limitations, and Societal Impact of Large Language Models* (Version 1). arXiv. <https://doi.org/10.48550/ARXIV.2102.02503>
- Thottoli, M. M., Alruqaishi, B. H. y Soosaimanickam, A. (2024). Robo academic advisor: Can chatbots and artificial intelligence replace human interaction? *Contemporary Educational Technology*, 16(1), ep485. <https://doi.org/10.30935/cedtech/13948>
- Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R. y Agvemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, 10(1), 15. <https://doi.org/10.1186/s40561-023-00237-x>
- Venkatesh, V. y Morris, M. G. (2000). Why Don't Men Ever Stop to Ask for Directions? Gender, Social Influence, and Their Role in Technology Acceptance and Usage Behavior. *MIS Quarterly*, 24(1), 115. <https://doi.org/10.2307/3250981>
- Vierhauser, M., Groher, I., Antensteiner, T. y Sauerwein, C. (2024). Towards Integrating Emerging AI Applications in SE Education. *2024 36th International Conference on Software Engineering Education and Training (CSEE&am-*

- p;T), 1-5. <https://doi.org/10.1109/CSEET62301.2024.10663045>
- Wang, T., Lund, B. D., Marengo, A., Pagano, A., Mannuru, N. R., Teel, Z. A. y Pange, J. (2023). Exploring the Potential Impact of Artificial Intelligence (AI) on International Students in Higher Education: Generative AI, Chatbots, Analytics, and International Student Success. *Applied Sciences*, 13(11), 6716. <https://doi.org/10.3390/app13116716>
- Wardat, Y., Tashtoush, M. A., AlAli, R. y Jarrah, A. M. (2023). ChatGPT: A revolutionary tool for teaching and learning mathematics. *Eurasia Journal of Mathematics, Science and Technology Education*, 19(7), em2286. <https://doi.org/10.29333/ejmste/13272>
- Wheaton, B. (1987). Assessment of Fit in Overidentified Models with Latent Variables. *Sociological Methods & Research*, 16(1), 118-154. <https://doi.org/10.1177/0049124187016001005>
- White, J., Fu, Q., Hays, S., Sandborn, M., Olea, C., Gilbert, H., Elnashar, A., Spencer-Smith, J. y Schmidt, D. C. (2023). *A Prompt Pattern Catalog to Enhance Prompt Engineering with ChatGPT* (arXiv:2302.11382). arXiv. <http://arxiv.org/abs/2302.11382>
- Winkler, R. y Soellner, M. (2018). Unleashing the Potential of Chatbots in Education: A State-Of-The-Art Analysis. *Academy of Management Proceedings*, 2018(1), 15903. <https://doi.org/10.5465/AMBPP.2018.15903abstract>
- Wolf, E. J., Harrington, K. M., Clark, S. L. y Miller, M. W. (2013). Sample Size Requirements for Structural Equation Models: An Evaluation of Power, Bias, and Solution Propriety. *Educational and Psychological Measurement*, 73(6), 913-934. <https://doi.org/10.1177/0013164413495237>
- Yung, Y. y Bentler, P. M. (1994). Bootstrap corrected ADF test statistics in covariance structure analysis. *British Journal of Mathematical and Statistical Psychology*, 47(1), 63-84. <https://doi.org/10.1111/j.2044-8317.1994.tb01025.x>
- Yurdugül, H. (2008). Minimum sample size for Cronbach's coefficient alpha: A Monte-Carlo study. *Hacettepe Üniversitesi eğitim fakültesi dergisi*, 35(35), 1-9.
- Zhao, W. X., Zhou, K., Li, J., Tang, T., Wang, X., Hou, Y., Min, Y., Zhang, B., Zhang, J., Dong, Z., Du, Y., Yang, C., Chen, Y., Chen, Z., Jiang, J., Ren, R., Li, Y., Tang, X., Liu, Z., ... Wen, J.-R. (2023). *A Survey of Large Language Models* (Versión 15). arXiv. <https://doi.org/10.48550/ARXIV.2303.18223>

Fecha de recepción del artículo: 1 de diciembre de 2024

Fecha de aceptación del artículo: 12 de marzo de 2025

Fecha de aprobación para maquetación: 23 de abril de 2025

Fecha de publicación en OnlineFirst: 30 de abril de 2025

Fecha de publicación: 1 de julio de 2025

Agente de Inteligencia Artificial Generativa en investigación científica. Un análisis explicativo del aprendizaje en el aula

Generative Artificial Intelligence agent in scientific research. An explanatory analysis of classroom learning

-----  -----

 Roberto Berrios Zepeda - *Universidad Nacional Autónoma de Nicaragua, UNAN (Nicaragua)*
 Lorgia Márquez Mora - *Universidad Nacional Autónoma de Nicaragua, UNAN (Nicaragua)*

RESUMEN

Existe un desconocimiento de la capacidad, utilidad y efectividad de algunos recursos tecnológicos como los agentes inteligentes con inteligencia artificial en contextos formativos en investigación científica. Esto motiva al desarrollo y análisis de una nueva estrategia pedagógica que utilice agentes inteligentes generativos con inteligencia artificial en la construcción de proyectos de investigación. Por tanto, se pretende verificar la efectividad de un nuevo procedimiento pedagógico y el diseño de actividades que utilicen agentes inteligentes generativos con inteligencia artificial para la mejora del aprendizaje en investigación científica. El método utilizado fue explicativo con diseño cuasi experimental de corte longitudinal y prospectivo. Se establecieron cuatro pasos del proyecto y sus respectivas hipótesis, fueron construidos y validados los instrumentos, se aplicaron a una muestra de 111 elementos de estudio organizados en un grupo de comparación y dos grupos de intervención, se aplicó un análisis de ANOVA de medidas repetidas. Se demostraron las diferencias significativas del avance en los grupos de intervención y el grupo de comparación en el aprendizaje, Idea de investigación, identificando el vacío y propósito de investigación; Planteamiento del estudio, identificando referencias bibliográficas y contexto del estudio; Diseño de investigación, determinando el método y procedimiento metodológico y Análisis de datos, interpretando datos de nivel descriptivo. La nueva metodología utilizada y asistida por inteligencia artificial obtuvo resultados generales satisfactorios.

Palabras clave: agentes de Inteligencia Artificial Generativa; aprendizaje; investigación científica.

ABSTRACT

There is a lack of knowledge regarding the capacity, usefulness and effectiveness of some technological resources, such as intelligent agents with artificial intelligence in educational contexts for scientific research. This motivates the development and analysis of a new pedagogical strategy that uses generative intelligent agents with artificial intelligence in the construction of research projects. Therefore, the objective is to verify the effectiveness of a new pedagogical procedure and the design of activities that employ generative intelligent agents with artificial intelligence to enhance learning in scientific research. The method used was explanatory with a quasi-experimental longitudinal and prospective design. Four project steps and their respective hypotheses were established, instruments were developed and validated and applied to a sample of 111 study elements organized into one comparison group and two intervention groups. A repeated measures ANOVA analysis was conducted. Significant differences were demonstrated in the progress of the intervention groups compared to the comparison group in learning, research idea development by identifying research gaps and objectives; study formulation by identifying bibliographic references and study context; research design by determining the method and methodological procedure; and data analysis by interpreting descriptive-level data. The new methodology used and assisted by artificial intelligence yielded satisfactory overall results.

Keywords: Generative Artificial Intelligence agents; learning; scientific research.

INTRODUCCIÓN

Según UNESCO IESALC (2020) la crisis de salud pública mundial ocasionada por el virus SARS-Cov-2 en el año 2019 acentuó una serie de retos al sistema de educación superior, la elaboración de medidas pedagógicas para evaluar formativamente la asimilación de los alumnos e incrementar el uso y diversidad de recursos digitales y el acceso a la información en cualquier lugar o momento. Otros autores como Kotler et al. (2021) coinciden en que la crisis sanitaria y las medidas de distanciamiento físico presionó a las instituciones para que se volvieran más tecnológicas, a este escenario se le incluye los avances y robustez de la capacidad de cómputo, plataformas de código abierto, conexiones a la web, capacidad y acceso a la nube, electrónica de móviles y big data que permiten el avance de tecnologías que pretenden imitar a las personas, entre las que sobresalen inteligencia de las máquinas, lenguaje natural de las máquinas, sensores electrónicos, autómatas mecánicos, realidad aumentada y virtual, conexión de las cosas y blockchain (Kotler et al., 2021 y Liu et al., 2022).

Según Salmerón et al. (2023); Yang et al. (2021) y Alhayani et al. (2021), los avances tecnológicos y las experiencias de aplicación de nuevas tecnologías a los procesos de educación y formación profesional son más acciones específicas que procesos administrados y gestionados para el fortalecimiento de la educación.

Una de las áreas de conocimiento de mayor interés a nivel internacional, es la inteligencia de las máquinas (IA), aunque, la comunidad científica no tiene aún un concepto definitivo de inteligencia artificial. Sin embargo, es reconocida como una ciencia interdisciplinaria con múltiples enfoques, entre los que sobresalen el pensamiento y acción humana y racional, además, de su aplicación en los procesos de percepción, razonamiento y aprendizaje en diferentes áreas de conocimientos (García-Peñalvo, 2023 y DataScientest, 2023).

Según Sánchez (2023) existe un desconocimiento de la capacidad y utilidad de algunos recursos tecnológicos como los agentes inteligentes con inteligencia artificial que resultan interesantes y motivadores para la educación y el desempeño laboral. Además, se desconoce la productividad de las herramientas en diferentes áreas de formación profesional como la investigación científica siendo interesante explicar, ¿Cuál es la efectividad de un nuevo procedimiento pedagógico que utiliza agentes inteligentes generativos para mejorar la asimilación de los procesos de investigación científica en alumnos de grado?

Según Sánchez (2023) existe una actitud positiva hacia el uso de ChatGPT en procesos didácticos, fortaleciendo el aprendizaje adaptado, la ayuda en la escritura, la creación de novedosas ideas y las competencias en investigación. Otros autores como González Sánchez et al. (2023) identifican la necesidad de comprender el efecto real de la IA en la generación de conocimiento significativo, este contexto promueve la importancia de analizar el efecto de nuevas estrategias que utilicen agentes inteligentes generativos para la mejora del aprendizaje en investigación científica motivando el desarrollo de este trabajo.

Según Litardo et al. (2023) argumentan que la inteligencia artificial puede mejorar el aprendizaje y la adaptación a las preferencias de los alumnos, lo que puede llevar a un mayor compromiso y rendimiento académico.

Este trabajo busca analizar estrategias que optimizan el uso de recursos tecnológicos para la elaboración de proyectos de investigación. Por tanto, el objetivo de esta investigación es verificar la efectividad de un nuevo procedimiento pedagógico que utiliza agentes inteligentes generativos para mejorar la asimilación de los procesos de investigación.

En las siguientes secciones, se describen temáticas sobre inteligencia de las máquinas IA, su conexión con la educación superior y la investigación científica. Se presenta el método y procedimiento aplicado, se presentan los resultados y discusiones, conclusiones y oportunidades para nuevos trabajos de investigación.

Inteligencia artificial

La IA tiene sus raíces en la década de 1950, con pioneros como Turing (1950) y McCarthy et al. (1955) sentando las bases teóricas. En esta etapa, se exploraron conceptos como el aprendizaje automático y la lógica simbólica. Aunque algunos autores establecen que es en 1943 donde tiene sus inicios la inteligencia artificial con el trabajo de McCulloch y Pitts (1943) quienes presentaron por primera vez un modelo matemático para el diseño de una red neuronal (DataScientest, 2023)

Definir la inteligencia de las máquinas es complicado porque existen diferentes aproximaciones en su desarrollo (Nilsson, 1982; García-Peñalvo, 2023). Para algunos autores puede considerarse una extensión de la informática, cuyo propósito es desarrollar máquinas que puedan hacer acciones que tradicionalmente requerían razonamiento de las personas, códigos activados por restricciones expuestos por modelos que conectan la percepción, el pensamiento y la acción o recursos electrónicos que responden a simulaciones humanas con capacidad de observación, análisis e intención; la ingeniería de la creación de máquinas o programas informáticos inteligentes.

En el proceso de mejora de la inteligencia artificial se identifican una serie de áreas de interés donde la inteligencia de las máquinas puede contribuir de manera destacada tal y como evidencian algunos trabajos en áreas como la investigación científica (Díaz, 2024), la investigación comercial para optimizar procesos de negocio y mejorar la toma de decisiones (Yu y Sup, 2021) y organizaciones que promueven investigaciones en diferentes sectores de la sociedad (UNESCO, 2021).

Según García-Peñalvo et al. (2024) hay un crecimiento exponencial de herramientas informáticas con características inteligentes gracias a la popularidad de los modelos de aprendizaje profundos grandes o LLM (Gruetzemacher y Paradice, 2022) y en especial a uno de los modelos transformadores preentrenados generativos o GPT (Brown et al., 2020). Esta diversidad de trabajos en áreas estratégicas de la sociedad permite reconocer ventajas funcionales y utilitarias importantes para

el desarrollo de procesos aplicados al desarrollo integral y sostenible de diversos campos sociales.

Inteligencia artificial y educación

La inteligencia de las máquinas en la educación universitaria es un área multifacética que está experimentando un desarrollo significativo.

Villarroel (2021) destaca los enfoques de IA incorporados para hacer más eficiente la enseñanza y el aprendizaje remoto. En este contexto se establece el reto por parte de la UNESCO para promover las tecnologías de inteligencia artificial (IA) regidas por los principios de equidad e inclusión y orientados a la consecución de los objetivos de desarrollo sostenible (ODS) a través de la agenda de educación 2030. Por tanto, algunos trabajos de investigación como Jia et al. (2022) establecen la importancia del análisis de datos educativos, a través de la exploración y el descubrimiento de conocimientos en bases de datos educativas para comprender los estándares de comportamiento de los alumnos y mejorar la gestión del sistema educativo. Por su parte García-Peñalvo (2020) y Lang et al. (2022) establecen la importancia de la analítica de aprendizaje para determinar los estilos de aprendizaje y facilitar la colaboración entre los estudiantes, lo que contribuye a un proceso educativo más dinámico y efectivo. En este contexto, crece el interés de como la IA contribuye al aprendizaje a través de sistemas inteligentes y automatización de contenidos (Ma et al., 2014; Yilmaz et al., 2022) fomentando así un aprendizaje más activo y autónomo.

Sari y Purwanta (2021) establecen que la IA puede potenciar el aprendizaje creativo en el aula. Otros autores como García Rosado (2024) proponen que el uso de estas herramientas inteligentes ayuda a la construcción de vínculos de confianza con el alumnado, el proceso pedagógico centrado en las personas, donde la evaluación no sea un mecanismo de control, sino un proceso de aprendizaje en sí mismo (Rudolph et al., 2023). Por tanto, surge el interés de utilizar herramientas de inteligencia artificial para mejorar la productividad de los procesos de enseñanza y asimilación, permitiendo la realimentación y asesoramiento a los estudiantes (Baker, 2016; Zawacki-Richter et al., 2019 y Villarroel, 2021).

Inteligencia artificial e investigación

Según López Martín (2023), el uso de inteligencia de las máquinas puede aportar valor en la producción y edición de manuscritos y su difusión después de su publicación. Por su parte, los trabajos de Lalaleo et al. (2024) establecen que la IA debe ser una herramienta que optimice la escritura esencial en la generación de conocimiento científico, en coordinación con la experiencia del docente. Otros trabajos como los de García Rosado (2024) identifican retos para caracterizar y sistematizar experiencias en el desarrollo de recursos didácticos y contenido teórico práctico de la IA en metodología de la investigación. El trabajo de Vera (2023)

establece que la inteligencia de las máquinas permite procesar grandes cantidades de datos e identificar patrones y tendencias, facilitando la generación de conocimiento y la toma de decisiones sustentadas.

Existen pocos trabajos que contribuyen a conocer como el uso de herramientas inteligentes ayudan a la mejora de los proyectos de investigación, parte de la complejidad de estos procesos recae en comprender que los proyectos de investigación son procesos que se construyen según los propósitos, variables y poblaciones de estudio definidas por el investigador. Se requiere un esfuerzo para identificar limitaciones o vacíos de información de una línea de investigación y que ayude a plantear correctamente el título del estudio (Ayala, 2020). Por su parte, Carvajal (2023) logra aplicar un procedimiento para sistematizar, delimitar y ajustar un tema de investigación haciendo uso de agentes de inteligencia generativo. Por tanto, se plantea lo siguiente.

- H1: Existen diferencias relevantes de las respuestas del pretest y posttest, aprendizaje idea de investigación reportados por los participantes, dependiendo si han recibido o no el nuevo procedimiento metodológico con IA.

Existe la dificultad de comprender la complejidad del contexto de un estudio y plantear un problema a resolver correctamente, esto se vuelve más complejo cuando no se tiene información y las herramientas necesarias para desarrollar esta etapa del proceso de investigación. Algunos autores como Ayala (2020) establecen como centro de una investigación el correcto planteamiento del problema. Por su parte, el trabajo de Carvajal (2023) establece un procedimiento para sistematizar e identificar información para construir una parte del planteamiento del problema centrado en los objetivos, preguntas y posibles hipótesis de investigación. Por tanto, se plantea lo siguiente.

- H2: Existen diferencias relevantes de las respuestas del pretest y posttest, aprendizaje planteamiento del estudio reportados por los participantes, dependiendo si han recibido o no el nuevo procedimiento metodológico con IA.

Existen limitaciones en la comprensión de los protocolos del diseño de investigación correctos y acertados, asociados a la taxonomía de los conceptos y las habilidades empíricas. El trabajo de Carvajal (2023) sistematiza la teoría y los procedimientos para la construcción de un diseño de investigación asistido por agentes inteligentes generativos. Por tanto, se plantea lo siguiente.

- H3: Existen diferencias relevantes de las respuestas del pretest y posttest, aprendizaje diseño de investigación reportados por los participantes, dependiendo si han recibido o no el nuevo procedimiento metodológico con IA.

Existen limitaciones en la comprensión del uso de la técnica correcta de análisis estadístico que fortalezca la confianza y seguridad del conocimiento generado para ser aplicado o replicado. El trabajo de Carvajal (2023) logra extraer, sintetizar y resumir información de análisis exploratoria con el uso de GPT. Por tanto, se plantea lo siguiente.

H4: Existen diferencias relevantes de las respuestas del pretest y postest, aprendizaje análisis de datos reportados por los participantes, dependiendo si han recibido o no el nuevo procedimiento metodológico con IA.

MATERIAL Y MÉTODO

Método

El estudio es de diseño explicativo de tipo cuasi experimental con intervención longitudinal y prospectivo.

Participantes

Las unidades de análisis son todos los estudiantes del curso de grado en investigación de mercados formada por 111 estudiantes. Se organizaron dos grupos experimentales, grupo experimental 1 con 32 estudiantes, grupo experimental 2 con 31 alumnos y un grupo control conformado por 48 alumnos. Los grupos se asignaron con base en el registro de matrícula ya existente, lo que limita la asignación aleatoria y aumenta el riesgo de sesgos por factores externos, aunque existe homogeneidad de los grupos y un nivel de similitud de competencias académicas.

Tabla 1
Distribución de los grupos por género y edad

	EDAD			GEN		
	G. CON-TROL	G. EXP 1	G. EXP 2	G. CON-TROL	G. EXP 1	G. EXP 2
Válido	48	32	31	48	32	31
Ausente	0	0	0	0	0	0
Media	19.875	20.563	19.839	1.521	1.469	1.742
Desviación Típica	1.196	1.883	1.344	0.505	0.507	0.445

Fuente: datos del estudio.

Instrumentos

Evaluamos cuatro pasos del proceso de investigación científica. Ver Figura 1.

El primer paso, idea de investigación, se desarrolla a través de un cuestionario de 17 ítems, con una fiabilidad de 0.765 McDonald ω , aceptable. El segundo paso, planteamiento del estudio con un cuestionario de 12 ítems y una fiabilidad de 0.81 McDonald ω , bueno. El tercer paso diseño de investigación con un cuestionario de 14 ítems y una fiabilidad de 0.80 McDonald ω , bueno. El cuarto paso análisis de datos con un cuestionario de 12 ítems y una fiabilidad moderada de 0.72 McDonald ω , adaptada de la planificación curricular por competencias en investigación de mercados de Sandino et al. (2019).

Se utiliza una escala tipo Likert de 1 a 5 puntos donde 1 es Definitivamente no y 5 Definitivamente sí.

Procedimiento

Se proponen cuatros pasos y acciones del proceso de investigación científica, para su desarrollo y análisis. Ver Figura 1.

Figura 1

Etapas de la investigación científica y procesos de sistematización para los agentes inteligentes generativos



Fuente: elaboración propia a partir de UNESCO IESALC (2023) y Salmerón et al. (2023).

Se inicia cada paso con un (pretest) aplicado en el aula de clases, la metodología docente empleada incluyó clases magistrales, luego se orienta una guía metodológica para el desarrollo del nuevo procedimiento en los grupos experimentales, en el grupo de control se desarrolló el procedimiento tradicional que consistió en conferencia y orientación de trabajo independiente en grupos fuera del aula. Los procedimientos

para cada paso tienen una duración de 15 días, al culminar cada paso se autoevalúa el estudiante con un cuestionario (postest) aplicado en el aula de clases.

El nuevo procedimiento consiste en el desarrollo de una guía compuesta por un input o prompt de perspectiva profesional según Morales-Chan (2023) y el uso de un patrón de lenguaje que utilizó tópico, forma, acentuación y detalles del contexto según Dathathri et al. (2019) para cada paso y un output o resultado de la búsqueda para la construcción del proyecto de investigación. El recurso tecnológico utilizado es Perplexity IA ® un motor de búsqueda de fuentes y citas con enlaces web. El modelo de acceso abierto de Perplexity está basado en el modelo GPT-3.5 de Open AI ®, combinado con el modelo de aprendizaje profundo grande (LLM) independiente de la compañía. Perplexity Pro tiene acceso premium a GPT-4 ® y Claude 3 ®.

Análisis de los datos

La comparación de las respuestas entre pretest y postest en cada paso del proyecto de investigación de los grupos se realizó a través de dos pruebas de ANOVA para muestras repetidas entre los grupos y probar las hipótesis definidas. Se utilizó el programa multiplataforma Jeffreys' s Amazing Statistics Program (JASP 0.18.1.0) ® para el análisis de los datos.

RESULTADOS

A continuación, se muestran los resultados por cada paso metodológico y pruebas de hipótesis.

Paso 1. Idea de investigación

Antes de iniciar el análisis se realizó la verificación del supuesto a través del contraste de varianza de Levene, para los resultados del pretest de $p = 0.11$ y postest $p = 0.20$ ambos mayor a $\alpha = 0.05$, cumpliendo el supuesto de igualdad de varianza.

Hay diferencias a nivel general en los niveles del aprendizaje, Idea de investigación de los grupos, las diferencias son significativas $p < 0.001$ menor a $\alpha = 0.05$. El resultado muestra que, sí existen diferencias significativas entre la puntuación pretest y postest, sin separar a los sujetos por grupos control y experimentales. Además, se indica la interacción entre la variable aprendizaje idea de investigación y los grupos, si las diferencias pretest y postest son diferente en función al grupo, vemos el p valor < 0.001 que es menor a $\alpha = 0.05$, por tanto, sí existen diferencias significativas. El criterio que más contribuye al factor Idea de investigación es, buscar el vacío en la línea de investigación. El 21 % de la variabilidad en el nivel, aprendizaje Idea de investigación es explicada en el momento de la medición ($\eta^2 = 0.21$). Ver Tabla 2 y 3.

Tabla 2*Efectos dentro de los sujetos*

Casos	Suma de Cuadrados	gl	Cuadrado Medio	F	p	η^2
Idea de investigación	1931.993	1	1931.993	124.758	< .001	0.182
Idea de investigación * GRUPOS	1854.668	2	927.334	59.882	< .001	0.174
Residuals	1842.827	119	15.486			

Nota: Suma de Cuadrados Tipo III**Tabla 3***Efectos entre sujetos*

Casos	Suma de Cuadrados	gl	Cuadrado Medio	F	p	η^2
GRUPOS	2270.885	2	1135.442	49.365	< .001	0.213
Residuals	2737.136	119	23.001			

Nota: Suma de Cuadrados Tipo III

Paso 2. Planteamiento del estudio

Antes de iniciar el análisis se realizó la verificación del supuesto a través del contraste de varianza de Levene, para los resultados del pretest de $p = 0.60$ y posttest $p = 0.96$ ambos mayor a $\alpha = 0.05$, cumpliendo el supuesto de igualdad de varianza.

Hay diferencias a nivel general en los niveles del procedimiento Planteamiento del estudio en los grupos, las diferencias son significativas $p < 0.001$ menor a $\alpha = 0.05$. El resultado muestra que, sí existen diferencias significativas entre la puntuación pretest y posttest, sin separar a los sujetos por grupos control y experimentales, es decir que los sujetos a nivel general independiente del grupo tienen valores más altos en el posttest que el pretest. Además, se indica la interacción entre la variable aprendizaje Planteamiento del estudio y los grupos, vemos que el p valor < 0.001 es menor a $\alpha = 0.05$, por tanto, sí existen diferencias significativas. El criterio que más contribuye al factor Planteamiento del Estudio es, versatilidad en buscar información, determinar el contexto y plantear la problemática del estudio. El 24.8 % de la variabilidad en el nivel, aprendizaje del planteamiento del estudio es explicada en el momento de la medición ($\eta^2 = 0.248$). Ver Tablas 4 y 5.

Tabla 4
Efectos dentro de los sujetos

Casos	Suma de Cuadrados	gl	Cuadrado Medio	F	p	η^2
Planteamiento del estudio	2466.212	1	2466.212	153.207	< .001	0.217
Planteamiento del estudio * GRUPO	1320.991	2	660.495	41.032	< .001	0.116
Residuals	1818.988	113	16.097			

Nota: Suma de Cuadrados Tipo III

Tabla 5
Efectos entre sujetos

Casos	Suma de Cuadrados	gl	Cuadrado Medio	F	p	η^2
GRUPO	2821.059	2	1410.529	54.246	< .001	0.248
Residuals	2938.299	113	26.003			

Nota: Suma de Cuadrados Tipo III

Paso 3. Diseño de investigación

Se aplica el contraste de igualdad de varianzas de Levene, el p valor 0.232 para el pretest y 0.089 del posttest está por encima de α 0.05, es decir se cumple el supuesto de igualdad de varianzas entre los grupos.

Hay diferencias a nivel general en los niveles del aprendizaje, Diseño de investigación en los grupos, las diferencias son significativas $p < 0.001$ menor a α 0.05. El resultado muestra que, sí existen diferencias significativas entre la puntuación pretest y posttest, sin separar a los sujetos por grupos control y experimentales. El p valor es < 0.001 es menor a α 0.05, es significativa, es decir que los sujetos a nivel general independiente del grupo tienen valores más altos en el posttest que el pretest. Además, se indica la interacción entre la variable aprendizaje diseño de investigación y los grupos, si las diferencias pretest y posttest son diferentes en función al grupo, vemos el p valor < 0.001 que es menor a α 0.05, por tanto, sí existen diferencias de valores entre las respuestas. El criterio que más contribuye al factor Diseño de investigación es, el procedimiento metodológico es dinámico e interactivo en la búsqueda de información científica para describir el método y procedimiento de

estudio. El 1.3 % de la variabilidad en el nivel aprendizaje, Diseño de investigación es explicada en el momento de la medición ($\eta^2 = 0.013$). Ver Tablas 6 y 7.

Tabla 6
Efectos dentro de los sujetos

Casos	Suma de Cuadrados	gl	Cuadrado Medio	F	p	η^2
Diseño de investigación PRE POS	1657.366	1	1657.366	191.990	< .001	0.397
Diseño de investigación PRE POS * GRUPO	779.316	2	389.658	45.138	< .001	0.186
Residuals	923.684	107	8.633			

Nota: Suma de Cuadrados Tipo III

Tabla 7
Efectos entre sujetos

Casos	Suma de Cuadrados	gl	Cuadrado Medio	F	p	η^2
GRUPO	52.483	2	26.241	3.661	0.029	0.013
Residuals	767.044	107	7.169			

Nota: Suma de Cuadrados Tipo III

Paso 4. Análisis de datos

Se aplica el contraste de igualdad de varianza de Levene, el p valor 0.073 para el pretest y 0.423 del posttest está por encima de α 0.05, es decir se cumple el supuesto de igualdad de varianza entre los grupos.

Hay diferencias a nivel general en los niveles del procedimiento, Análisis de datos en los grupos, las diferencias son significativas $p < 0.001$ menor a α 0.05. El resultado muestra que, sí existen diferencias significativas entre la puntuación pretest y posttest, sin separar a los sujetos por grupos control y experimentales. El p valor es < 0.001 es menor a α 0.05, es significativa, es decir que los sujetos a nivel general independiente del grupo tienen valores más altos en el posttest que el pretest. Además, se indica la interacción entre la variable Aprendizaje análisis de datos y los grupos, si las diferencias pretest y posttest son diferentes en función al grupo,

vemos el p valor < 0.001 que es menor a α 0.05, por tanto, sí existen diferencias significativas.

El criterio que más contribuye al factor Análisis de datos es, el procedimiento es enriquecedor en la búsqueda de información científica para conocer el significado e interpretar el estadístico. El 34 % de la variabilidad en el nivel, aprendizaje análisis de datos es explicada en el momento de la medición ($\eta^2 = 0.34$). Ver Tablas 8 y 9.

Tabla 8

Efectos dentro de los sujetos

Casos	Suma de Cuadrados	gl	Cuadrado Medio	F	p	η^2
Análisis de respuestas Pre Posttest	111.110	1	111.110	31.318	< .001	0.040
Análisis de respuestas Pre Posttest * Grupo	246.912	2	123.456	34.798	< .001	0.089
Residuals	379.616	107	3.548			

Nota: Suma de Cuadrados Tipo III

Tabla 9

Efectos entre sujetos

Casos	Suma de Cuadrados	gl	Cuadrado Medio	F	p	η^2
Grupo	948.082	2	474.041	46.127	< .001	0.340
Residuals	1099.627	107	10.277			

Nota: Suma de Cuadrados Tipo III

DISCUSIÓN Y CONCLUSIONES

En el proceso, Idea de investigación el efecto entre los sujetos muestra que, sí existen diferencias relevantes entre los resultados de primer cuestionario aplicado y el posterior cuestionario aplicado para los grupos, confirmando la hipótesis aprendizaje Idea de investigación. Además, la interacción entre la variable Idea de investigación y los grupos muestran que existen diferencias significativas, es decir que, el avance que han tenido los grupos experimentales entre el pretest y posttest es significativamente superior que el avance del grupo control y se confirma que en general el nuevo tratamiento con IA de los grupos experimentales ha sido más

efectivo para el aprendizaje Idea de investigación. Se observa un avance significativo del nuevo procedimiento con IA en la búsqueda de información, esclarecimiento del enfoque y propósito del estudio que ayuda a definir mejor el título del estudio en consideración con los avances del grupo de comparación que utilizó el procedimiento tradicional.

El nuevo procedimiento aporta a la mejora de algunas áreas identificadas en los trabajos de Aldana et al. (2020) y Bozkurt et al. (2023) al incrementar la productividad en la construcción del proyecto, la realimentación diversificada con el docente y la calidad de la estructura y contenido del proyecto desde el inicio contribuyendo a la asimilación del conocimiento en investigación.

En el proceso, Aprendizaje planteamiento del estudio el efecto entre los sujetos muestra que, sí existen diferencias relevantes entre los resultados del primer cuestionario aplicado y el posterior cuestionario aplicado para los grupos, afirmando la hipótesis aprendizaje planteamiento del estudio. Además, la interacción entre la variable planteamiento del estudio y los grupos muestran que existen diferencias significativas, es decir que, el avance que han tenido los grupos experimentales entre el pretest y posttest es significativamente superior que el avance del grupo control y se confirma que en general el nuevo tratamiento con IA de los grupos experimentales ha sido más efectivo para el aprendizaje. Se observan avances significativos en el planteamiento del estudio asistido por el nuevo procedimiento que utiliza IA, esto ayudó a diversificar la revisión teórica y optimizar la escritura a como lo menciona Lalaleo et al. (2024), describir el contexto, plantear el problema, definir el propósito y los objetivos del estudio. Aunque, se identifica la necesidad de fortalecer el análisis crítico, la redacción científica y la reducción del sesgo de información en la construcción del contexto del estudio, aportando a reducir la omisión por desconocimiento metodológico.

Se obtiene un leve avance del grupo de control que utilizó el procedimiento tradicional relacionado a la realimentación con el docente tutor.

En el proceso, aprendizaje Diseño de investigación el efecto entre los sujetos muestra que, sí existen diferencias relevantes entre los resultados del primer cuestionario aplicado y el posterior cuestionario aplicado para los grupos, afirmando la hipótesis aprendizaje diseño de investigación. Además, la interacción entre la variable Diseño de investigación y los grupos muestran que existen diferencias significativas, es decir que, el avance que han tenido los grupos experimentales entre el pretest y posttest es superior que el avance del grupo control y se confirma que en general el nuevo tratamiento con IA de los grupos experimentales ha sido más efectivo para el aprendizaje diseño de investigación. Se observan débiles avances en el diseño de investigación asistido por el nuevo procedimiento que utiliza IA principalmente en la búsqueda de información estructurada y coherencia entre el método y el procedimiento metodológico, estableciendo la necesidad de profundizar el análisis para mejorar la calidad del informe del proyecto según la naturaleza,

propósito y nivel del estudio, características propias de la metodología programática de investigación científica (Supo y Zacarías, 2020).

En el proceso, aprendizaje Análisis de datos de la investigación el efecto entre los sujetos muestra que, si existen diferencias relevantes entre los resultados del primer cuestionario aplicado y el posterior cuestionario aplicado para los grupos, afirmando la hipótesis aprendizaje análisis de datos. Además, la interacción entre la variable análisis de datos de investigación y los grupos muestran que existen diferencias significativas, es decir que, el avance que han tenido los grupos experimentales entre el pretest y posttest es significativamente superior que el avance del grupo control y se confirma que en general el nuevo tratamiento con IA de los grupos experimentales ha sido más efectivo para el aprendizaje análisis de datos. Se observan avances significativos en el análisis de datos asistidos por el nuevo procedimiento que utiliza IA especialmente en el análisis descriptivo uni-variable y la lectura de los datos en consideración con los avances del grupo de comparación que utilizó el procedimiento tradicional. Existe un aporte del nuevo procedimiento a la lectura de datos a través de gráficos y la propuesta de ideas coherentes y acertadas a la naturaleza, propósito y nivel del estudio mejorando el aprendizaje creativo (Sari y Purwanta, 2021).

El nuevo procedimiento utilizado asistido por IA obtuvo resultados satisfactorios para el contraste de las hipótesis y su aporte original a la investigación científica a través de un moderno diseño de actividades y metodología pedagógica en el aula de clase.

Existen limitaciones por la conformación de los grupos sin aleatoriedad, el uso de datos autoinformados, la memoria selectiva de los participantes, tendencia a responder de manera positiva, el limitado acceso a conexión de internet y los sesgos de las respuestas de los algoritmos. Se recomienda para futuros proyectos de investigación incluir la discusión de los informes finales de los proyectos, la realimentación con el docente tutor y coevaluación adaptada.

Sería interesante utilizar los avances de agentes automáticos con IA para evaluar la comprensión crítica en un contexto de aprendizaje adaptado en investigación y definir patrones de rutas de asimilación del conocimiento.

REFERENCIAS

- Aldana, G. M., Babativa, D. A., Caraballo, G. J. y Rey, C. A. (2020). Escala de actitudes hacia la investigación (EACIN): Evaluación de sus propiedades psicométricas en una muestra colombiana. *Revista CES Psicología*, 13(1), 89-103. <https://doi.org/10.21615/cesp.13.1.6>
- Alhayani, B., Mohammed, H. J., Chaloob, I. Z. y Ahmed, J. S. (2021). Effectiveness of artificial intelligence techniques against cyber security risks apply of IT industry. *Materials Today: Proceedings*, 44, 3356-3361. <https://doi.org/10.1016/j.matpr.2021.02.531>
- Ayala, O. (2020). Competencias informacionales y competencias investigativas en estudiantes universitarios. *Revista Innova Educación*, 2(4), 668-679. <https://doi.org/10.35622/j.rie.2020.04.011>

- Baker, R. S. (2016). Stupid Tutoring Systems, Intelligent Humans. *International Journal of Artificial Intelligence in Education*, 26(2), 600-614. <https://doi.org/10.1007/s40593-016-0105-0>
- Bozkurt, A., Junhong, X., Lambert, S., Pazurek, A., Crompton, H., Koseoglu, S., Farrow, R., Bond, M., Nerantzi, C., Honeychurch, S., Bali, M., Dron, J., Mir, K., Gray, B. C., Stewart, B. P., Costello, E., Mason, J., Stracke, C. M., Romero-Hall, E., ... Brooks, C. (2023). Speculative futures on ChatGPT and generative artificial intelligence (AI): A collective reflection from the educational landscape. *Asian Journal of Distance Education*, 18(1), 53-130. <https://doi.org/10.5281/zenodo.7636568>
- Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., ... Amodei, D. (2020). *Language Models are Few-Shot Learners*. arXiv. <https://doi.org/10.48550/arXiv.2005.14165>
- Carvajal, I. C. M. (2023). Inteligencia artificial (IA) en la investigación científica: Sistematización y reflexiones sobre experiencias educativas. *Revista EDUCARE - UPEL-IPB - Segunda Nueva Etapa 2.0*, 27(3), 112-137. <https://doi.org/10.46498/reduipb.v27i3.2050>
- DataScientest. (2023). *Inteligencia artificial: definición, historia, usos y peligros*. DataScientest. <https://datascientest.com/es/inteligencia-artificial-definicion>
- Dathathri, S., Madotto, A., Lan, Z., Fung, P. y Neubig, G. (2019). Plug and play language models: A simple approach to controlled text generation. *arXiv*. <https://arxiv.org/abs/1912.02164>
- Díaz, L. A. (2024). El uso de la inteligencia artificial en la investigación científica. *Revista Historia de la Educación Latinoamericana*, 26(43), 1-15. <https://doi.org/10.19053/upte.01227238.18014>
- García-Peñalvo, F. J. (2020). Learning Analytics as a Breakthrough in Educational Improvement. En D. Burgos (Ed.), *Radical Solutions and Learning Analytics: Personalised Learning and Teaching Through Big Data* (pp. 1-15). Springer Singapore. https://doi.org/10.1007/978-981-15-4526-9_1
- García-Peñalvo, F. J. (2023). *Redefiniendo la relación del profesorado con la inteligencia artificial*. II Congreso Internacional de Educación. <https://bit.ly/46Y8Y77>
- García-Peñalvo, F. J., Llorens-Largo, F. y Vidal, J. (2024). La nueva realidad de la educación ante los avances de la inteligencia artificial generativa. *RIED. Revista Iberoamericana de Educación a Distancia*, 27(1), 9-39. <https://doi.org/10.5944/ried.27.1.37716>
- García Rosado, L. F. (2024). Inteligencia artificial, una mirada desde la asignatura de Metodología de la investigación científica: Relato de experiencia docente. *Educación Superior*, (37), 11-34. <https://doi.org/10.56918/es.2024.i37.pp11-34>
- González Sánchez, J. L., Villota García, F. R., Moscoso Parra, A. E., Garces Calva, S. W. y Bazurto Arévalo, B. M. (2023). Aplicación de la Inteligencia Artificial en la Educación Superior. *Dominio de las Ciencias*, 9(3), 1097-1108. <https://doi.org/10.23857/dc.v9i3.3488>
- Gruetzemacher, R. y Paradise, D. (2022). Deep Transfer Learning & Beyond: Transformer Language Models in Information Systems Research. *ACM Computing Surveys*, 54(10S), 1-30. <https://doi.org/10.1145/3505245>
- Jia, K., Wang, P., Li, Y., Chen, Z., Jiang, X., Lin, C. L. y Chin, T. (2022). Research landscape of artificial intelligence and e-learning: A bibliometric research. *Frontiers in Psychology*, 13, 795039. <https://doi.org/10.3389/fpsyg.2022.795039>

- Kotler, P., Kartajaya, H. y Setiawan, I. (2021). *Marketing 5.0: Technology for Humanity*. John Wiley & Sons.
- Lalaleo, F. R., Carrera, F. A. y Martínez, A. P. (2024). La IA como herramienta de apoyo en la investigación científica en los docentes investigadores del ISTE. *Espíritu Emprendedor TES*, 8(1), 97-110. <https://doi.org/10.33970/eetes.v8.n1.2024.377>
- Lang, C., Siemens, G., Wise, A. F., Gašević, D. y Merceron, A. (Eds.). (2022). *The Handbook of Learning Analytics*. Society for Learning Analytics Research (SoLAR). <https://doi.org/10.18608/hla22>
- Litardo, J. T., Wong, C. R., Ruiz, S. M. y Benites, K. P. (2023). Retos y oportunidades docente en la implementación de la inteligencia artificial en la educación superior ecuatoriana. *South Florida Journal of Development*, 4(2), 867-889. <https://doi.org/10.46932/sfjdv4n2-020>
- Liu, Y., Chen, L. y Yao, Z. (2022). The application of artificial intelligence assistant to deep learning in teachers' teaching and students' learning processes. *Frontiers in Psychology*, 13, 929175. <https://doi.org/10.3389/fpsyg.2022.929175>
- López Martín, E. (2023). El papel de la inteligencia artificial generativa en la publicación científica. *Educación XXI*, 27(1), 9-15. <https://doi.org/10.5944/educxx1.39205>
- Ma, W., Adesope, O. O., Nesbit, J. C. y Liu, Q. (2014). Intelligent tutoring systems and learning outcomes: A meta-analysis. *Journal of Educational Psychology*, 106(4), 901-918. <https://doi.org/10.1037/a0037123>
- McCarthy, J., Minsky, M. L., Rochester, N. y Shannon, C. E. (1955). A proposal for the Dartmouth summer research project on artificial intelligence. *AI Magazine*, 27(4), 12-14. <https://doi.org/10.1609/aimag.v27i4.1904>
- McCulloch, W. S. y Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, 5(4), 115-133. <https://doi.org/10.1007/bf02478259>
- Morales-Chan, M. (2023). Explorando el potencial de Chat GPT: una clasificación de Prompts efectivos para la enseñanza. *Paper GES*. <https://biblioteca.galileo.edu/tesario/handle/123456789/1348>
- Nilsson, N. J. (1982). Artificial intelligence: Engineering, science, or slogan? *AI Magazine*, 3(1), 2-17. <https://doi.org/10.1609/aimag.v3i1.359>
- Rudolph, J., Tan, S. y Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning and Teaching*, 6(1), 1-22. <https://doi.org/10.37074/jalt.2023.6.1.9>
- Salmerón Moreira, Y. M., Luna Álvarez, H. E., Murillo Encarnación, W. G. y Pacheco Gómez, V. A. (2023). El futuro de la Inteligencia Artificial para la educación en las instituciones de Educación Superior. *Revista Conrado*, 19(93), 27-34.
- Sánchez, O. V. G. (2023). Uso y percepción de ChatGPT en la educación superior. *Revista de Investigación en Tecnologías de la Información*, 11(23), 98-107. <https://doi.org/10.36825/RITI.11.23.009>
- Sandino, M., Espinoza, M. J. y Berrios, R. (2019). *Microprogramación del componente de investigación de mercados. Proyecto curricular para docencia*. Universidad Nacional Autónoma de Nicaragua UNAN León.
- Sari, J. M. y Purwanta, E. (2021). The Implementation of Artificial Intelligence in Creative Learning Based on Stem in the Era of Society 5.0. *Tadris: Jurnal Keguruan dan Ilmu Tarbiyah*, 6(2), 433-440. <https://doi.org/10.24042/tadris.v6i2.10135>
- Supo, J. y Zacarías, H. (2020). *Metodología de la investigación científica. Seminarios de investigación científica* (3ra ed.). Sociedad Hispana de Investigación

- Científica, Editorial BIOESTADISTICO EEDU EIRL.
- Turing, A. M. (1950). Computing Machinery and Intelligence. *Mind*, 59(236), 433-460. <https://doi.org/10.1093/mind/lix.236.433>
- UNESCO IESALC. (2020). *COVID-19 y educación superior: De los efectos inmediatos al día después; Análisis de impactos, respuestas políticas y recomendaciones*. <https://www.iesalc.unesco.org/wp-content/uploads/2020/05/COVID-19-ES-130520.pdf>
- UNESCO. (2021). *El aporte de la inteligencia artificial y las TIC avanzadas a las sociedades del conocimiento: Una perspectiva de derechos, apertura, acceso y múltiples actores*. <https://unesdoc.unesco.org/ark:/48223/pf0000375796>
- UNESCO IESALC. (2023). *ChatGPT e Inteligencia artificial en la educación superior. Guía de inicio rápida*. <https://www.iesalc.unesco.org/wp-content/uploads/2023/05/ChatGPT-e-Inteligencia-artificial-en-la-educacion-superior.-Gui%CC%81a-de-inicio-ra%CC%81pida.pdf>
- Vera, F. (2023). Integración de la Inteligencia Artificial en la Educación superior: Desafíos y oportunidades. *Transformar*, 4(1), 17-34. <https://doi.org/10.56219/dialectica.vii21.2322>
- Villarroel, J. J. G. (2021). Implicancia de la inteligencia artificial en las aulas virtuales para la educación superior. *Orbis Tertius-UPAL*, 5(10), 31-52. <https://doi.org/10.59748/ot.v5i10.98>
- Yang, S. J., Ogata, H., Matsui, T. y Chen, N. S. (2021). Human-centered artificial intelligence in education: Seeing the invisible through the visible. *Computers and Education: Artificial Intelligence*, 2, 100008. <https://doi.org/10.1016/j.caeai.2021.100008>
- Yilmaz, R., Yurdugül, H., Karaoğlu Yilmaz, F. G., Şahin, M., Sulak, S., Aydın, F., Tepgeç, M., Terzi Müftüoğlu, C. y Oral, Ö. (2022). Smart MOOC integrated with intelligent tutoring: A system architecture and framework model proposal. *Computers and Education: Artificial Intelligence*, 3, 100092. <https://doi.org/10.1016/j.caeai.2022.100092>
- Yu, W. y Sup, W. (2021). Artificial intelligence for the development of university education management. *Frontiers in Educational Research*, 4(1), 120-125. <https://doi.org/10.25236/FER.2021.040120>
- Zawacki-Richter, O., Marín, V. I., Bond, M. y Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education—where are the educators? *International Journal of Educational Technology in Higher Education*, 16(1), 1-27. <https://doi.org/10.1186/s41239-019-0171-0>

Fecha de recepción del artículo: 1 de diciembre de 2024

Fecha de aceptación del artículo: 28 de febrero de 2025

Fecha de aprobación para maquetación: 23 de marzo de 2025

Fecha de publicación en OnlineFirst: 8 de abril de 2025

Fecha de publicación: 1 de julio de 2025

Diseño y análisis de narrativas enriquecidas con inteligencia artificial. Proyecto de transferencia UNED - UTMACH

Design and analysis of narratives enriched with artificial intelligence. UNED - UTMACH transfer project



- Javier Gil Quintana - *Universidad Nacional de Educación a Distancia, UNED (España)*
- Eduardo García Blázquez - *Universidad Nacional de Educación a Distancia, UNED (España)*
- Carmen Cantillo Valero - *Universidad Nacional de Educación a Distancia, UNED (España)*
- José Javier Hueso Romero - *Universidad Nacional de Educación a Distancia, UNED (España)*

RESUMEN

Este artículo analiza los resultados de un Proyecto de Transferencia desarrollado desde la Universidad Nacional de Educación a Distancia (UNED, España) para la Universidad Técnica de Machala (UTMACH, Ecuador) de integración de la Inteligencia Artificial (IA) en la elaboración de 22 narrativas educativas por alumnado universitario, evaluando su impacto en la claridad, coherencia y personalización del contenido académico. La IA facilita la creación de materiales educativos claros y coherentes que se comprueba al analizar el rendimiento y necesidades del alumnado universitario, permitiendo adaptación del contenido a su nivel de comprensión y corrigiendo errores. La IA también posibilita la personalización de la enseñanza, ajustando las narrativas a los intereses y estilos de aprendizaje individuales, lo que incrementa la relevancia y compromiso de las y los estudiantes. La incorporación de materiales con IA en escenarios virtuales de aprendizaje como los sNOOC hace que el proceso de aprendizaje sea más dinámico y participativo. El método utilizado ha sido mixto basado en la recolección de datos, análisis estadístico e interpretación que han proporcionado los Círculos de Evaluación por Pares formados por 200 investigadores. Los resultados del estudio indican una fuerte correlación entre la relevancia educativa y la claridad del contenido con las narrativas valoradas como relevantes tendiendo a ser percibidas como claras y coherentes. El estudio subraya la importancia de una base ética sólida en el uso de la IA para garantizar una comprensión adecuada y equitativa del contenido, a pesar de que a veces la ética y la claridad no están completamente alineadas.

Palabras clave: inteligencia artificial; narrativas digitales; diseño visual; diseño ético.

ABSTRACT

This article analyzes the results of a Transfer Project developed by the Universidad Nacional de Educación a Distancia (UNED, Spain) for the Universidad Técnica de Machala (UTMACH, Ecuador) on the integration of Artificial Intelligence (AI) in the elaboration of 22 educational narratives by university students, evaluating its impact on the clarity, coherence, and personalization of academic content. AI facilitates the creation of clear and coherent educational materials tested by analyzing university students' performance and needs, allowing adaptation of the content to their level of understanding and correcting errors. AI also enables the personalization of teaching, tailoring narratives to individual interests and learning styles and increasing learners' relevance and engagement. Incorporating AI materials in virtual learning scenarios such as sNOOCs makes learning more dynamic and participatory. A mixed-methods approach was employed, combining data collection, statistical analysis, and evaluations by Peer Evaluation Circles comprising 200 researchers. The study results indicate a strong correlation between educational relevance and content clarity, as narratives rated as relevant tend to be perceived as clear and coherent. The study underscores the importance of a strong ethical foundation in using AI to ensure adequate and equitable understanding of content, even though sometimes ethics and clarity are not entirely aligned.

Keywords: artificial intelligence; digital narratives; visual design; ethical design.

INTRODUCCIÓN

Las narrativas enriquecidas son historias donde no sólo el texto forma parte del relato, sino que se apoyan también en otros recursos interactivos con el fin de que la experiencia sea más inmersiva y dinámica. Estos relatos en el ámbito digital tradicionalmente han combinado imágenes, vídeos, sonidos o animaciones, incluso enlaces que han convertido determinadas historias en una navegación hipervinculada. La inteligencia artificial (IA) ha revolucionado múltiples sectores, y su aplicación en el ámbito educativo se presenta como una de las innovaciones más prometedoras para mejorar la calidad del aprendizaje (Romero, 2024). En particular, la creación de narrativas mediante IA (Moscoso et al., 2024) ofrece un enfoque novedoso que puede transformar la manera en que se presenta y, como consecuencia, se asimila el contenido académico, alejándose así de los peligros que consideran la integración curricular de esta en los procesos de aprendizaje (Abbas et al., 2024).

La IA puede ayudar a crear narrativas más claras y coherentes, al analizar grandes cantidades de datos sobre el rendimiento y las necesidades de las y los estudiantes (Van Vaerenbergh, 2024). Esto permite a los sistemas de IA generar contenido que sigue una estructura lógica y adaptada al nivel de comprensión del estudiantado. La IA puede identificar y corregir errores o inconsistencias en el contenido educativo, mejorando la calidad del material presentado. Además de estas cuestiones, otra de las principales ventajas de la IA en la educación es su capacidad para personalizar el aprendizaje (López et al., 2023) en determinados entornos formativos virtuales. Los sistemas de IA también pueden adaptar las narrativas educativas a las temáticas especializadas que ofrecen determinados modelos como son *Nano Open Online Courses* (Clark, 2013; Basantes Andrade et al., 2020), concretamente su tipología sNOOC (Gil-Quintana, 2024), teniendo presente las necesidades individuales del alumnado que consume estos productos, sus intereses, estilos de aprendizaje y niveles de conocimiento previos. Esto puede resultar una experiencia de aprendizaje más relevante y atractiva, no sólo por la narrativa creada, sino también por el modelo específico y minimalista sNOOC, que puede ayudar a evitar las altas tasas de abandono en la formación masiva (Ratnasari et al., 2024). Si a estos modelos les añadimos prácticas de pedagogías innovadoras, de construcción colectiva del conocimiento y de empoderamiento del alumnado, donde estos contenidos son creados por las y los estudiantes, generan interactividad y compromiso (Fondevila-Gascón et al., 2024; Tamayo, 2019). El itinerario de aprendizaje donde se presentan estos contenidos no sólo aumenta la calidad narrativa de estos, su diseño e incluso la accesibilidad de los recursos, sino también en su difusión en determinados contextos postdigitales (Escaño, 2023), entre los que se encuentran las redes sociales.

Este estudio mixto analiza los resultados del proyecto de transferencia “Diseño de productos, planificación y evaluación educativa con IA” desarrollado en el curso 2023/2024 desde la Universidad Nacional de Educación a Distancia (UNED, España) para la Universidad Técnica de Machala (UTMACH, Ecuador). Este proyecto tuvo

como finalidad la integración de la IA en la creación de narrativas educativas por alumnado universitario, narrativas que son la base para el desarrollo posterior de sNOOC (Gil-Quintana, 2024) en la plataforma tmooc.es. La transferencia del conocimiento entre universidades españolas hacia universidades privadas extranjeras tiene gran importancia tanto para el ámbito académico como para el sector industrial y científico de ambos países. Esta colaboración, no sólo contribuye a potenciar la innovación en la educación a distancia, sino que mejora la calidad de los procesos formativos creando sinergias de colaboración, fomentando la calidad educativa y el prestigio internacional. La integración de la IA en la elaboración de narrativas que se presenta en este estudio parte de una experiencia de aprendizaje colaborativo (Tan et al., 2022), evaluando su impacto en aspectos fundamentales como la claridad, coherencia y personalización del aprendizaje (Baidoo-Anu y Ansah, 2023). Con más detalle, este estudio examina desde un planteamiento mixto, a través de Círculos de Análisis por Pares (Schön, 1983; Boud et al., 2021), cómo estos materiales creados de forma colaborativa con IA influyen en la relevancia educativa y la claridad de las narrativas generadas, destacando una fuerte correlación entre ambas.

Los resultados de este estudio sugieren que las narrativas valoradas como relevantes también tienden a ser percibidas como claras y coherentes. La alta correlación entre el uso eficaz de la IA y la personalización del aprendizaje indica que la IA tiene un impacto positivo en la adaptación del contenido a las necesidades individuales, lo que puede aumentar el compromiso y la efectividad formativa. La relación entre la ética y la transparencia en el uso de la IA también se revela crucial, con una correlación positiva que sugiere que, una sólida base ética, facilita una mejor comprensión del funcionamiento de la IA. A pesar de los avances, se observan excepciones donde la ética y la claridad en la explicación no siempre están alineadas, subrayando la necesidad de mejorar estas dimensiones para ofrecer una experiencia educativa coherente. La IA ofrece un potencial considerable para enriquecer narrativas personalizadas y claras, pero su implementación como se observa en los datos cualitativos debe ser realizada con una atención meticulosa a la ética, la claridad y la adaptación a las necesidades individuales para asegurar una integración efectiva y equitativa.

MÉTODOLÓGIA

Este estudio se presenta desde un enfoque multidisciplinario (Pereyra, 2022) para analizar la integración de la IA en la creación de narrativas para entornos sNOOC. El estudio se efectúa mediante el análisis de diferentes aspectos como la pertinencia y claridad del contenido, la facilidad de uso y accesibilidad, la precisión y relevancia del contenido generado, así como las consideraciones éticas y el compromiso con la calidad educativa. Partiendo de esta premisa, se establecen los siguientes objetivos:

- Objetivo 1 (O1): Evaluar el impacto de la IA en la personalización de contenidos educativos, analizando cómo esta tecnología de vanguardia contribuye a mejorar la relevancia y claridad percibida por las personas participantes en la experiencia de aprendizaje sNOOC.
- Objetivo 2 (O2): Analizar cómo el diseño visual y la experiencia de usuario influyen en la facilidad de uso y accesibilidad, y cómo afecta la percepción de precisión y relevancia en narrativas creadas con IA.
- Objetivo 3 (O3): Examinar si la transferencia en narrativas colaborativas con IA fortalece la percepción ética y claridad, explorando la relación entre ética y comprensión del funcionamiento de la IA.

Para responder de forma más precisa a este estudio y teniendo presente los objetivos anteriores, se presentan las siguientes hipótesis:

- Hipótesis 1 (H1): La utilización efectiva de la IA en el diseño de narrativas educativas mejora la personalización del aprendizaje y la percepción de relevancia y claridad.
- Hipótesis 2 (H2): El diseño visual y la experiencia de usuario tienen una correlación positiva significativa con la facilidad de uso y accesibilidad, mejorando la percepción de precisión y relevancia de la narrativa creada con IA.
- Hipótesis 3 (H3): Las narrativas creadas desde el aprendizaje colaborativo con IA demuestran un fuerte compromiso con la ética al ser más transparentes en la explicación de su funcionamiento, estableciéndose una correlación positiva entre ética y claridad.

El estudio de corte mixto presenta datos cuantitativos recopilados y organizados por una matriz de valoración como herramienta que nos permite evaluar y comparar la información de manera estructurada, posibilitando un análisis más claro y objetivo. A través de los instrumentos hemos cuantificado también las variables referidas a la calidad, como materiales educativos de las narrativas creadas, organizadas en torno a categorías presentadas en el apartado de análisis de resultados. El valor del Alfa de Cronbach (Rodríguez y Álvarez, 2020), refleja la confiabilidad de la matriz de valoración en su conjunto. Estos valores cercanos a 0,98 indican una alta consistencia interna y respaldan la validez y confiabilidad de las mediciones utilizadas en el estudio. En cuanto a la confiabilidad y consistencia de las mediciones, se analiza la relación entre cada ítem específico y el puntaje total del instrumento, con valores altos indicando una mayor contribución al constructo evaluado. Además, se han incorporado datos cualitativos procedentes del discurso generado en los foros de debate que se ha ido generando a la vez que se creaban las narrativas educativas.

Los datos se obtienen partiendo del proyecto de transferencia “Diseño de productos, planificación y evaluación educativa con IA” se desarrolló a lo largo de 18 sesiones comprendidas en los meses de febrero a mayo de 2024, tuvo como

finalidad la integración de la IA en la elaboración de narrativas educativas por alumnado universitario en enseñanzas de posgrado a través de la plataforma Moodle de UTMACH, cuyo trabajo fue la base para el desarrollo posterior de sNOOC (Gil-Quintana, 2024) en la plataforma externa a esta institución: [tmooc.es](https://mooc.es). El diseño de estas creaciones se realizó con aplicaciones como *MyHeritage*, *Voki*, *Fotor*, *Canva*, *Dall-e*, *Padlet*, *Tome* y *Elevanlabs*. Los datos se obtienen a partir de datos recopilados por medio de SPSS de 22 Círculos de Análisis de narrativas por Pares (CAP) formados por 200 personas que desarrollaron un total de 2110 análisis y las aportaciones en 154 foros de debate. Los CAP es una estructura colaborativa y democrática (Schön, 1983; Boud et al., 2021) donde un grupo de expertos investigadores analizan críticamente un hecho concreto desde el principio de revisión por pares, pero con un formato más interactivo, de valoración crítica y de enriquecimiento mutuo. Este conjunto de expertos, aún con una ubicación geográfica y contexto sociocultural ecuatoriano, a nivel de representatividad, está vinculado a distintos niveles del sistema educativo y años de experiencia en sus diferentes profesiones; además, sus áreas de especialización están relacionadas con Educación, Humanidades y Artes, Ciencias sociales, Educación comercial y derecho, Ciencias, Ingeniería industrial, Agricultura, Salud y servicios sociales. Más específicamente, la muestra se obtiene de un tamaño $n > 200$ personal investigador y se presentan de manera concisa y relevante para informar sobre la efectividad y desafíos de la implementación de la IA en la creación de narrativas educativas en torno a un determinado contenido destinado al desarrollo del itinerario de aprendizaje en sNOOC. La selección de esta muestra puede introducir varios sesgos como es ser expertos con experiencia en educación que puede influir en el análisis según criterios más técnicos o pedagógicos, ser homogéneos a nivel profesional compartiendo enfoques y marcos similares, ser más críticos o analíticos al estar acostumbrados a la evaluación de trabajos académicos y, de vital importancia, estar condicionados por opiniones acerca de cómo la IA puede influir negativa o positivamente en la educación.

En la Tabla 1, se recopilan los contenidos interdisciplinares creados por el alumnado como parte de su participación en las actividades formativas. Estos recursos han sido generados y alojados en la plataforma educativa en línea [tmooc.es](https://mooc.es), donde se encuentran disponibles para consulta y revisión como sNOOC, uniéndose a otras experiencias realizadas con esta perspectiva (Hueso-Romero et al., 2024). La contribución de las y los estudiantes ha sido clave en la construcción de estos materiales, que se presentan como parte integral de su proceso de aprendizaje colaborativo y que se encuentran publicados en abierto dentro de itinerarios de aprendizaje de la formación masiva.

Tabla 1
Círculo de Análisis de las narrativas por Pares

Código Creaciones	Narrativa creada
A01	Mentes Creativas Motivadoras
A02	Gamificación Educativa
A03	Producción Agropecuaria
A04	Mentes Creativas Negocios Inteligentes
A05	Agronomía
A06	Ciberseguridad
A07	Ciencias Experimentales
A08	Mundo App - Dóminus
A09	Salud Mental
A10	Bioquímica y Farmacia
A11	Enfermería Cuidados Artificiales
A12	Cambio Climático - Pérdida de Biodiversidad
A13	Medicina Salud (Morfofisiología)
A14	Tecnologías de la Información
A15	Actividad Física y Deporte
A16	Contabilidad y Auditoría Mercadotecnia
A17	FCE!
A18	Ingeniería de Alimentos
A19	Genius Alliance
A20	Ingeniería Civil
A21	Economía y Comercio Exterior
A22	Haciendo historia

Fuente: elaborado por los autores.

Para evaluar la distribución de los datos y verificar si estos siguen una distribución normal (Tapia y Cevallos, 2021), se realizó la prueba de normalidad utilizando la prueba de Kolmogorov-Smirnov (K-S), presentados en la Tabla 2 y Figura 1, debido a que $N=2110$, con una fiabilidad del 95 % por lo que ($\alpha = 0.05$). Para la prueba, se establecen las siguientes hipótesis: Hipótesis nula: (H_0) = $p > \alpha = 0.05$; e Hipótesis alternativa: (H_1): = $p < \alpha = 0.05$. El resultado de la prueba arrojó un valor de $p = 0$, lo que indica que es extremadamente improbable que los datos sigan una distribución normal. Dado que el valor de p es menor que el umbral común de significancia ($\alpha = 0.05$), se rechaza la hipótesis nula, lo que sugiere

que los datos no siguen una distribución normal. Este resultado justifica el uso de métodos estadísticos no paramétricos en los análisis posteriores. Este hallazgo tiene implicaciones metodológicas importantes al asumir normalidad de los datos en los procedimientos estadísticos. Al no cumplirse esta condición, se justifica el uso en este estudio de pruebas no paramétricas para el análisis, al no requerir normalidad en la distribución de los datos.

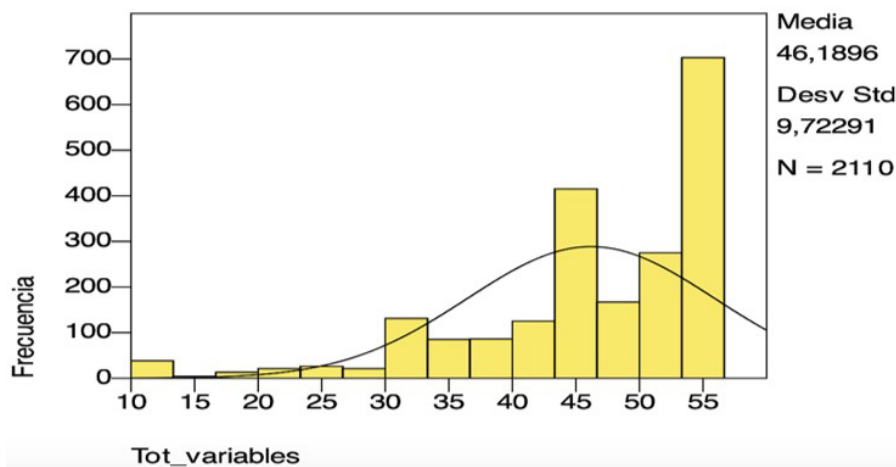
Tabla 2
Resultados prueba Kolmogorov-Smirnov

	N	Variable
		Mentes Creativas Motivadoras
Parámetros Normal	Media	46,19
	Desviación Estándar	9,72
Diferencias Más Extremas	Absoluto	,18
	Positivo	,18
	Negativo	-,15
Z de Kolmogorov-Smirnov		8,38
Sig. Asint (2-colas)		,000

Fuente: elaborado por los autores.

Figura 1

Histograma resultados curva normal



Fuente: elaborado por los autores.

El análisis de normalidad es un paso clave para la selección de las técnicas estadísticas utilizadas. Dado que la prueba K-S indicó una distribución no normal, se ha optado por enfoques no paramétricos que permiten un análisis válido de las variables estudiadas. La falta de ajuste de la curva presentada en la Figura 1 fortalece la validez de la opción metodológica, asegurando resultados estadísticamente sólidos y apropiados.

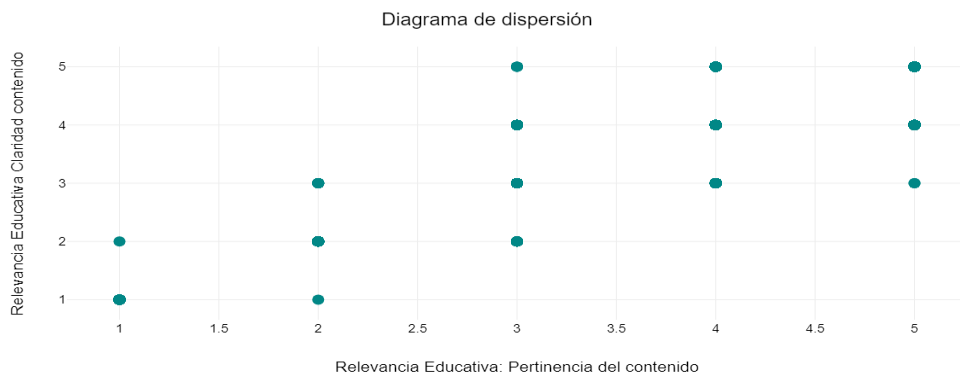
RESULTADOS

Categoría 1: Diseño de narrativas enriquecidas con IA: claridad y personalización

La correlación entre la relevancia educativa y la claridad del contenido es de 0.91, lo que indica una fuerte relación positiva entre ambas variables y una significancia estadística. Los CAP consideran las narrativas como relevantes y adecuadas también tienden a percibir las como claras y coherentes. Los CAP, referidos a A01 y A11 muestran esta relación consistente y favorable. En contraste, los grupos A15 y A17 tienen puntuaciones más bajas en ambas variables, lo que indica una mayor variabilidad en las opiniones sobre la claridad y la pertinencia, según se indica en la Figura 2.

Figura 2

Diagrama de dispersión, relevancia educativa y pertinencia de contenido



Fuente: elaborado por los autores.

El análisis de la pregunta sobre la relevancia de las narrativas muestra una valoración media (VM) como A01 (4.25) y A11 (4.1), con valoraciones altas y consistentes. Los CAP A06 (4.63) y A07 (4.54) también destacan por sus valores medios elevados y baja variabilidad; A03 (4.5) y A04 (4.13) muestran valoraciones positivas, pero con mayor dispersión; A18 (4.67) y A08 (4.67) perciben la narrativa como altamente pertinente; A15 (3.96) y A17 (3.93) tienen puntuaciones más bajas y dispersas.

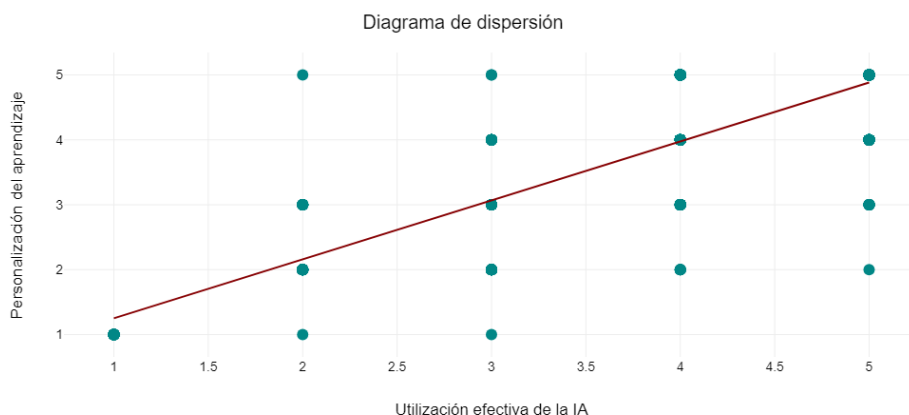
Respecto a la claridad y coherencia de las narrativas creadas, A06 (4.59) y A07 (4.47) tienen valoraciones medias altas con baja variabilidad. A01 (4.25) y A11 (4.09) también presentan valoraciones positivas. A03 (4.56) y A10 (4.07), muestran alta valoración, pero con más variabilidad. A18 (4.67) y A01 (5) tienen valoraciones muy altas, mientras que A15 (3.92) y A17 (3.88) presentan puntuaciones más bajas y mayor diversidad de opiniones.

La utilización efectiva de la IA en la educación no sólo mejora la claridad y coherencia del contenido presentado a través de la narrativa, sino que también permite una personalización del aprendizaje que se adapta a las necesidades individuales. Según los resultados analizados, los CAP valoran positivamente la efectividad de la IA en términos de claridad y coherencia del contenido, como A08 (VM 4.75), A06 (VM 4.51) y A10 (VM 4.36), también muestran una alta aceptación de la personalización del aprendizaje con IA, con puntuaciones medias de 4.58, 4.51 y 4.29, respectivamente. Por lo que una integración eficaz de la IA contribuye significativamente a adaptar los procesos de aprendizaje a las necesidades individuales de las personas participantes, mejorando la experiencia educativa de manera personalizada.

La alta correlación de 0.88 entre la utilización efectiva de la IA y la personalización del aprendizaje subraya esta relación positiva. Los que experimentan una implementación efectiva de la IA tienden a percibir una mayor adaptación del contenido en forma de narrativa a sus necesidades, lo que se traduce en una experiencia de aprendizaje más ajustada y efectiva. Sin embargo, algunas variaciones en las percepciones, como en A17 (VM 3.75) y A04 (VM 1.00), indican que ciertas narrativas experimentan dificultades en la aplicación de la IA, mostrando una mayor diversidad de opiniones y, en algunos casos, una percepción menos positiva sobre la relevancia educativa de la IA. Estos resultados destacan la importancia de una implementación cuidadosamente ajustada de la IA para maximizar tanto la claridad del contenido como la personalización del aprendizaje. Mientras que la mayoría valoran positivamente la integración de la IA, los casos con puntuaciones más bajas sugieren que se deben abordar áreas específicas para mejorar la experiencia de aprendizaje personalizada y asegurar que todos se beneficien de manera equitativa, según se manifiesta en la Figura 3.

Figura 3

Diagrama de dispersión, personalización del aprendizaje y utilización efectiva de la IA



Fuente: elaborado por los autores.

El análisis de regresión examina cómo la interactividad y retroalimentación (X) y la utilización efectiva de la narrativa creada con IA (M) influyen en el resultado del aprendizaje (Y), a través de dos modelos distintos que ilustran estas relaciones. La figura del mediador muestra la utilización efectiva de la IA (M) como una variable mediadora en la relación entre la interactividad y retroalimentación (X) y el resultado del aprendizaje (Y). A partir de la figura del mediador, se plantean dos

modelos a analizar. En el primer modelo, $M = aX + iM$, M es la variable dependiente y se predice a partir de X con $R=0.88$, la variable independiente, con la ecuación: $M = aX + iM$, según se presenta en la Tabla 3.

Tabla 3
Resumen del modelo

R	R²	Ajustado R²	Error estándar de la estimación	F	p
0.88	0.77	0.77	0.42	6866.57	<.001

Fuente: elaborado por los autores.

El coeficiente de correlación ($R = 0.88$) indica una relación fuerte entre (X) y (M). El coeficiente de determinación ($R^2 = 0.77$) muestra que el 77 % de la variabilidad en la utilización efectiva de la IA (M) puede ser explicada por la interactividad y retroalimentación (X). El coeficiente de regresión de X sobre M es 0.87 ($B=0.87$), y el coeficiente estandarizado ($Beta = 0.88$), lo que significa que una unidad adicional en la interactividad y retroalimentación aumenta la utilización efectiva de la IA en 0.87 uds. Un error estándar de la estimación de 0.42 refleja la precisión de las predicciones del modelo. El valor F de 6866.57 y un $p < .001$ confirman que el modelo es altamente significativo y que la relación entre X y M es estadísticamente robusta.

Por su parte, el segundo modelo investiga el impacto de X y M sobre el resultado en el aprendizaje (Y), con la ecuación: $Y = cX + bM + iY$. En este modelo, ($R = 1$) y ($R^2 = 1$), se indica una correlación perfecta y que el modelo explica el 100 % de la variabilidad en el resultado del aprendizaje (Y). El coeficiente para la interactividad y retroalimentación X es 1 ($B=1$), y el coeficiente estandarizado 1 ($Beta=1$), indicando que cada unidad de mejora en X resulta en una mejora directa de una unidad en Y . No obstante, el coeficiente para la utilización efectiva de la IA (M) es 0 ($B=0$), lo que sugiere que M no contribuye adicionalmente al resultado del aprendizaje Y cuando la interactividad y retroalimentación (X) están presentes. El efecto total de X sobre Y es 1, el efecto directo de X sobre Y también es 1, y el efecto indirecto de X sobre Y a través de M es 0, indicando que no hay influencia indirecta significativa de X en Y a través de M .

Ambos modelos demuestran que la interactividad y retroalimentación tienen un efecto determinante y directo en el resultado del aprendizaje. El primer modelo destaca que la calidad de la interactividad y retroalimentación mejora la utilización efectiva de la IA; mientras que el segundo modelo revela que, una vez controlado por la interactividad y retroalimentación, la utilización efectiva de la IA no añade un valor adicional significativo al resultado del aprendizaje. Por lo tanto, para optimizar el impacto educativo de la narrativa creada con IA, es crucial enfocarse en mejorar

los aspectos de interactividad y retroalimentación, ya que estos factores son más relevantes para el éxito en el aprendizaje que la propia utilización efectiva de la IA.

El estudio cualitativo pone de manifiesto que la IA mejora significativamente la personalización de contenidos educativos, adaptándose a las necesidades individuales del alumnado [F5:1]. Este hecho aumenta la relevancia y la claridad percibida, optimizando el proceso de aprendizaje y permitiendo que sea eficiente y efectiva [F7:12]; “La personalización del aprendizaje mediante IA mejora la relevancia y claridad percibida por los estudiantes, al adaptarse a sus estilos y ritmos de aprendizaje” [F3:12]. El diseño visual y la experiencia de usuario son clave para la facilidad de uso y accesibilidad de narrativas creadas con IA [F10:22]. Un diseño atractivo, intuitivo y bien estructurado mejora la percepción de precisión y relevancia [F9:12], fomentando la participación [F7:19] y garantizando una experiencia inclusiva para todas las personas [F11:23]; “La claridad en la narrativa y el diseño visual influyen en la percepción de precisión y relevancia de los contenidos educativos creados con IA” [F5:16].

Categoría 2: Percepción del diseño visual y experiencia de usuario en la narrativa creada con IA

El diseño visual y la experiencia del usuario (UX) son cruciales para la eficacia de las plataformas educativas, influyendo directamente en la facilidad de uso y la satisfacción del usuario. Los resultados indican que A01 y A02 tienen una alta valoración (VM 4.50, DT 0.00), lo que refleja una excelente percepción del diseño visual y UX. Asimismo, A03 y A04 muestran valoraciones altas (VM 4.50, DT 0.00; VM 4.23, DT 0.00), respectivamente. A05 tiene una valoración (VM 4.35, DT 0.78) que refleja una alta satisfacción y experiencia uniforme. Destaca A06 (VM 4.55, DT 0.78), indicando una percepción muy positiva del diseño visual y UX. A07 (VM 4.46, DT 0.77) y A08 (VM 4.36, DT 0.76), sugiere una valoración positiva y adecuada del diseño visual y UX. A09 (VM 4.22, DT 0.97) y A10 (VM 4.20, DT 0.98) muestra percepciones generalmente satisfactorias, aunque con cierta variabilidad. A11 (VM 4.00, DT 0.89) señala una percepción positiva con margen para mejoras. A12 (VM 4.13, DT 0.90) indica una apreciación positiva con necesidad de interfaces más especializadas. A13 (VM 4.24, DT 0.90) refleja una buena percepción con cierta variabilidad. A14 (VM 4.22, DT 1.00) indica una percepción positiva con variabilidad en herramientas y plataformas utilizadas. A15 (VM 3.97, DT 0.91) y A16 (VM 3.96, DT 0.95) reflejan percepciones positivas, pero con espacio para mejoras. A17 (VM 3.76, DT 1.16) evidencia una percepción relativamente baja y variada. Otros como A18 (VM 4.22, DT 0.85), A19 (VM 4.16, DT 0.91), A20 (VM 4.20, DT 0.87), y A21 (VM 4.14, DT 0.88) tienen puntuaciones favorables, reflejando satisfacción general con el diseño visual y UX. Así, la percepción del diseño visual y la experiencia del usuario varía entre unos, mostrando alta satisfacción, y otros, sobre áreas con puntuaciones más bajas. La consistencia en la experiencia del usuario y la adaptación del diseño a

las necesidades específicas de cada campo son claves para mejorar la efectividad de las herramientas educativas basadas en IA.

La facilidad de uso y accesibilidad de las herramientas basadas en IA varía significativamente entre áreas temáticas. La puntuación de A01 (VM 4.18, DT 0.83) refleja una experiencia positiva con algunas variaciones. A11 (VM 4.03, DT 0.84), muestra satisfacción general, pero con mayor variabilidad. A12 (VM 4.15, DT 0.92) muestra la necesidad de ajustes específicos. A10 y A09 (VM 4.27, DT 0.94) y (VM 4.24, DT 0.89) respectivamente, muestran buena percepción con moderada variabilidad. A06 (VM 4.51, DT 0.82) destaca con una experiencia uniforme y satisfactoria. En contraste, A16 (VM 3.93, DT 0.96), evidencia áreas de mejora. A03 muestra puntuación alta (VM 4.39, DT 0.82), reflejando experiencias positivas. A18 y A17 tienen puntuaciones más bajas (VM 3.93, DT 0.98) y (VM 3.84, DT 1.05), indicando más dificultades. Por ello, la facilidad de uso y accesibilidad varía ampliamente, sugiriendo la necesidad de ajustar y personalizar las herramientas de IA para satisfacer mejor las necesidades de cada comité, mejorando la consistencia y adaptabilidad.

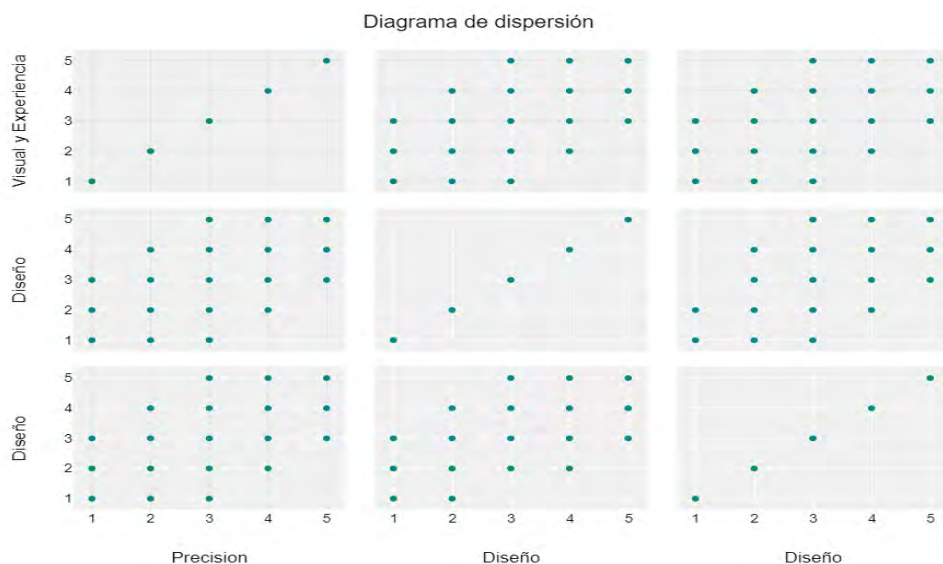
En cuanto a la precisión y relevancia de las narrativas creadas por IA varía entre áreas temáticas. A01 (VM 4.18, DT 0.82) denota alta percepción con cierta variabilidad. A11 (VM 4.03, DT 0.84) revela mayor variabilidad en percepciones. A12 (VM 4.15, DT 0.92) sugiere necesidad de ajustes. A10 y A09 (VM 4.27, DT 0.94) y (VM 4.24, DT 0.89), respectivamente, indican buena percepción con variabilidad. A06 (VM 4.51, DT 0.82) muestra percepción consistente y satisfactoria. A16 (VM 3.93, DT 0.96) señala preocupación por la precisión y relevancia. Los CAP A03 y A22 muestran puntuaciones altas (VM 4.39, DT 0.82) y (VM 4.44, DT 0.79), indicando relevancia y precisión. A18 y A17 tienen puntuaciones más bajas (VM 3.93, DT 0.98) y (VM 3.84, DT 1.05), indicando la necesidad de mejorar la alineación del contenido presentado en la narrativa. Con estos resultados, la percepción de la precisión y relevancia del contenido generado por IA varía ampliamente, sugiriendo la necesidad de personalizar y ajustar el contenido para mejorar la satisfacción en diferentes áreas temáticas. Mejora en precisión y relevancia es crucial para asegurar valor educativo de las herramientas de IA.

El análisis de correlaciones de estas 3 variables: “diseño visual y experiencia de usuario”, “facilidad de uso y accesibilidad” y “precisión y relevancia de contenido”, revela que el diseño visual y la experiencia del usuario tienen una fuerte relación positiva con la facilidad de uso y accesibilidad (0.73), señalando que, una mejora en estos aspectos, tiende a hacer las herramientas más fáciles de usar y accesibles. La correlación con la precisión y relevancia del contenido es moderada (0.68), indicando que un buen diseño visual también contribuye a una mejor percepción de la narrativa, aunque no tanto como la facilidad de uso. Por otro lado, la facilidad de uso y accesibilidad muestra una correlación moderada con la precisión y relevancia del contenido (0.67), indicando que, aunque influyen en la percepción del contenido, su impacto es menor en comparación con el diseño visual. En conjunto,

estos resultados destacan que el diseño visual y la experiencia del usuario son los factores más determinantes para una buena percepción general de las herramientas educativas basadas en IA, indicado en la Figura 4.

Figura 4

Diagrama de dispersión entre las variables diseño visual y experiencia de usuario en relación con la facilidad de uso y accesibilidad, precisión y relevancia de contenido



Fuente: elaborado por los autores.

La evaluación de la originalidad y creatividad de la narrativa generada por IA muestra una variedad de percepciones entre las diferentes narrativas, reflejando diferencias significativas en la apreciación de la innovación y creatividad del contenido proporcionado. Para profundizar en esta evaluación, se ha empleado un *Diagrama de Bland-Altman* (DBA) que permite analizar la concordancia entre las percepciones de originalidad y creatividad entre las diferentes narrativas creadas. Este diagrama ilustra la relación entre las medias de las puntuaciones de originalidad y creatividad y sus respectivas desviaciones estándar (DT). En general, las narrativas creadas por AO1 y AO6 tienen puntuaciones altas en originalidad y creatividad, con medias de (VM 4.14, DT 0.85) y (VM 4.48, DT 0.85) respectivamente. Estas creaciones parecen destacarse por ofrecer contenido innovador y creativo, siendo percibidos positivamente por las personas participantes. El DBA revela que la baja DT en AO6

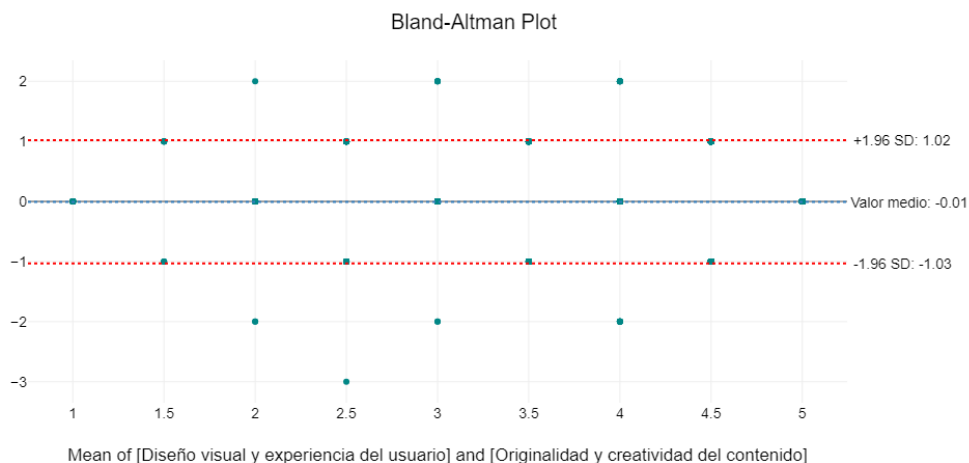
sugiere una percepción uniforme de la creatividad entre las y los participantes, mientras que en A01, la desviación es algo mayor, indicando una ligera variabilidad en la percepción. Los grupos A09 y A10 también tienen puntuaciones destacadas (VM 4.32, DT 0.83) y (VM 4.28, DT 0.85) respectivamente. Estos resultados muestran una alta valoración de la originalidad y creatividad del contenido, aunque la DT es algo más alta, lo que sugiere cierta diversidad en la percepción de la creatividad del contenido. El DBA confirma que la variabilidad en la percepción es mayor en estos grupos, comparado con A06, pero aún dentro de un rango aceptable. Por otro lado, A11 y A12 presentan puntuaciones medias más modestas (VM 3.93, DT 0.92; VM 4.01, DT 0.94) respectivamente. Aunque las personas consideran que el contenido es útil, hay una percepción menor de originalidad y creatividad. La alta DT indica una mayor variabilidad en la percepción, lo que podría sugerir que el contenido podría beneficiarse de enfoques más innovadores. El DBA muestra que la alta variabilidad en estas puntuaciones puede ser un factor en la menor percepción de creatividad. Por su parte, A16 muestra una puntuación media (VM 4.04, DT 1.03), indicando una percepción moderada de creatividad con una alta variabilidad en la respuesta de las y los usuarios. Esto apunta a que, mientras algunas personas encuentran el contenido creativo, otras podrían sentir que falta innovación. El diagrama refuerza la idea de alta variabilidad en las percepciones para A16. En contraste, se sitúan A08 y A03 con puntuaciones elevadas (VM 4.41, DT 0.80; VM 4.39, DT 0.78) respectivamente. Las narrativas creadas son percibidas como altamente originales y creativas por la mayoría de las personas que forman los CAP, con una percepción consistente de su calidad innovadora. A03 y A01 también destacan por sus puntuaciones (VM 4.33, DT 0.78; VM 4.57, DT 0.65) respectivamente, lo que indica una fuerte apreciación por la originalidad y creatividad del contenido. La baja DT en A01 sugiere una percepción muy uniforme de la creatividad del contenido, como se confirma en el DBA.

Por otro lado, las narrativas creadas por A17 y A11 tienen puntuaciones más bajas (VM 3.81, DT 1.08; VM 4.09, DT 0.94) respectivamente, lo que refleja una percepción menor y más variable en cuanto a la originalidad y creatividad del contenido. Estas narrativas podrían beneficiarse de una revisión para incrementar su nivel de innovación y creatividad. El diagrama destaca la variabilidad en la percepción de las creaciones, sugiriendo que se deben abordar áreas específicas para mejorar.

Finalmente, las narrativas creadas con puntuaciones extremadamente altas y bajas, como A19 con una puntuación perfecta de 5 y A20 con una puntuación de 1, subrayan la variabilidad extrema en la percepción de creatividad y originalidad. El DBA manifiesta que algunas narrativas son consideradas ejemplares en estos aspectos, mientras que otras no cumplen con las expectativas de los usuarios, mostrando una variabilidad significativa en las percepciones de originalidad y creatividad, según refleja la Figura 5.

Figura 5

Bland-Altman Plot. Diseño visual y experiencia del usuario / Originalidad y creatividad del contenido



Fuente: elaborado por los autores.

Los datos cualitativos nos ofrecen información interesante acerca de esta categoría. La IA para las personas participantes mejora la personalización de los contenidos educativos, adaptándose a las necesidades individuales y mejorando la relevancia y claridad de los contenidos [F12:6]. El storytelling digital, con un diseño visual atractivo y bien estructurado, fomenta habilidades de comunicación y colaboración, enriqueciendo la experiencia de aprendizaje [F5:4; F8:9]. De la misma forma, se destaca en las aportaciones cómo el rol del profesorado es clave para el éxito en estas innovaciones; “El rol del docente como guía es fundamental para el éxito de las experiencias de innovación docente en contextos digitales” [F7:5]. Referido al diseño visual y la experiencia de usuario se consideran fundamentales para la facilidad de uso y accesibilidad de las narrativas creadas con IA [F10:5]; un diseño intuitivo, atractivo y bien estructurado mejora la percepción de precisión y relevancia, mientras que una narrativa clara y coherente fomenta la participación [F10:8]; “El diseño visual y la experiencia de usuario deben ser intuitivos y accesibles, fomentando la facilidad de uso y la participación” [F11:3].

Categoría 3: Ética, responsabilidad y compromiso

El análisis de las consideraciones éticas y la claridad en la explicación del funcionamiento de la inteligencia artificial (IA) en las narrativas creadas revela

diferencias significativas en la percepción entre las personas participantes. En términos de ética, las narrativas A07 y A06 se destacan con altas puntuaciones (VM 4.55, DT 0.76) y (VM 4.54, DT 0.83), respectivamente, indicando un fuerte compromiso con las normas éticas en la IA. La baja desviación típica (DT) en A07 muestra una percepción consistente, mientras que en A06 hay una ligera variabilidad, aunque todavía positiva. Las narrativas A09 y A10 también obtienen altas puntuaciones (VM 4.44, DT 0.83; VM 4.49, DT 0.79), demostrando un sólido enfoque en ética. Por el contrario, A11 y A12 presentan puntuaciones más bajas (VM 4.15, DT 0.95; VM 4.21, DT 0.96), con mayor variabilidad en la percepción. Las narrativas A08 y A03, con sus respectivas puntuaciones (VM 4.46, DT 0.73; VM 4.33, DT 0.78), muestran una gestión efectiva de la ética; mientras que A17 y A16 presentan puntuaciones moderadas (VM 4.07, DT 1.03) y (VM 3.99, DT 1.04), con mayor variabilidad. Narrativas como A19 y A20 obtienen la puntuación máxima de 5, reflejando una percepción excelente, mientras que A21 y A13 tienen puntuaciones bajas de (VM 2.00, DT -) y (VM 3.00, DT -), indicando deficiencias en el tratamiento de la ética.

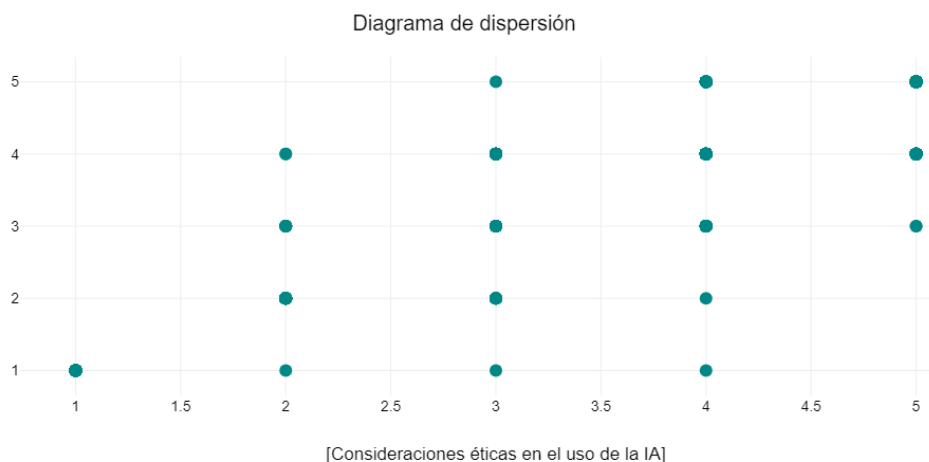
En cuanto a la transparencia y explicación del funcionamiento de la IA, las narrativas A07 y A06 tienen las puntuaciones más altas (VM 4.51, DT 0.78; VM 4.52, DT 0.87), demostrando efectividad en la claridad. La baja DT en A07 refleja una percepción uniforme, mientras que en A06 hay una ligera variabilidad. Las narrativas A09 y A10 (VM 4.38, DT 0.82; VM 4.44, DT 0.83), respectivamente, muestran una buena percepción de la transparencia. A08 y A03 (VM 4.46, DT 0.7; VM 4.44, DT 0.70), indican una alta transparencia, mientras que A01 y A13 (VM 4.50, DT 0.74; VM 4.18, DT 0.88), reflejan satisfacción con la explicación de la IA. Las narrativas A16 y A17 (VM 4.08, DT 0.95; VM 3.93, DT 1.03), muestran una mayor variabilidad en la percepción de la transparencia. Finalmente, narrativas como A19 y A20 obtienen la puntuación máxima de 5, indicando una excelente percepción, mientras que A21 y A13 tienen puntuaciones bajas de (VM 2.00, DT -; VM 3.00, DT -), sugiriendo una falta de claridad en la explicación de la IA.

La correlación entre las consideraciones éticas y la transparencia en la explicación del funcionamiento de la IA se visualiza en el diagrama de dispersión. Se observa cómo las narrativas que obtienen altas puntuaciones en ética también tienden a mostrar una buena claridad en la explicación del funcionamiento de la IA. Por ejemplo, las narrativas A07 y A06, que destacan en ética (VM 4.55, DT 0.76; VM 4.54, DT 0.83), también obtienen altos valores en transparencia (VM 4.51, DT 0.78; VM 4.52, DT 0.87) respectivamente. Esto demuestra una fuerte correlación positiva entre un sólido enfoque ético y una comunicación clara de la IA. Sin embargo, el diagrama también señala excepciones. Narrativas como A11 y A12, que tienen puntuaciones más bajas en ética (VM 4.15, DT 0.95; VM 4.21, DT 0.96), muestran una variabilidad similar en la transparencia (VM 4.08, DT 0.95; VM 3.93, DT 1.03) en la explicación de la IA, respectivamente. Estas narrativas reflejan una menor coherencia entre la ética y la claridad en la explicación, evidenciando que,

aunque la ética y la transparencia a menudo se alinean, hay narrativas donde estas dimensiones no están igualmente desarrolladas. En concreto, A07 y A06 sobresalen tanto en ética como en transparencia, mientras que A17 y A21 requieren mejoras en la integración de estos temas. La correlación positiva observada en el diagrama de dispersión refuerza la importancia de abordar simultáneamente las consideraciones éticas y la claridad en la explicación del funcionamiento de la IA para ofrecer una experiencia educativa completa y coherente.

Figura 6

Diagrama de dispersión, correlación consideraciones éticas, transparencia y explicación del funcionamiento de la IA



Fuente: elaborado por los autores.

Los datos cuantitativos presentados en esta categoría se ven reforzados por la investigación cualitativa en los foros. De estos datos se desprende que la IA transforma la educación al personalizar el aprendizaje, adaptándose a las necesidades del alumnado y proporcionando un estímulo inmediato [F3:3; F7:4; F10:14]. No podemos olvidar que la implementación de este recurso debe ser ética, garantizando la transparencia, la privacidad y evitando los sesgos [F3:7; F5:7; F14:3]. Se observa en las aportaciones de las personas participantes la demanda de la democratización del acceso a la IA para asegurar la equidad educativa, apostando por el desarrollo de procesos que ayuden a mitigar la brecha digital y fomentando la labor de los agentes educativos para concienciar en el uso responsable, el pensamiento crítico y la justicia en educación [F4:16; F8:8]. Como señalan algunas aportaciones en los foros: “La ética en la IA plantea desafíos como la transparencia, la equidad y la justicia en su aplicación, especialmente en la recopilación y uso de datos” [F1:4]; “Es necesario

establecer marcos regulatorios y mecanismos de supervisión para garantizar un uso ético y responsable de la IA en la educación” [F3:6]; “La ética en la IA representa uno de los desafíos más importantes de nuestra era digital, requiriendo atención y regulación para evitar sesgos y discriminaciones” [F6:8].

DISCUSIÓN

Los resultados de este estudio destacan la importancia de la IA en la creación de narrativas educativas interdisciplinarias (Yusuf et al., 2024) para los entornos sNOOC (Hueso-Romero et al., 2024), evidenciando tanto sus beneficios como los desafíos asociados. La fuerte correlación entre la claridad del contenido y la relevancia educativa sugiere que las narrativas generadas por IA son efectivas, a nivel lingüístico, en la comunicación de conceptos, percepción de la relevancia y claridad (H1), alineándose con investigaciones previas que enfatizan la claridad como clave para el aprendizaje (Ahmad et al., 2023; Baidoo-Anu y Ansah, 2023). No obstante, las variaciones en las puntuaciones de narrativas indican que no todas las implementaciones de IA son igualmente efectivas, lo que destaca la necesidad de un enfoque cuidadoso en su diseño y selección, exigiendo así una evaluación posterior (Félix y Webb, 2024).

La formación de agentes educadores en el uso de estas herramientas es esencial para maximizar su potencial y alinearlas con los objetivos pedagógicos, así como mejorar los resultados de la enseñanza y el aprendizaje en la educación universitaria (Eager y Brunton, 2023), construyendo colectivamente el conocimiento (Lee et al., 2023) y apostar la erradicación de la brecha digital (Forero-Corba y Negre Bennasar, 2024).

Desde una perspectiva ética, la fuerte correlación entre las consideraciones morales y la transparencia en la explicación del funcionamiento de la IA es notable (H3). La IA personaliza el aprendizaje y mejora la educación, pero debe usarse con ética, garantizando la transparencia y la equidad (Gallent-Torres et al., 2024). Las narrativas que priorizan los principios éticos también tienden a ser más comprensibles, lo que sugiere que un enfoque sólido en ética puede facilitar la comprensión por parte de las y los estudiantes. No obstante, las excepciones, demuestran que la ética y la claridad no siempre coinciden, lo que resalta la necesidad de abordar estas discrepancias.

La alta relación entre el uso eficaz de la inteligencia artificial y la personalización del aprendizaje sugiere que la tecnología puede ajustar el contenido a las necesidades individuales, mejorando el compromiso del alumnado (H1) con el aprendizaje (Bond et al., 2020). Los datos han puesto de manifiesto una fuerte correlación de 0.88 lo que se respalda la idea de que una implementación educada de la IA aporta una adaptabilidad en los procesos de aprendizaje individualizado; a mayor eficacia en el uso de la IA mayor personalización de la experiencia. Sin embargo, se ponen de

manifiesto también dificultades en su implementación, lo que indica que se requieren ajustes para maximizar la claridad y la personalización (H3).

El diseño visual y la experiencia de usuario son factores influyentes para la percepción de materiales creados con IA (H2) desde el aprendizaje colaborativo (Tan et al., 2022). Este factor muestra una fuerte relación positiva con la facilidad de uso y accesibilidad, sugiriendo una percepción positiva del diseño visual para mejorar la misma (0,7); aunque este diseño contribuye a mejorar la precisión y relevancia del contenido (H2), su influencia es más moderada (0.68). Partiendo de estos datos se vislumbra la necesidad de interfaces intuitivas y diseños atractivos para conseguir el máximo potencial de las narrativas generadas por IA.

Este estudio también sugiere la necesidad de futuras investigaciones sobre la efectividad de la IA en diversos contextos educativos y su comparación con métodos tradicionales. De hecho, la originalidad y la creatividad es percibida de diferentes formas dentro de las narrativas analizadas. La variabilidad en esta percepción hace considerar que, no solo por el hecho de utilizar IA una narrativa es innovadora, sino que se hace imprescindible incidir también en la diversificación y la adaptabilidad del contenido generado. Se recomienda seguir investigando sobre cómo optimizar el uso de la IA para la mejora de las prácticas educativas (Arango Pérez et al., 2024).

En resumen, la integración de la IA en narrativas educativas presenta un gran potencial, pero su revisión debe ser crítica (Delgado et al., 2024) y su implementación reflexiva, cuidadosamente diseñada y ética para garantizar resultados educativos óptimos. La efectividad del uso de la IA depende de su claridad, personalización (Guzmán Matute et al., 2024) y diseño visual; por ende, de la interactividad y la retroalimentación de estas narrativas emergen como factores clave en el éxito de los procesos de aprendizaje. No podemos olvidar que este éxito siempre estará condicionado por estrategias didácticas que fomenten la participación y la construcción colectiva del conocimiento como el aprendizaje personalizado (Luckin et al., 2016), la gamificación (Sailer y Homner, 2020), el Aprendizaje Basado en Problemas, el Aprendizaje Colaborativo (Dillenbourg, 2015), la Evaluación Formativa, etc.

CONCLUSIONES

La integración de la IA en la creación de narrativas educativas ha demostrado ser una herramienta valiosa para mejorar la claridad y personalización del aprendizaje en un entorno postdigital. La IA ha posibilitado que estas narrativas sean más personalizadas, interactivas, accesibles y dinámicas, abriendo múltiples posibilidades en la forma de presentar los relatos, contar historias y experimentar incorporaciones novedosas a las mismas. De esta forma, estas narrativas integradas en entornos sNOOC, no sólo facilitan el empoderamiento del alumnado, sino también la transformación de este modelo formativo en una experiencia inmersiva.

Los resultados de este estudio indican que las narrativas creadas con IA facilitan la comprensión de conceptos complejos y se adaptan a las necesidades individuales del alumnado. No obstante, es fundamental abordar las variaciones en la efectividad de estas narrativas, lo que subraya la necesidad de un diseño cuidadoso y de la formación continua del profesorado universitario en el uso de estas herramientas. La consideración de la personalización del aprendizaje y de los principios éticos en la implementación de la IA es crucial para garantizar una experiencia educativa comprensible y comprometida, aspectos que se podrían consolidar a través de estudios cualitativos.

Se recomienda que futuras investigaciones exploren la efectividad de la IA en diferentes contextos educativos y su comparación con métodos tradicionales, asegurando así una integración reflexiva y ética que potencie la calidad del aprendizaje. La investigación futura deberá centrarse en el desarrollo de marcos normativos, de herramientas de IA más transparentes y la mejora de experiencia de usuario en las narrativas creadas. Este nuevo lenguaje puede ser enriquecido a nivel educativo con la cocreación, la multimodalidad, el contenido dinámico, etc., que ayudarán a mejorar la calidad educativa y los itinerarios de aprendizaje.

Agradecimientos

Proyecto: “Diseño de productos, planificación y evaluación educativa con Inteligencia Artificial”, Vicerrectorado de Investigación, Transferencia y Divulgación Científica; Oficina de Transferencia del Conocimiento de la Universidad Nacional de Educación a Distancia. Grupo de Investigación SMEMIU (UNED).

REFERENCIAS

- Abbas, M., Jam, F. A. y Khan, T. I. (2024). Is it harmful or helpful? Examining the causes and consequences of generative AI usage among university students. *International Journal of Educational Technology in Higher Education*, 21(1). <https://doi.org/10.1186/s41239-024-00444-7>
- Ahmad, N., Murugesan, S. y Kshetri, N. (2023). Generative Artificial Intelligence and the Education Sector. *Computer*, 56(6), 72-76. <https://doi.org/10.1109/MC.2023.3263576>
- Arango Pérez, R., Lovato Sagrado, A., Ortega González, E. y Fontán de Bedout, L. (2024). Implicaciones filosóficas, éticas y pedagógicas del uso de la Inteligencia Artificial en Educación. *Digital Education Review*, 45. <https://doi.org/10.1344/der.2024.45.29-36>
- Baidoo-Anu, D. y Ansah, L. O. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI*, 7(1), 52-62. <https://doi.org/10.2139/ssrn.4337484>
- Basantes Andrade, A., Cabezas González, M. y Casillas Martín, S. (2020). Los nano-MOOC como herramienta de formación en competencia digital docente. *Revista Ibérica de Sistemas e Tecnologías de*

- Informação, E32, 202-214. <https://doi.org/10.14201/gredos.144006>
- Bond, M., Buntins, K., Bedenlier, S., Zawacki-Richter, O. y Kerres, M. (2020). Mapping research in student engagement and educational technology in higher education: a systematic evidence map. *International Journal of Educational Technology in Higher Education*, 17, 2. <https://doi.org/10.1186/s41239-019-0176-8>
- Boud, D., Cohen, R. y Sampson, J. (2021). *Peer Learning in Higher Education: Learning from and Each Other*. Routledge.
- Clark, D. (2013, January 29). MOOCs: taxonomy of 8 types of MOOC. *Donald Clark Plan B*. <http://donaldclarkplanb.blogspot.com/2013/01/moocs-taxonomy-of-8-types-of-mooc.html>
- Delgado, N., Campo Carrasco, L., Sainz de la Maza, M. y Etxabe-Urbieta, J. M. (2024). Aplicación de la Inteligencia Artificial (IA) en Educación: Los beneficios y limitaciones de la IA percibidos por el profesorado de educación primaria, educación secundaria y educación superior. *Revista Electrónica Interuniversitaria de Formación del Profesorado*, 27(1), 207-224. <https://doi.org/10.6018/reifop.577211>
- Dillenbourg, P. (2015). *Orchestration Graphs Modeling Scalable Education*. EPFL Press.
- Eager, B. y Brunton, R. (2023). Prompting Higher Education Towards AI-Augmented Teaching and Learning Practice. *Journal of University Teaching and Learning Practice*, 20(5). <https://doi.org/10.53761/1.20.5.02>
- Escaño, C. (2023). Educación Postdigital: un enfoque desde la pedagogía crítica y mediática para un contexto post-COVID19. *Revista Mediterránea de Comunicación*, 14(2), 243-257. <https://doi.org/10.14198/MEDCOM.23899>
- Félix, J. y Webb, L. (2024). Use of artificial intelligence in education delivery and assessment. *Research Briefing, UK Parliament*. <https://doi.org/10.58248/PN712>
- Fondevila-Gascón, J.-F., Martín-Guart, R. F., Carreras Alcalde, M. y Vila Márquez, F. (2024). Interactividad en educación: Aplicaciones de la Inteligencia Artificial y el HbbTV. *Didáctica, innovación y multimedia*, 42. <https://ddd.uab.cat/record/293485>
- Forero-Corba, W. y Negre Bennasar, F. (2024). Técnicas y aplicaciones del Machine Learning e Inteligencia Artificial en educación: una revisión sistemática. *RIED-Revista Iberoamericana de Educación a Distancia*, 27(1), 1-34. <https://doi.org/10.5944/ried.27.1.37491>
- Gallent-Torres, C., Arenas Romero, B., Vallespir Adillón, M. y Foltýnek, T. (2024). Inteligencia Artificial en educación: entre riesgos y potencialidades. *Práxis Educativa*, 19, e23760. <https://doi.org/10.5212/PraxEduc.v.19.23760.083>
- Gil-Quintana, J. (2024). Los sNOOC, un modelo educativo masivo, abierto y en línea, desde el empoderamiento social y el enfoque minimalista. *Revista Latinoamericana Ogmios*, 4(11). <https://doi.org/10.53595/rlo>
- Guzmán Matute, N. del R., Álvarez González, N. F. y Pacheco Pérez, X. O. (2024). Alcances y limitaciones de la IA en educación. *RECIMUNDO*, 8(1), 215-223. <https://doi.org/10.26820/recimundo/8.1.ene.2024.215-223>
- Hueso Romero, J. J., García Blázquez, E. y Gil Quintana, J. (2024). El Microaprendizaje servicio a través de los sNOOC: propuesta formativa para personas en riesgo de exclusión en México. *EduTec. Revista Electrónica de Tecnología Educativa*, 88, 42-61. <https://doi.org/10.21556/edutec.2024.88.3101>
- Lee, A. V. Y., Tan, S. C. y Teo, C. L. (2023). Designs and practices using generative AI for sustainable student discourse and knowledge creation. *Smart Learning*

- Environments*, 10, 59. <https://doi.org/10.1186/s40561-023-00279-1>
- López, H. L. L., Escalera, A. R. y García, C. R. C. (2023). Personalización del aprendizaje con inteligencia artificial en la educación superior. *Revista Digital de Tecnologías Informáticas y Sistemas*, 7(1), 123-128. <https://doi.org/10.61530/redtis.vol7.n1.2023.165.123-128>
- Luckin, R., Holmes, W., Griffiths, M. y Forcier, L. B. (2016). *Intelligence Unleashed: Un argumento a favor de la IA en la educación*. Pearson.
- Moscoso, M. C. M., Chacón, H. F. C., Oquendo, F. M. M., Logaña, M. A. P. y Ramos, E. G. C. (2024). Narrativas Digitales en el Área de Ciencias Sociales: Integración de Storytelling e Inteligencia Artificial. *Ciencia Latina Revista Científica Multidisciplinar*, 8(2), 4192-4209. https://doi.org/10.37811/cl_rcm.v8i2.10830
- Pereyra, L. E. (2022). *Metodología de la investigación*. Klik.
- Ratnasari W., Chou, T. C. y Huang, C. (2024). From hype to reality: the changing landscape of MOOC research. *Library Hi Tech*. <https://doi.org/10.1108/LHT-07-2023-0320>
- Rodríguez, J. R. y Álvarez, M. R. (2020). Calcular la fiabilidad de un cuestionario o escala mediante el SPSS: El coeficiente alfa de Cronbach. *REIRE: Revista d'Innovació i Recerca en Educació*, 13(2), 8. <https://doi.org/10.1344/reire2020.13.230048>
- Romero, M. Á. M. (2024). Aplicaciones de la Inteligencia Artificial para la investigación y la innovación en la educación superior. *Revista Social Fronteriza*, 4(4), e44336. [https://doi.org/10.59814/resofro.2024.4\(4\)336](https://doi.org/10.59814/resofro.2024.4(4)336)
- Sailer, M. y Homner, L. (2020). La gamificación del aprendizaje: un metaanálisis. *Educational Psychology Review*, 32, 77-112. <https://doi.org/10.1007/s10648-019-09498-w>
- Schön, D. A. (1983). *El profesional reflexivo. Cómo piensan los profesionales cuando actúan*. Paidós.
- Tamayo, J. L. R. (2019). Realidad extendida, interactividad y entornos inmersivos 3D: Revisión de la literatura y proyecciones. *Actas Icono 14*, 1(1), 396-415. <https://icono14.net/ojs/index.php/actas/article/view/1330>
- Tan, S. C., Lee, A. V. Y. y Lee, M. (2022). A systematic review of artificial intelligence techniques for collaborative learning over the past two decades. *Computers and Education: Artificial Intelligence*, 100097. <https://doi.org/10.1016/j.caeai.2022.100097>
- Tapia, C. E. F. y Cevallos, K. L. F. (2021). Pruebas para comprobar la normalidad de datos en procesos productivos: Anderson-Darling, Ryan-Joiner, Shapiro-Wilk y Kolmogórov-Smirnov. *Societas*, 23(2), 83-106. <https://revistas.up.ac.pa/index.php/societas/article/view/2302>
- Van Vaerenbergh, S. (2024). Inteligencia artificial para potenciar la creatividad y la innovación educativa. *Revista INFAD de Psicología. International Journal of Developmental and Educational Psychology*, 1(1), 507-513. <https://doi.org/10.17060/ijodaep.2024.n1.v1.2644>
- Yusuf, A., Pervin, N., Román-González, M. y Noor, N. M. (2024). IA generativa en educación e investigación: revisión de mapeo sistemático. *Review of Education*, 12(2). <https://doi.org/10.1002/rev3.3489>

Fecha de recepción del artículo: 1 de diciembre de 2024

Fecha de aceptación del artículo: 4 de marzo de 2025

Fecha de aprobación para maquetación: 26 de marzo de 2025

Fecha de publicación en OnlineFirst: 9 de abril de 2025

Fecha de publicación: 1 de julio de 2025

Inteligencia Artificial y chatbots para una educación superior sostenible: una revisión sistemática

Artificial Intelligence and chatbots for sustainable higher education: a systematic review



-  Eva Jiménez-García - *Universidad Europea de Madrid, UEM (España)*
-  Judit Ruiz-Lázaro - *Universidad Nacional de Educación a Distancia, UNED (España)*
-  Sonia Martínez-Requejo - *Universidad Europea de Madrid, UEM (España)*
-  Sara Redondo-Duarte - *Universidad Complutense de Madrid, UCM (España)*

RESUMEN

La influencia de la Inteligencia Artificial y de los Chatbots han generado una transformación significativa en los procedimientos de enseñanza y aprendizaje en Educación superior. Este estudio tiene como objetivo analizar la producción científica sobre el papel que desempeñan los Chatbots basados en inteligencia artificial en el contexto de la educación superior y su vinculación con los Objetivos de Desarrollo Sostenible (ODS). Para ello, se lleva a cabo una revisión sistemática siguiendo la metodología establecida en la Declaración PRISMA. La búsqueda inicial proporcionó un conjunto inicial de 136 referencias, de las cuales, tras un proceso de depuración, se seleccionaron 42 para un riguroso análisis bibliométrico detallado (co-palabras, citas y fuentes) y su relación con los ODS. Los resultados evidencian patrones de co-ocurrencia en temáticas clave, como el uso de Chatbots en diferentes áreas de educación superior, sus desafíos e impacto en la motivación del alumnado. En relación con los ODS, el ODS 4 (Educación de Calidad) es el más abordado, seguido por el ODS 3 (Salud y Bienestar), ODS 7 (Energía Asequible y No Contaminante), ODS 10 (Reducción de Desigualdades) y ODS 13 (Acción Climática). Como conclusión, este estudio ofrece una visión sobre las tendencias en el uso de Chatbots en educación superior, evidenciando su impacto y potencial en el desarrollo educativo universitario. Se espera que este estudio contribuya a la elaboración de políticas educativas orientadas a una integración sostenible de la inteligencia artificial en el entorno universitario.

Palabras clave: inteligencia artificial; chatbots; educación superior; sostenibilidad.

ABSTRACT

The influence of artificial intelligence and chatbots has significantly transformed teaching and learning processes in higher education. This study aims to analyze the scientific production on the role of AI-based chatbots in higher education and their connection to the Sustainable Development Goals (SDGs). To achieve this, a systematic review was conducted following the methodology established in the PRISMA Declaration. The initial search yielded 136 references, of which, after a screening process, 42 were selected for a detailed bibliometric analysis (co-word analysis, citations, and sources) and their relationship with the SDGs. The results reveal co-occurrence patterns in key topics, such as the use of chatbots in different areas of higher education, their challenges, and impact on student motivation. Regarding the SDGs, SDG 4 (Quality Education) is the most frequently addressed, followed by SDG 3 (Good Health and Well-being), SDG 7 (Affordable and Clean Energy), SDG 10 (Reduced Inequalities), and SDG 13 (Climate Action). In conclusion, this study provides an overview of trends in the use of chatbots in higher education, highlighting their impact and potential in university-level educational development. This study is expected to contribute to the development of educational policies that promote the sustainable integration of artificial intelligence in the university setting.

Keywords: artificial intelligence; chatbots; higher education; sustainability.

INTRODUCCIÓN

En las últimas décadas, la tecnología educativa ha transformado los procedimientos de enseñanza y aprendizaje en educación superior, siendo una de las innovaciones más relevantes en este ámbito (Zorrilla-Puerto et al., 2023). En este sentido, destaca la reciente cantidad de estudios que avalan la influencia de la inteligencia artificial (de ahora en adelante IA) en el ámbito universitario y que está marcando una transformación significativa en la forma en que se concibe el proceso educativo (Jiménez-García et al., 2024). Este creciente interés se ha visto impulsado por la disponibilidad generalizada de grandes modelos de lenguaje (Large Language Model, LLM) lo que ha generado nuevas oportunidades para enriquecer la enseñanza, adaptar el proceso de aprendizaje y ofrecer un apoyo más eficiente tanto al profesorado como al alumnado.

En esta línea, destaca la introducción de sistemas inteligentes como los Chatbots, Gemini, ChatGPT de OpenAI (Zou y Huang, 2023) y Copilot de Microsoft (Barrett y Pack, 2023), capaces de adaptarse a las necesidades individuales del alumnado. Estas herramientas de IA Generativa (GenAI), concretamente los Chatbots basados en algoritmos de procesamiento del lenguaje natural y aprendizaje automático, tienen la capacidad de interactuar de manera contextualizada y personalizada con el alumnado, facilitando interacciones personalizadas y contextuales entre estudiantes y sistemas automatizados de aprendizaje (Redondo-Duarte et al., 2024) y aportando respuestas instantáneas y adaptadas, lo que facilita la comunicación estudiante-máquina (Okonkwo y Ade-Ibijola, 2021). En este sentido, ofrecen a los usuarios la capacidad de crear textos inmediatos, u otros formatos, sobre una variedad de temas a partir de indicaciones simples (Rahim et al., 2022). Los Chatbots desempeñan múltiples roles en el ámbito educativo superior, como orientación inicial al alumnado, tutoría virtual, asistencia técnica en el uso de plataformas educativas, apoyo académico personalizado, asesoramiento académico, entre otros (Anghelescu y Nicolaescu, 2018; Alkhoori et al., 2020; Allison, 2012; Dube y Jacobs, 2023).

Recientemente, diversas investigaciones han demostrado que los Chatbots están impactando significativamente en el ámbito educativo (Ojeda et al., 2023). Por ejemplo, estudios como los llevados a cabo por Kuhail et al. (2023) y Pérez et al. (2020) han subrayado aspectos como la mejora de la experiencia y resultados de aprendizaje. Otros han observado que la satisfacción percibida por los usuarios con los Chatbots influye directamente en su intención de seguir utilizándolos, mientras que la innovación personal es un factor clave en este proceso (Tian et al., 2024). Además, los Chatbots han demostrado mejorar el aprendizaje personalizado y adaptativo, aumentar la satisfacción del alumnado y fortalecer los servicios virtuales dentro de la educación superior (Oqaidi et al., 2024). Desde una perspectiva operativa, estas herramientas han mostrado un gran potencial para agilizar tareas administrativas, enriquecer las experiencias de aprendizaje en entornos universitarios (Sain et al., 2024). Asimismo, su integración en plataformas de recomendación educativa ha

resultado en un alto nivel de influencia en la toma de decisiones de los estudiantes, con un 86,15 % de ellos basando sus elecciones en las sugerencias proporcionadas por estos sistemas (Kingchang et al., 2024).

Otro caso es el de EduChat, desarrollado por Dinh y Tran (2023), cuyo Chatbot responde preguntas relacionadas con la universidad utilizando un enfoque híbrido que combina métodos basados en reglas y técnicas de aprendizaje automático. Otro caso es el estudio de Rodríguez et al. (2021), que analiza el uso de Chatbot para el aprendizaje autónomo, facilitando la educación personalizada. Además, Alt e Ibolya (2021) identificaron usuarios potenciales de Chatbots bancarios, lo que puede extrapolarse a la educación superior para diseñar Chatbots que se adapten a las necesidades específicas de diferentes grupos de estudiantes. Estos ejemplos ilustran cómo los Chatbots pueden integrarse en diversas áreas de la educación superior, desde la asistencia académica hasta la gestión administrativa, mejorando así la eficiencia y la satisfacción de los estudiantes.

Ejemplos adicionales de estos avances incluyen estudios como el de Malik et al. (2021), que evidencian mejoras en el rendimiento académico, o el de Kuhail et al. (2023), que destaca el impacto positivo de los Chatbots en el aprendizaje y el compromiso estudiantil en comparación con métodos tradicionales de enseñanza. La investigación realizada por Vázquez-Cano et al. (2021) también pone de manifiesto su uso pedagógico en la mejora del acceso universitario en la asignatura de Lengua Española, mientras que Romero-Rodríguez et al. (2023) subrayan el valor de ChatGPT para fomentar el pensamiento complejo. Adicionalmente, Xiao y Zhi (2023) exploran su contribución en la enseñanza de idiomas, Weeks et al. (2022) examinan su utilidad en la comunicación sobre salud pública, y Moldt et al. (2023) resaltan su importancia en la formación médica.

Por otro lado, la integración de la IA y los Chatbots en la educación superior, están vinculados de forma directa con la Agenda 2030 y los Objetivos de Desarrollo Sostenible (ODS, Naciones Unidas, 2018) y, más específicamente, al ODS 9 sobre Industria, Innovación e Infraestructura debido al fomento de la innovación y la tecnología en la educación. Gracias a las evidencias científicas, el uso de Chatbots educativos permite mejorar el acceso a una educación más personalizada y efectiva, dando respuesta al ODS 4 sobre una Educación de Calidad. Un ejemplo de ello es el estudio de Robayo-Pinzon et al. (2023), quienes subrayan su potencial para co-crear valor en la educación superior, ya que estas tecnologías pueden mejorar el acceso a la educación, personalizar la experiencia de aprendizaje, proporcionar retroalimentación instantánea y ofrecer materiales educativos adaptados, promoviendo así una educación inclusiva, equitativa y de calidad.

Asimismo, se pueden equiparar las oportunidades educativas, así como mejorar las experiencias en cuanto al bienestar, al posibilitar la interacción personalizada entre el alumnado y los sistemas educativos automatizados, el desarrollo e implementación de los Chatbots puede ayudar a la Reducción de Desigualdades (ODS 10) y a la Salud y al Bienestar (ODS 3). En este sentido, los Chatbots ofrecen una serie

de beneficios educativos como su flexibilidad, accesibilidad y disponibilidad, aunque resulta imprescindible garantizar una implementación equitativa, responsable y ética (Pack y Maloney, 2023; Tseng y Warschauer, 2023; Yeo, 2023), así como una mejora en la comprensión de las emociones humanas (Okonkwo y Ade-Ibijola, 2021). Finalmente, el impacto positivo de los Chatbots en la educación superior puede mejorar las habilidades y competencias del alumnado, preparándolos para el mercado laboral (Celik et al., 2024) y, por tanto, contribuyendo al ODS 8 sobre Trabajo Decente y Crecimiento Económico.

A pesar del creciente interés en esta área, persisten vacíos en la literatura académica. Hasta la fecha, no se ha realizado una revisión exhaustiva que identifique las principales tendencias temáticas en el uso de Chatbots educativos, sus fuentes de co-citación o su relación con la Agenda 2030 y su conexión con los ODS. Además, existe una escasez de estudios que aporten evidencia rigurosa sobre su efectividad y desafíos específicos en el contexto universitario. Por lo tanto, este estudio busca abordar estas lagunas mediante un análisis detallado de las interacciones de los Chatbots con el alumnado y una evaluación crítica de la literatura existente. Este enfoque permitirá desarrollar un marco comprensivo y basado en evidencias que facilite la implementación y evaluación efectiva de estas tecnologías en la educación superior. Asimismo, se espera que este estudio sirva como base para futuras investigaciones y para la formulación de políticas educativas que fomenten la integración sostenible de la IA en el ámbito universitario.

Para ello, se plantean los siguientes objetivos: 1) Analizar la evolución cronológica de la productividad científica de la IA y los Chatbots en la educación superior. 2) Identificar la producción científica tomando como referencia el país de origen de las publicaciones. 3) Analizar la producción científica sobre la IA y Chatbots en la educación superior y su conexión con los ODS. 4) Identificar las principales fuentes donde se publica sobre la IA y Chatbots en la educación superior teniendo en cuenta el número de citas recibidas y el número de artículos científicos en cada una de ellas. 5) Analizar las asociaciones de co-ocurrencia entre palabras clave relevantes en los artículos académicos sobre IA y Chatbots en el contexto de la educación superior.

METODOLOGÍA

Protocolo y registro

Para dar respuesta a los objetivos de investigación formulados, en este estudio se lleva a cabo una revisión sistemática siguiendo las pautas establecidas en la Declaración PRISMA, en su versión más actualizada y enfocada al ámbito educativo (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*; Sánchez-Serrano et al., 2022). Este enfoque se considera un protocolo para realizar una investigación detallada, sistemática y rigurosa sobre la producción científica (Uman, 2011).

Criterios de elegibilidad

En la Tabla 1, se muestran los criterios de selección utilizados para la identificación de los estudios que formarán parte de la revisión sistemática.

Tabla 1
Criterios de inclusión y exclusión

Criterios de inclusión	Criterios de exclusión
<i>Filtros iniciales</i>	
Artículos publicados con texto completo.	Artículos que no cuentan con el texto completo.
Artículos de investigación publicados en revistas científicas.	Trabajos que no son publicaciones científicas: tesis doctorales, libros, actas de congresos...
<i>Otros filtros</i>	
Artículos realizados en contexto universitario.	Artículos no realizados en contexto universitario.
<i>Fecha de búsqueda: diciembre 2023</i>	

Fuente: elaboración propia.

Estrategia de búsqueda

A continuación, se aplicó la siguiente frase de búsqueda: (AI OR Artificial Intelligence) AND (chatbot) AND (“Higher education” OR university) en la base de datos Web of Science (WoS) para alinear las temáticas abordadas en los artículos con los ODS. Para la búsqueda se ha utilizado el filtro “topic” y todos los registros de la Core Collection para asegurar que se contemplan todas las áreas en el estudio.

El estudio se centra en la educación superior debido a su relevancia en el contexto de la sostenibilidad y la innovación tecnológica. Los términos utilizados en la búsqueda, como “Higher education” y “university”, están alineados con el objetivo principal de analizar el uso de inteligencia artificial y Chatbots en universidades e instituciones de educación superior. Otros niveles educativos fueron excluidos de esta revisión, ya que no constituyen el enfoque principal del análisis, que busca evaluar específicamente la implementación de estas tecnologías en el ámbito universitario. Por otro lado, dado que es un área de investigación en continuo desarrollo, se ha optado por incluir estudios sin restricciones temporales para garantizar una revisión lo más amplia y representativa posible, permitiendo identificar tanto investigaciones pioneras como tendencias.

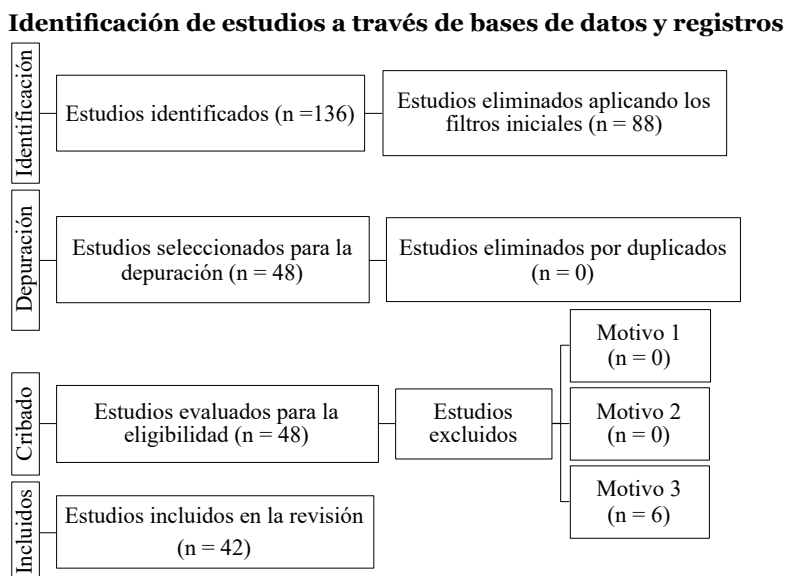
En la fase de identificación, se recuperaron 136 estudios a partir de la ecuación de búsqueda establecida. Posteriormente, en la fase de depuración, se aplicaron filtros

iniciales de inclusión y exclusión, eliminando 88 estudios que no cumplieran con los criterios establecidos, como no pertenecer al ámbito de la educación superior o no estar relacionados con Chatbots e IA. Los 48 estudios restantes fueron seleccionados para su evaluación y, en esta etapa, no se identificaron estudios duplicados.

En la fase de cribado, la selección de registros se realizó utilizando el software Covidence, especialmente diseñado para llevar a cabo análisis de revisiones sistemáticas entre varios investigadores con la finalidad de evitar sesgos en la selección (Kellermeyer et al., 2018). La evaluación de los estudios fue realizada por dos revisores de forma independiente, quienes analizaron los títulos, resúmenes y texto completo de los artículos siguiendo los criterios de inclusión y exclusión previamente establecidos. Los estudios se clasificaron en las siguientes categorías (incluir, excluir o indeterminado) según los criterios de inclusión predefinidos, y cualquier discrepancia se resolvió mediante discusión. Este enfoque garantizó un análisis sistemático y riguroso, minimizando errores y asegurando la transparencia en el proceso de selección. Así, de los 48 estudios restantes, 6 fueron excluidos por no estar directamente relacionados con el ámbito universitario (motivo 3). Finalmente, en la fase de inclusión, 42 estudios fueron seleccionados para la revisión sistemática y utilizados en el análisis bibliométrico.

En la Figura 1 se pueden observar las diferentes fases llevadas a cabo.

Figura 1
Diagrama de flujo



Fuente: elaboración propia adaptado de Page et al. (2021).

Análisis de datos

Por un lado, se lleva a cabo un análisis descriptivo y comparativo de la evolución temporal de la productividad científica a lo largo de los años, así como de la producción científica en función de la ubicación geográfica de cada publicación. Esta fase permite realizar un análisis comparativo entre países para comprender su contribución relativa y la distribución global de la investigación en educación superior relacionada con la IA y los Chatbots. A continuación, mediante un análisis de frecuencias, se muestra la producción científica en relación con los ODS. Este análisis involucra la identificación de las temáticas abordadas en los artículos según la clasificación llevada a cabo por WoS que permite la alineación con los ODS.

Además, se lleva a cabo un análisis detallado para identificar las principales fuentes de publicación en este campo. Esta evaluación se basa en el número de citas recibidas y el número de artículos científicos publicados en revistas específicas, lo que permite determinar las fuentes más influyentes y relevantes en la divulgación de la investigación sobre IA y Chatbots en educación superior. Para finalizar, se ha llevado a cabo un análisis bibliométrico mediante el software VOSviewer (Van-Eck y Waltman, 2013) aplicando el método “Association Strength”, el cual no permite la fusión automática de datos provenientes de múltiples bases de datos. La combinación de información de diferentes fuentes requiere un proceso manual de normalización y eliminación de duplicados, lo que podría introducir sesgos e inconsistencias en los resultados. Por ello, se optó por trabajar con una única base de datos que garantizara homogeneidad y rigor en el análisis de redes de co-ocurrencia de palabras clave.

Este análisis permite no solo identificar las palabras clave más relevantes, sino también entender cómo estas palabras se relacionan entre sí en el contexto de la IA y los Chatbots en la educación superior. Adicionalmente, se ha generado un mapa de red de co-ocurrencia para visualizar gráficamente las relaciones entre las palabras clave del estudio, así como la interconexión y relevancia de los términos clave en el contexto de la investigación.

RESULTADOS

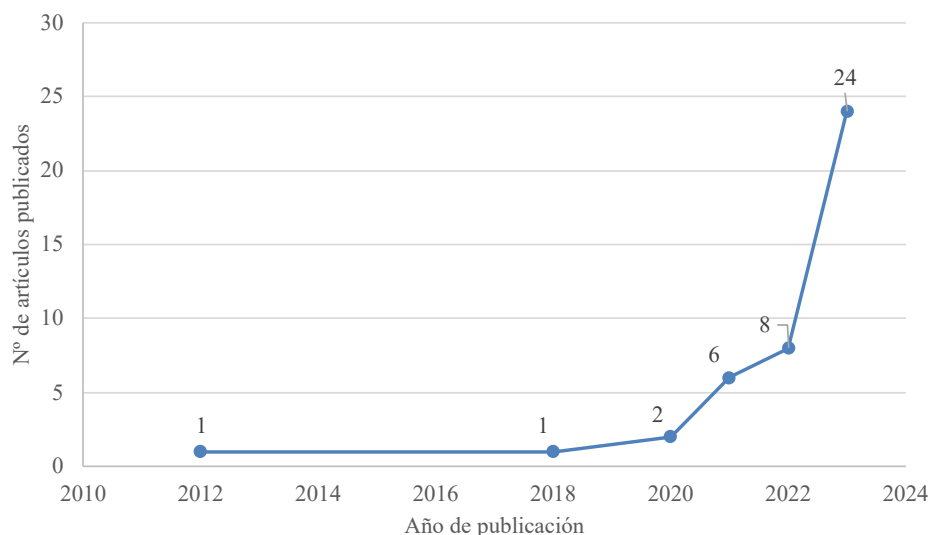
En este apartado, se muestran los resultados atendiendo a los objetivos planteados.

Evolución cronológica de la producción científica

En relación con el primer objetivo planteado, analizar la cronología de la productividad científica en relación con la IA y los Chatbots en la educación superior, se lleva a cabo un análisis descriptivo para examinar el crecimiento y la evolución del número de artículos publicados y su distribución a lo largo de los años.

La Figura 2 presenta visualmente la evolución cronológica observada a lo largo de los años, destacando una evolución notable en la productividad a lo largo de los años.

Figura 2
Número de publicaciones por año



Fuente: elaboración propia.

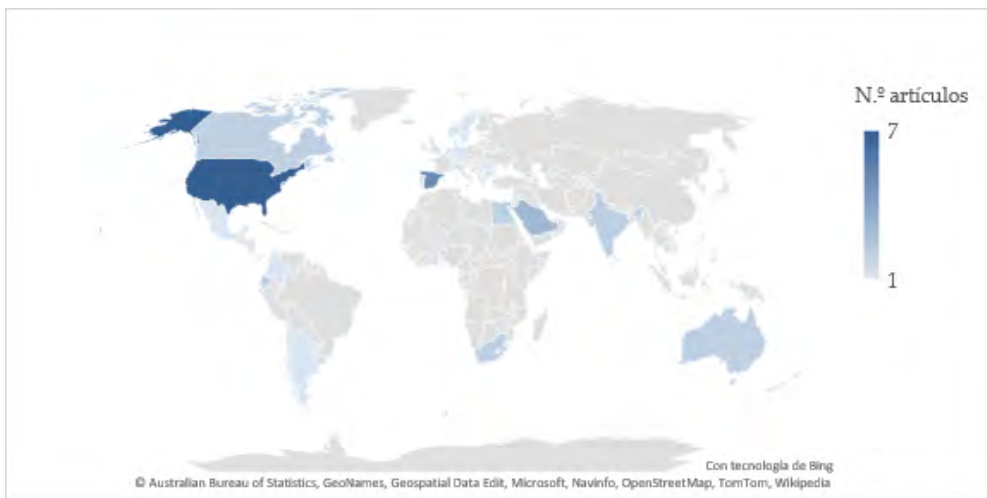
Específicamente, destaca un notorio aumento en la producción a partir del año 2020, seguido por un considerable incremento en 2021 y una aceleración aún más pronunciada en 2022. Sin embargo, el año 2023 se revela como un punto culminante en términos de productividad científica, representando un hito significativo en el número de estudios publicados sobre IA y Chatbots en la educación superior. Con un total de 24 artículos durante este periodo, este año experimenta un crecimiento notable en comparación con los años anteriores, representando más de la mitad (57,14 %) del conjunto total de artículos analizados.

Productividad científica atendiendo al país de publicación

El análisis de la distribución geográfica de la productividad científica en el ámbito de la IA y los Chatbots en educación superior, muestra una distribución variada pero no uniforme de la producción científica a nivel internacional con una clara concentración en ciertos países, tal y como se observa en la Figura 3.

Figura 3

Número de publicaciones por año



Fuente: elaboración propia.

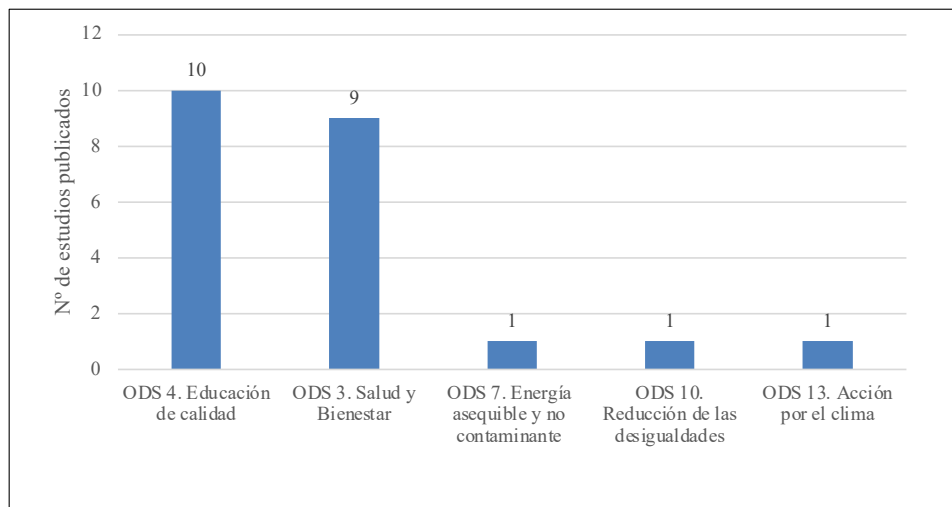
Destaca Estados Unidos quien lidera en términos de número de artículos, con 7 contribuciones de investigaciones provenientes de instituciones o autores afiliados a entidades estadounidenses en esta área de estudio. Por otro lado, los países como España y China muestran una presencia considerable con 5 y 4 artículos respectivamente. Sin embargo, se observa que varios países (Argentina, Chile, Colombia, Inglaterra, Alemania, Ghana, Israel, Japón, Malasia, México, Nueva Zelanda, Noruega, Rumania, Singapur, Corea del Sur, Suecia, Suiza, Taiwán, Vietnam) tienen una contribución de un solo artículo, lo que indica una participación más limitada pero aún existente en esta área de investigación.

Producción científica en función de los Objetivos de Desarrollo Sostenible (ODS)

En el siguiente apartado se analiza la producción científica sobre la IA y Chatbots en la educación superior y su conexión con los ODS propuestos por las Naciones Unidas. El objetivo de ese análisis es el de comprender de qué manera la integración de la IA y los Chatbots en el ámbito de la educación superior puede impulsar el avance de los ODS.

Para ello, de los 42 artículos científicos, en WoS se han identificado 21 artículos con el registro del ODS al que está asociado. En la Figura 4 se muestra un resumen del total de artículos asociados a cada ODS.

Figura 4
Número de publicaciones por año



Fuente: elaboración propia.

Los resultados muestran que el mayor número de artículos, con 10 artículos asociados, se encuentra alineado con el ODS 4 (Educación de Calidad), señalando una atención significativa hacia la mejora de la calidad educativa mediante la investigación en tecnologías educativas, como la IA y los Chatbots.

Por otro lado, se identifican un número más reducido de registros relacionados con los siguientes ODS. Un artículo está asociado al ODS 3 (Salud y Bienestar) lo que refleja un interés y dedicación en la investigación científica relacionada con la implementación de la IA y los Chatbots en la promoción de sistemas de apoyo para la salud, la medicina preventiva y el bienestar general dentro de entornos educativos. Otro artículo se asocia al ODS 7 (Energía asequible y no contaminante) mostrando cómo el uso de la IA y los Chatbots en la educación superior puede contribuir a la conciencia y la educación sobre la importancia de la energía sostenible y no contaminante. Con respecto al ODS 10 (Reducción de las desigualdades) hay un artículo alineado con la reducción de las brechas educativas existentes. Por último, un artículo asociado al ODS 13 (Acción para el clima) que muestra el interés ante la conciencia y la educación sobre el cambio climático.

Análisis de las fuentes y citas

Para conocer las fuentes donde más se publica sobre IA y Chatbot en el ámbito de la educación superior, se ha realizado un análisis de frecuencias atendiendo a las citas recibidas en las 42 referencias publicadas en un total de 35 fuentes.

En la Tabla 2 se muestran las fuentes que han recibido al menos 1 cita en relación con el número de artículos asociados a cada una de ellas.

Los resultados muestran la distribución de las citas y la cantidad de artículos relacionados con la IA y la tecnología aplicada a la educación superior y a la salud mental en varias revistas académicas. JMIR Mental Health lidera con una alta cantidad de citas, exactamente 161 citas a pesar de tener solo un artículo asociado (Fulmer et al., 2018). Este artículo, centrado en el uso de la IA para mitigar síntomas de ansiedad y depresión en estudiantes universitarios, se ha convertido en una contribución significativa en este campo.

Revistas como *Sustainability* e *International Journal of Educational Technology in Higher Education* también destacan, acumulando entre 50 y 54 citas cada una a través de tres artículos asociados. Sus contribuciones, en el caso de *Sustainability* (Park y Kim, 2023; Rahim et al., 2022; Villegas-Ch et al., 2020) y en el caso de *International Journal of Educational Technology in Higher Education* (Barrett y Pack, 2023; Essel et al., 2022; Vázquez-Cano et al., 2021) abordan temas relevantes en la aplicación de la IA en educación superior.

Tabla 2

Fuentes, citas y número de artículos

Fuentes	n.º de citas	n.º de artículos
JMIR Mental Health	161	1
Sustainability	54	3
International Journal of Educational Technology in Higher Education	50	3
Journal of University Teaching and Learning Practice	28	1
Library Hi Tech	24	1
Internet Interventions-The Application of Information Technology in Mental and Behavioural Health	24	1
Jmir Formative Research	20	1
Open Praxis	18	1
International Journal of Emerging Technologies in Learning	14	1
European Journal of Engineering Education	11	1
Future Internet	9	1

Fuentes	n.º de citas	n.º de artículos
Journal of Medical Internet Research	8	2
Medical Education Online	7	1
Library Management	5	1
Journal of King Saud University-Computer and Information Sciences	5	1
Processes	4	1
Jmir Research Protocols	4	1
ACM Transactions on Interactive Intelligent Systems	4	1
Wmu Journal of Maritime Affairs	3	1
Frontiers in Psychology	2	2
Pixel-Bit. Revista de Medios y Educacion	2	1
Education Sciences	2	2
Languages	2	1
Market-Trziste	2	1
Telecommunications Policy	1	1
Assessment & Evaluation in Higher Education	1	1
Education and Information Technologies	1	1
JMIR Medical Education	1	1

Fuente: elaboración propia.

Otras fuentes, como *Journal of University Teaching and Learning Practice*, *Library Hi Tech*, *Internet Interventions*, *JMIR Formative Research*, entre otros, también han recibido un número considerable de citas, lo que indica la relevancia y el impacto de sus investigaciones en este ámbito.

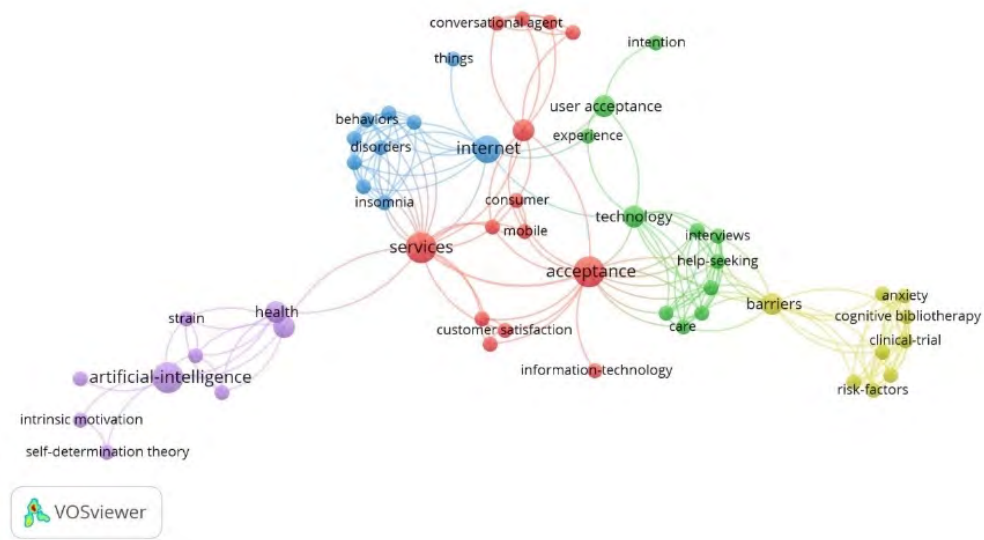
Análisis de co-ocurrencia de palabras clave

A continuación, se dará respuesta al segundo objetivo planteado: Identificar y analizar las asociaciones de co-ocurrencia entre palabras clave relevantes en los artículos académicos sobre IA y Chatbots en el contexto de la educación superior.

Para realizar este análisis, se examinaron 42 artículos, considerando todas las palabras clave, incluidas las palabras clave de autor y las palabras clave PLUS. Esto permitió un estudio detallado de las relaciones de co-ocurrencia entre términos clave, obteniendo un total de 82 palabras analizadas.

En la Figura 5, los nodos más grandes indican las palabras clave con mayor frecuencia de citación en los artículos analizados.

Figura 5
Red de palabras claves (VOSviewer)



Fuente: elaboración propia.

A partir de este mapa, se identificaron cuatro clústeres, cada uno señalado con un color distinto (rojo, verde, azul y amarillo), destacando así las interrelaciones en el ámbito de la investigación. A continuación, en la Tabla 3 se presenta el detalle de la red, mostrando las categorías temáticas derivadas de las palabras clave.

Tabla 3
Categorías temáticas de palabras clave según la co-ocurrencia

Categoría	Descripción del clúster	Palabras clave
Tecnología en la educación superior (clúster color rojo)	Este clúster se centra en las percepciones, actitudes y comportamientos relacionados con la adopción y aceptación de tecnologías de IA y chatbot en el ámbito de la educación superior.	acceptance; services; model; consumer; conversational agent; customer satisfaction; design; display; information-technology; internet banking; mobile; quality; science; utaut2.

Categoría	Descripción del clúster	Palabras clave
Tecnologías de apoyo en educación superior (clúster color verde)	Este clúster trata sobre la disposición y aceptación de los estudiantes hacia las tecnologías de apoyo, incluyendo IA y Chatbots, resaltando la importancia de su aceptación para su eficaz implementación en entornos educativos.	technology; user acceptance; care; conversational agents; experience; help-seeking; information; intention; interviews; self-disclosure; virtual agent.
Chatbot en salud y bienestar (clúster color azul)	Este clúster está centrado en la aplicación y el uso de IA y chatbot en aplicaciones relacionadas con la salud, tanto física como mental, así como en intervenciones y prevención en el ámbito universitario.	internet; behaviors; community-health; disorders; framework; insomnia; older-adults awareness; randomized controlled-trial; things; web
Desafíos en la integración educativa con IA y Chatbots (clúster color amarillo)	Este clúster analiza los desafíos psicológicos y las barreras asociadas con la integración de tecnología, específicamente IA y chatbots, en la educación superior, incluyendo aspectos como ansiedad, factores de riesgo y problemas sistémicos.	barriers; anxiety; clinical-trial; cognitive bibliotherapy; psychometric properties; risk-factors; symptoms; system; working alliance.
Impacto de la IA en la salud y motivación de estudiantes (cluster color morado)	Este clúster está enfocado en comprender cómo la IA afecta a la salud mental, la motivación y la dinámica universitaria, destacando su influencia en la motivación intrínseca de los estudiantes y su bienestar general.	artificial-intelligence; health; stress; education; higher-education; intrinsic motivation; prevalence; self-determination theory; strain

Fuente: elaboración propia.

DISCUSIÓN Y CONCLUSIONES

El análisis exhaustivo realizado en el marco de este estudio se ha enfocado en la exploración detallada de la función desempeñada por los Chatbots basados en IA dentro del ámbito de la educación superior. En las siguientes líneas se muestran las conclusiones más relevantes del estudio llevado a cabo.

En primer lugar, los datos revelan un incremento significativo en la publicación de artículos sobre IA y Chatbots en educación superior en los últimos años, particularmente en 2022 y 2023, lo que indica una creciente tendencia investigativa

en el tema. Esta dinámica es consistente con lo reportado por Redondo-Duarte et al. (2024), quienes identificaron una transición hacia áreas más especializadas dentro de este campo, tales como la personalización del aprendizaje y la tutoría inteligente.

Además, se aprecia un uso de Chatbots en diferentes ámbitos de la educación superior, como, por ejemplo, Noble et al. (2022), quienes presentaron un Chatbot orientado a la navegación de recursos de salud mental para trabajadores de la salud y sus familias, evidenciando el potencial de la IA para el bienestar emocional. Asimismo, Sharma et al. (2023) desarrollaron un Chatbot especializado en la formación en Regulaciones para la Prevención de Colisiones para la industria marítima, lo que subraya la capacidad de estos sistemas para adaptarse a sectores específicos y mejorar los procesos de enseñanza y aprendizaje. En el ámbito pedagógico, Shim et al. (2023) evaluaron la efectividad de talleres de Chatbots como herramientas de enseñanza experiencial, mostrando cómo pueden fortalecer el compromiso de los estudiantes con su aprendizaje. De manera similar, Tamayo et al. (2020) diseñaron un Chatbot como asistente de aprendizaje a distancia, destacando su potencial en la educación en línea y en la mejora de la accesibilidad a los contenidos educativos.

Otro aspecto relevante es el uso de Chatbots con enfoques más avanzados, como el desarrollado por Trappey et al. (2022), que incorpora análisis de diálogos sentimentales para ofrecer consejería empática a los usuarios. Este tipo de aplicaciones pone de manifiesto la capacidad de los Chatbots para adaptarse a interacciones más complejas y responder a necesidades emocionales y psicológicas en el ámbito educativo. Además, Villegas-Ch et al. (2021) implementaron un asistente virtual con IA para la gestión académica en una universidad, mostrando la utilidad de estas herramientas para mejorar la eficiencia administrativa y optimizar los procesos internos de las instituciones educativas.

En segundo lugar, el análisis de la producción científica según la distribución geográfica evidencia la participación de diversos países y regiones. Estados Unidos es el líder en este campo con una fuerte contribución y presencia de instituciones y autores estadounidenses en el desarrollo y estudio de la IA y los Chatbots en la educación superior. La prominencia de estudios estadounidenses en la investigación educativa respalda su influencia en el desarrollo y la adopción de soluciones basadas en IA en el entorno educativo superior. No obstante, como mencionan Kuhail et al. (2023), no se debe pasar por alto la creciente participación de países europeos como España o asiáticos como China, lo que indica un interés y una implicación cada vez mayores en esta área.

Esta diversidad geográfica subraya la importancia de la colaboración internacional y la multiplicidad de perspectivas y enfoques, enriqueciendo así el desarrollo de las aplicaciones de IA y Chatbots en la educación superior a nivel global.

En tercer lugar, el análisis de los estudios relacionados con la IA y los Chatbots en la educación superior, en el marco de los ODS, destaca la disparidad en la atención otorgada a diferentes ODS. El ODS 4, centrado en la Educación de Calidad, emerge como el foco principal de la investigación científica, indicando un interés

considerable en comprender y aprovechar el potencial de la IA y los Chatbots para la mejora de la calidad del sistema educativo. Este enfoque es coherente con la importancia estratégica de la educación en el logro de un desarrollo sostenible (Naciones Unidas, 2018). No obstante, Howell y Potgieter (2023) advierten sobre la necesidad de abordar el posible mal uso de la IA y los Chatbots para garantizar la integridad académica y el cumplimiento de los principios del ODS 4.

El ODS 3 (Salud y Bienestar) también se ve beneficiado por la integración de Chatbots en plataformas educativas, ya que pueden proporcionar información sobre bienestar emocional, prevención de enfermedades y promoción de hábitos saludables entre estudiantes. Además, pueden facilitar entornos educativos saludables mediante recursos sobre salud mental, reducción del estrés y orientación médica preventiva (Koulouri et al., 2022).

En cuanto al ODS 7 (Energía Asequible y No Contaminante), los Chatbots ofrecen un papel clave en la concienciación sobre el uso responsable de la energía, ofreciendo información educativa sobre energías renovables y sostenibilidad en los entornos académicos. Mediante estas herramientas, se pueden difundir prácticas y conocimientos sobre eficiencia energética y reducción de la huella de carbono (Naciones Unidas, 2020).

El impacto en el ODS 10 (Reducción de las Desigualdades) con el uso de Chatbots en plataformas de educación en línea, facilitan la superación de barreras geográficas y socioeconómicas. Esto permite un acceso más equitativo a la educación superior, especialmente en comunidades con recursos limitados, ayudando a reducir las brechas en la educación de calidad (UNESCO, 2021).

Finalmente, el ODS 13 (Acción por el Clima) se vincula con el uso de Chatbots en la sensibilización sobre el cambio climático y la promoción de comportamientos sostenibles. El estudio de Menkhoff y Gan (2023) muestra que los chatbots pueden involucrar a los estudiantes de manera interactiva, proporcionando información relevante y personalizada sobre temas ambientales y climáticos. Esto no solo aumenta la conciencia sobre la importancia de la acción climática, sino que también fomenta comportamientos sostenibles entre los estudiantes.

Es importante destacar la ausencia total de artículos relacionados con varios ODS (1, 2, 5, 6, 8, 9, 11, 12, 14, 15, 16 y 17). Esta carencia de investigación específica subraya la falta de exploración sobre la aplicación de la IA y los chatbots en áreas fundamentales para la Educación de Desarrollo Sostenible (EDS).

En cuarto lugar, las conclusiones extraídas del análisis de fuentes y citas muestran la importancia y el impacto diferencial de diversas fuentes en el campo de la IA aplicada a la educación superior, destacando: *JMIR Mental Health, Sustainability e International Journal of Educational Technology in Higher Education*. Estos resultados evidencian una diversidad significativa en las publicaciones provenientes de diversas fuentes interesadas en impulsar y fomentar la investigación en esta área en constante crecimiento. Estos hallazgos respaldan la importancia de seguir fomentando y respaldando una investigación interdisciplinaria y colaborativa para

garantizar el progreso continuo en el uso efectivo de la IA en la educación superior (Jiménez-García et al., 2024; Kuhail et al., 2023; Redondo-Duarte et al., 2024).

En quinto lugar, el análisis de co-ocurrencia de palabras clave en los clústeres identificados revela la relevancia y el potencial de la IA y los Chatbots en la educación superior, pero enfatizan la necesidad de abordar desafíos y aspectos psicológicos para una implementación adecuada y un impacto positivo en el entorno educativo. En esta línea, los Clústeres 1 “Tecnología en la Educación Superior” y 2 “Tecnologías de Apoyo en Educación Superior” resaltan la importancia de comprender las percepciones, actitudes y disposición de los estudiantes (Al-Abdullatif, 2023) y docentes (Farazouli et al., 2023) hacia la adopción de tecnologías de IA y Chatbots en el ámbito universitario. Asimismo, el Clúster 3 “Chatbot en Salud y Bienestar” destaca un interés específico en la aplicación de Chatbots y tecnologías de IA en la salud (Guckenberger et al., 2023; Hirose et al., 2023; Huang et al., 2023), tanto física como mental (Liu et al., 2022), así como en intervenciones y prevención en el contexto universitario (Jackson-Triche et al., 2023); esto indica un énfasis en la aplicación de estas tecnologías para potenciar el bienestar integral de los estudiantes (Liu et al., 2022).

Por último, los Clústeres 4 “Desafíos en la integración educativa con IA y Chatbots” y 5 “Impacto de la IA en la salud y motivación de estudiantes” abordan los desafíos psicológicos, (Crawford et al., 2023; Hasanein y Sobaih, 2023) y el impacto de la IA en la salud y motivación de los estudiantes (Al-Abdullatif et al., 2023). Estos clústeres resaltan la importancia de superar desafíos como la ansiedad, los factores de riesgo y los problemas sistémicos asociados con la integración de IA y Chatbots en la educación superior. Además, destacan la importancia de entender el impacto de estas tecnologías en la salud mental y la motivación de los estudiantes en el entorno universitario.

Atendiendo a estos resultados, se concluye que los Chatbots basados en IA tienen un impacto significativo en la educación superior, con aplicaciones concretas en asistencia académica, salud mental y automatización de procesos educativos.

Sin embargo, persisten retos importantes en su diseño e implementación. En cuanto a las barreras y desafíos, la implementación de Chatbots en educación superior enfrenta limitaciones tecnológicas, éticas y psicológicas. Desde una perspectiva tecnológica, la falta de interoperabilidad entre sistemas educativos y la necesidad de grandes volúmenes de datos para el entrenamiento de modelos de IA siguen siendo obstáculos importantes (Kuhail et al., 2023). Singh y Beniwal (2022) muestran algunas limitaciones de los Chatbots en cuanto a la coherencia y la naturalidad de las interacciones.

En el ámbito ético, preocupa la privacidad de los datos y los fraudes (Lozano y Fontao, 2023), así como la integridad académica y la autoría (Nam y Bai, 2023). El estudio de Nikolic et al. (2023) analiza el impacto de ChatGPT en la evaluación. En concreto los resultados muestran la necesidad de modificar las prácticas de evaluación actuales para mantener la integridad académica. En términos

psicológicos, la aceptación de los Chatbots por parte de estudiantes y docentes es variable, influenciada por la confianza en la tecnología y la percepción de su efectividad (Hasanein y Sobaih, 2023).

Para finalizar, cabe destacar algunas limitaciones que presenta este estudio y que se centran en el uso exclusivo de la base de datos Web of Science. Si bien esto garantizó un marco analítico estructurado, alineado con los ODS, pero la exclusión de otras bases de datos puede haber limitado la cobertura de la literatura analizada. Además, el uso de VOSviewer restringió la posibilidad de combinar datos de múltiples fuentes, ya que esta herramienta no permite la fusión automática de registros. La integración manual podría haber introducido sesgos, por lo que se optó por trabajar con una única base de datos para asegurar la coherencia del análisis. Si bien estas limitaciones no afectan la validez de los resultados obtenidos, futuras investigaciones podrían explorar estrategias para integrar múltiples fuentes de datos, lo que permitiría ampliar el alcance del análisis bibliométrico en este campo.

En este contexto, se sugiere orientar futuras investigaciones hacia el impacto de estas tecnologías en otras etapas educativas, con el fin de ampliar el alcance de los análisis. Asimismo, se propone profundizar en la brecha identificada en el estudio, especialmente en el ámbito de la Educación para el Desarrollo Sostenible, destacando la necesidad de expandir el enfoque científico en áreas menos exploradas de la aplicación de la inteligencia artificial y los Chatbots.

La exploración de estas áreas estratégicas podría ofrecer a la comunidad educativa y científica soluciones tecnológicas más integrales, contribuyendo de manera significativa a la consecución de los ODS y al bienestar global. Con respecto a la implementación de los Chatbots, la IA también plantea desafíos, como la necesidad de garantizar interacciones efectivas y la preocupación por la privacidad de los usuarios. Además, se sugiere investigar estrategias para mejorar la aceptación y confianza en estas tecnologías, así como evaluar su impacto en el rendimiento académico y bienestar estudiantil.

NOTAS

1. Este artículo ha sido realizado en el marco del Proyecto de Investigación XSAN002310 “Chatbot para el acompañamiento y evaluación formativa del alumnado. Análisis, Diseño, Implementación” financiado por la Universidad Europea de Madrid y el Banco Santander.

REFERENCIAS

- Al-Abdullatif, A. M. (2023). Modeling students' perceptions of chatbots in learning: Integrating technology acceptance with the value-based adoption model. *Education Sciences*, 13(11), 1151. <https://doi.org/10.3390/educsci13111151>
- Al-Abdullatif, A. M., Al-Dokhny, A. A. y Drwish, A. M. (2023). Implementing the Bashayer chatbot in Saudi higher education: Measuring the influence on students' motivation and learning strategies. *Frontiers in Psychology*,

- 14, 1129070. <https://doi.org/10.3389/fpsyg.2023.1129070>
- Alkhoodi, A., Kuhail, M. A. y Alkhoodi, A. (2020). Unibud: A virtual academic adviser. En *2020 12th Annual Undergraduate Research Conference on Applied Computing (URC)* (pp. 1-4). IEEE. <https://doi.org/10.1109/URC49805.2020.9099191>
- Allison, D. (2012). Chatbots in the library: Is it time? *Library Hi Tech*, 30(1), 95-107. <https://doi.org/10.1108/07378831211213238>
- Alt, M. A. y Ibolya, V. (2021). Identifying relevant segments of potential banking chatbot users based on technology adoption behavior. *Market-Trziste*, 33(2), 165-183. <https://doi.org/10.22598/mt/2021.33.2.165>
- Anghelescu, P. y Nicolaescu, S. V. (2018). Chatbot application using search engines and teaching methods. En *2018 10th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)* (pp. 1-6). IEEE. <https://doi.org/10.1109/ECAI.2018.8678948>
- Barrett, A. y Pack, A. (2023). Not quite eye to AI: Student and teacher perspectives on the use of generative artificial intelligence in the writing process. *International Journal of Educational Technology in Higher Education*, 20, Artículo 59. <https://doi.org/10.1186/s41239-023-00427-0>
- Celik, I., Gedrimiene, E., Siklander, S. y Muukkonen, H. (2024). Las posibilidades de las herramientas basadas en inteligencia artificial para apoyar las habilidades del siglo XXI: Una revisión sistemática de la investigación empírica en la educación superior. *Australasian Journal of Educational Technology*, 40(3), 19-38. <https://doi.org/10.14742/ajet.9069>
- Crawford, J., Cowling, M. y Allen, K. A. (2023). Leadership is needed for ethical ChatGPT: Character, assessment, and learning using artificial intelligence (AI). *Journal of University Teaching and Learning Practice*, 20(3), Artículo 2. <https://doi.org/10.53761/1.20.3.02>
- Dinh, H. y Tran, T. K. (2023). EduChat: An AI-based chatbot for university-related information using a hybrid approach. *Applied Sciences-Basel*, 13(22), 12446. <https://doi.org/10.3390/app132212446>
- Dube, T. V. y Jacobs, L. (2023). Academic library services extension during the COVID-19 pandemic: Considerations in higher education institutions in the Gauteng Province, South Africa. *Library Management*, 44(1/2), 17-39. <https://doi.org/10.1108/LM-04-2022-0039>
- Essel, H. B., Vlachopoulos, D., Tachie-Menson, A., Johnson, E. E. y Baah, P. K. (2022). The impact of a virtual teaching assistant (chatbot) on students' learning in Ghanaian higher education. *International Journal of Educational Technology in Higher Education*, 19, Artículo 55. <https://doi.org/10.1186/s41239-022-00362-6>
- Farazouli, A., Cerratto-Pargman, T., Bolander-Laksov, K. y McGrath, C. (2023). Hello GPT! Goodbye home examination? An exploratory study of AI chatbots impact on university teachers' assessment practices. *Assessment & Evaluation in Higher Education*. <https://doi.org/10.1080/02602938.2023.2241676>
- Fulmer, R., Joerin, A., Gentile, B., Lakerink, L. y Rauws, M. (2018). Using psychological artificial intelligence (Tess) to relieve symptoms of depression and anxiety: Randomized controlled trial. *JMIR Mental Health*, 5(4), Artículo e10118. <https://doi.org/10.2196/mental.9782>
- Guckenberger, M., Andratschke, N., Ahmadsei, M., Christ, S. M., Heusel, A. E., Kamal, S., Kroese, T. E., Looman, E. L., Reichl, S., Badra, E. V, von der Grun, J., Willmann, J., Tanadini-Lang, S. y Mayinger, M. (2023). Potential of ChatGPT in facilitating research in

- radiation oncology? *Radiotherapy and Oncology*, 188, 109894. <https://doi.org/10.1016/j.radonc.2023.109894>
- Hasanein, A. M. y Sobaih, A. E. (2023). Drivers and consequences of ChatGPT use in higher education: Key stakeholder perspectives. *European Journal of Investigation in Health Psychology and Education*, 13(11), 2599-2614. <https://doi.org/10.3390/ejihpe13110181>
- Hirosawa, T., Kawamura, R., Harada, Y., Mizuta, K., Tokumasu, K., Kaji, Y., Suzuki, T. y Shimizu, T. (2023). ChatGPT-generated differential diagnosis lists for complex case-derived clinical vignettes: Diagnostic accuracy evaluation. *JMIR Medical Informatics*, 11, Artículo e48808. <https://doi.org/10.2196/48808>
- Howell, B. E. y Potgieter, P. H. (2023). What do telecommunications policy academics have to fear from GPT-3? *Telecommunications Policy*, 47(7), 102576. <https://doi.org/10.1016/j.telpol.2023.102576>
- Huang, R. S. T., Lu, K. J. Q., Meaney, C., Kempainen, J., Punnett, A. y Leung, F. H. (2023). Assessment of resident and AI chatbot performance on the University of Toronto family medicine residency progress test: Comparative study. *JMIR Medical Education*, 9, Artículo e50514. <https://doi.org/10.2196/50514>
- Jackson-Triche, M., Vetat, D., Turner, E. M., Dahiya, P. y Mangurian, C. (2023). Meeting the behavioral health needs of health care workers during COVID-19 by leveraging chatbot technology: Development and usability study. *Journal of Medical Internet Research*, 25, Artículo e40635. <https://doi.org/10.2196/40635>
- Jiménez-García, E., Orenes-Martínez, N. y López-Fraile, L. A. (2024). Rueda de la Pedagogía para la Inteligencia Artificial: Adaptación de la Rueda de Carrington. *RIED-Revista Iberoamericana de Educación a Distancia*, 27(1), 87-113. <https://doi.org/10.5944/ried.27.1.37622>
- Kellermeyer, R., Heydman, L. M., Mastick, G. S. y Kidd, T. (2018). *The role of apoptotic signaling*. Preprints.org. <https://doi.org/10.20944/preprints201809.0281.v1>
- Kingchang, T., Chatwattana, P. y Wannapiroon, P. (2024). Artificial intelligence chatbot platform: AI chatbot platform for educational recommendations in higher education. *International Journal of Information and Education Technology*, 14(1), 1-7. <https://doi.org/10.18178/ijiet.2024.14.1.2021>
- Koulouri, T., Macredie, R. D. y Olakitan, D. (2022). Chatbots to support young adults' mental health: An exploratory study of acceptability. *ACM Transactions on Interactive Intelligent Systems*, 12(2), Artículo 13. <https://doi.org/10.1145/3485874>
- Kuhail, M. A., Alturki, N., Alramlawi, S. y Alhejori, K. (2023). Interacting with educational chatbots: A systematic review. *Education and Information Technologies*, 28(1), 973-1018. <https://doi.org/10.1007/s10639-022-11177-3>
- Liu, H., Peng, H. M., Song, X. Y., Xu, C. Z. y Zhang, M. (2022). Using AI chatbots to provide self-help depression interventions for university students: A randomized trial of effectiveness. *Internet Interventions-the Application of Information Technology in Mental and Behavioural Health*, 27, 100495. <https://doi.org/10.1016/j.invent.2022.100495>
- Lozano, A. y Fontao, C. B. (2023). Is the education system prepared for the irruption of artificial intelligence? A study on the perceptions of students of primary education degree from a dual perspective: Current pupils and future teachers. *Education Sciences*, 13(7), 733. <https://doi.org/10.3390/educsci13070733>
- Malik, R., Sharma, A., Trivedi, S. y Mishra, R. (2021). Adoption of chatbots for learning among university students: Role of perceived convenience and enhanced performance. *International Journal of*

- Emerging Technologies in Learning*, 16(18), 200-212. <https://doi.org/10.3991/ijet.v16i18.24315>
- Menkhoff, T. y Gan, B. (2023). Involucrar a los estudiantes a través de chatbots conversacionales y contenido digital: Una perspectiva de acción climática. *Human Interaction and Emerging Technologies (IHET)*, 70, 334-347. <https://doi.org/10.54941/ahfe1002960>
- Moldt, J. A., Festl-Wietek, T., Mamlouk, A. M., Nieselt, K., Fuhl, W. y Herrmann-Werner, A. (2023). Chatbots for future docs: Exploring medical students' attitudes and knowledge towards artificial intelligence and medical chatbots. *Medical Education Online*, 28(1), 2182659. <https://doi.org/10.1080/10872981.2023.2182659>
- Naciones Unidas. (2018). *La Agenda 2030 y los Objetivos de Desarrollo Sostenible: Una oportunidad para América Latina y el Caribe*. CEPAL.
- Naciones Unidas. (2020). *Energía asequible y sostenible en la educación*. <https://www.un.org/sustainabledevelopment/es/energy/>
- Nam, B. H. y Bai, Q. (2023). ChatGPT and its ethical implications for STEM research and higher education: A media discourse analysis. *International Journal of STEM Education*, 10(1), Artículo 61. <https://doi.org/10.1186/s40594-023-00452-5>
- Nikolic, S., Daniel, S., Haque, R., Belkina, M., Hassan, G. M., Grundy, S., Lyden, S., Neal, P. y Sandison, C. (2023). ChatGPT versus engineering education assessment: A multidisciplinary and multi-institutional benchmarking and analysis of this generative artificial intelligence tool to investigate assessment integrity. *European Journal of Engineering Education*, 48(4), 559-614. <https://doi.org/10.1080/03043797.2023.2213169>
- Noble, J. M., Zamani, A., Gharaat, M., Merrick, D., Maeda, N., Foster, A. L., Nikolaidis, I., Goud, R., Stroulia, E., Agyapong, V. I. O., Greenshaw, A. J., Lambert, S., Gallson, D., Porter, K., Turner, D. y Zaiane, O. (2022). Developing, implementing, and evaluating an artificial intelligence-guided-mental health resource navigation chatbot for health care workers and their families during and following the COVID-19 pandemic: Protocol for a cross-sectional study. *JMIR Research Protocols*, 11(7), Artículo e33717. <https://doi.org/10.2196/33717>
- Ojeda, A. D., Solano-Barliza, A. D., Alvarez, D. O. y Cárcamo, E. B. (2023). Análisis del impacto de la inteligencia artificial ChatGPT en los procesos de enseñanza y aprendizaje en la educación universitaria. *Formación universitaria*, 16(6), 61-70. <https://doi.org/10.4067/S0718-50062023000600061>
- Okonkwo, C. W. y Ade-Ibijola, A. (2021). Chatbots applications in education: A systematic review. *Computers and Education: Artificial Intelligence*, 2, 100033. <https://doi.org/10.1016/j.caeai.2021.100033>
- Oqaidi, K., Aouhassi, S. y Mansouri, K. (2024). Are chatbots the future of higher education? Unveiling their impact, challenges, and prospects. En *2024 4th International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET)* (pp. 1-6). IEEE. <https://doi.org/10.1109/IRASET60544.2024.10548820>
- Pack, A. y Maloney, J. (2023). Using generative artificial intelligence for language education research: Insights from using OpenAI's ChatGPT. *TESOL Quarterly*. <https://doi.org/10.1002/tesq.3253>
- 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., ... Moher, D. (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>
- Park, D. Y. y Kim, H. (2023). Determinants of intentions to use digital mental healthcare content among university students, faculty, and staff: Motivation, perceived usefulness, perceived ease of use, and parasocial interaction with AI chatbot. *Sustainability*, 15(1), 872. <https://doi.org/10.3390/su15010872>
- Pérez, J. Q., Daradoumis, T. y Puig, J. M. M. (2020). Rediscovering the use of chatbots in education: A systematic literature review. *Computer Applications in Engineering Education*, 28(6), 1549-1565. <https://doi.org/10.1002/cae.22326>
- Rahim, N. I. M., 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>
- Redondo-Duarte, S., Martínez-Requejo, S., Jiménez-García, E. y Ruiz-Lázaro, J. (2024). The potential of educational chatbots for the support and formative assessment of students. En M. Ibrahim, M. Aydoğmuş y Y. Tükel (Eds.), *New trends and promising directions in modern education*, (pp. 113-148). Palet Yayınları.
- Robayo-Pinzon, O., Rojas-Berrio, S., Rincon-Novoa, J. y Ramirez-Barrera, A. (2023). Artificial intelligence and the value co-creation process in higher education institutions. *International Journal of Human-Computer Interaction*. <https://doi.org/10.1080/10447318.2023.2259722>
- Rodríguez, J. A., Santana, M. G., Perera, M. V. A. y Pulido, J. R. (2021). Embodied conversational agents: Artificial intelligence for autonomous learning. *Pixel-Bit. Revista de Medios y Educacion*, 62, 107-144. <https://doi.org/10.12795/pixelbit.86171>
- 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>
- Sain, Z., Vasudevan, A. y Lama, A. (2024). The emerging future of AI chatbots in higher education. *Jurnal Ilmiah Didaktika: Media Ilmiah Pendidikan dan Pengajaran*, 25(1), 93-107. <https://doi.org/10.22373/jid.v25i1.25583>
- Sánchez-Serrano, S., Pedraza-Navarro, I. y Donoso-González, M. (2022). ¿Cómo hacer una revisión sistemática siguiendo el protocolo PRISMA? Usos y estrategias fundamentales para su aplicación en el ámbito educativo a través de un caso práctico. *Bordón. Revista de Pedagogía*, 74(3), 51-66. <https://doi.org/10.13042/Bordon.2022.95090>
- Sharma, A., Undheim, P. E. y Nazir, S. (2023). Design and implementation of AI chatbot for COLREGs training. *WMU Journal of Maritime Affairs*, 22(1), 107-123. <https://doi.org/10.1007/s13437-022-00284-0>
- Shim, K. J., Menkhoff, T., Teo, L. Y. Q. y Ong, C. S. Q. (2023). Assessing the effectiveness of a chatbot workshop as experiential teaching and learning tool to engage undergraduate students. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-023-11795-5>
- Singh, S. y Beniwal, H. (2022). A survey on near-human conversational agents. *Journal of King Saud University-Computer and Information Sciences*, 34(10, Part B), 8852-8866. <https://doi.org/10.1016/j.jksuci.2021.10.013>
- Tamayo, P. A., Herrero, A., Martín, J., Navarro, C. y Tránchez, J. M. (2020). Design of a chatbot as a distance learning assistant. *Open Praxis*, 12(1),

- 145-153. <https://doi.org/10.5944/openpraxis.12.1.1063>
- Tian, W., Ge, J., Zhao, Y. y Zheng, X. (2024). AI chatbots in Chinese higher education: Adoption, perception, and influence among graduate students-An integrated analysis utilizing UTAUT and ECM models. *Frontiers in Psychology*, 15, 1268549. <https://doi.org/10.3389/fpsyg.2024.1268549>
- Trappey, A. J. C., Lin, A. P. C., Hsu, K. Y. K., Trappey, C. V. y Tu, K. L. K. (2022). Development of an empathy-centric counseling chatbot system capable of sentimental dialogue analysis. *Processes*, 10(5), 930. <https://doi.org/10.3390/pr10050930>
- Tseng, W. y Warschauer, M. (2023). AI-writing tools in education: If you can't beat them, join them. *Journal of China Computer-Assisted Language Learning*, 3(2), 258-262. <https://doi.org/10.1515/jccall-2023-0008>
- Uman, L. S. (2011). Systematic reviews and meta-analyses. *Journal of the Canadian Academy of Child & Adolescent Psychiatry*, 20(1), 57-59.
- UNESCO. (2021). *Recomendación sobre la Ética de la Inteligencia Artificial*. https://unesdoc.unesco.org/ark:/48223/pf0000381137_spa
- Van-Eck, N. J. y Waltman, L. (2013). *VOSviewer manual*. Leiden University Centre for Science and Technology Studies (CWTS).
- Vázquez-Cano, E., Mengual-Andrés, S. y López-Meneses, E. (2021). Chatbot to improve learning punctuation in Spanish and to enhance open and flexible learning environments. *International Journal of Educational Technology in Higher Education*, 18, Artículo 31. <https://doi.org/10.1186/s41239-021-00269-8>
- Villegas-Ch, W., Arias-Navarrete, A. y Palacios-Pacheco, X. (2020). Proposal of an architecture for the integration of a chatbot with artificial intelligence in a smart campus for the improvement of learning. *Sustainability*, 12(4), 1500. <https://doi.org/10.3390/su12041500>
- Villegas-Ch, W., García-Ortiz, J., Mullo-Ca, K., Sánchez-Viteri, S. y Roman-Cañizares, M. (2021). Implementation of a virtual assistant for the academic management of a university with the use of artificial intelligence. *Future Internet*, 13(4), 97. <https://doi.org/10.3390/fi13040097>
- Weeks, R., Cooper, L., Sangha, P., Sedoc, J., White, S., Toledo, A., Gretz, S., Lahav, D., Martin, N., Michel, A., Lee, J. H., Slonim, N. y Bar-Zeev, N. (2022). Chatbot-delivered COVID-19 vaccine communication message preferences of young adults and public health workers in urban American communities: Qualitative study. *Journal of Medical Internet Research*, 24(7), Artículo e38418. <https://doi.org/10.2196/38418>
- Xiao, Y. Y. y Zhi, Y. Y. (2023). An exploratory study of EFL learners' use of ChatGPT for language learning tasks: Experience and perceptions. *Languages*, 8(3), 212. <https://doi.org/10.3390/languages8030212>
- Yeo, M. A. (2023). Academic integrity in the age of artificial intelligence (AI) authoring apps. *TESOL Journal*. <https://doi.org/10.1002/tesj.716>
- Zorrilla-Puerto, J., Lores-Gómez, B., Martínez-Requejo, S. y Ruiz-Lázaro, J. (2023). El papel de la robótica en Educación Infantil: Revisión sistemática para el desarrollo de habilidades. *RiiTE Revista Interuniversitaria de Investigación en Tecnología Educativa*, (15), 188-194. <https://doi.org/10.6018/riite.586601>
- Zou, M. y Huang, L. (2023). To use or not to use? Understanding doctoral students' acceptance of ChatGPT in writing through technology acceptance model. *Frontiers in Psychology*, 14, 1259531. <https://doi.org/10.3389/fpsyg.2023.1259531>

Fecha de recepción del artículo: 1 de diciembre de 2024

Fecha de aceptación del artículo: 11 de marzo de 2025

Fecha de aprobación para maquetación: 31 de marzo de 2025

Fecha de publicación en OnlineFirst: 23 de abril de 2025

Fecha de publicación: 1 de julio de 2025

Percepciones de futuros docentes y pedagogos sobre uso responsable de la IA. Un instrumento de medida

Perceptions of future teachers and pedagogues on responsible AI. A measurement instrument



- Melchor Gómez-García - *Universidad Autónoma de Madrid, UAM (España)*
- Julio Ruiz-Palmero - *Universidad de Málaga, UMA (España)*
- Moussa Boumadan-Hamed - *Universidad Autónoma de Madrid, UAM (España)*
- Roberto Soto-Varela - *Universidad de Valladolid, UVA (España)*

RESUMEN

Este estudio presenta la validez y fiabilidad en la creación de un instrumento diseñado para evaluar las percepciones de docentes y pedagogos en formación hacia la integración de la Inteligencia Artificial en tareas relacionadas con su profesión docente, teniendo en cuenta factores intrínsecos como la actitud hacia su uso responsable, el nivel de creatividad en la creación de material didáctico con estas herramientas, el disfrute asociado en el uso de estas herramientas, y el nivel de ansiedad al enfrentarse al aprendizaje de esta tecnología emergente en su formación académica y su relevancia en su futuro mercado laboral. Fue utilizado un diseño no experimental ex post facto a través de encuestas con un muestreo no probabilístico por conveniencia, con un total de 548 docentes y pedagogos en formación de facultades de Ciencias de la Educación del territorio español. Para la elaboración del instrumento, se utilizaron medidas de fiabilidad y validez. Respecto a la fiabilidad, fueron utilizados los índices Alfa de Cronbach, Coeficiente Spearman-Brown, Dos Mitades de Guttman y fiabilidad compuesta. Respecto a la validez, se utilizaron la validez de comprensión, constructo, convergente y discriminante. Los resultados demostraron una fiabilidad altamente satisfactoria, y en términos de validez se observó un buen ajuste del modelo. La versión final del instrumento consta de 25 ítems clasificados en cinco factores latentes.

Palabras clave: inteligencia artificial; docentes en formación; pedagogos en formación; instrumento psicométrico.

ABSTRACT

This study presents the validity and reliability in the creation of an instrument designed to evaluate the perceptions of teachers and pedagogues in training towards the integration of Artificial Intelligence in tasks related to their teaching profession, taking into account intrinsic factors such as the attitude towards its responsible use, the level of creativity in the creation of didactic material with these tools, the associated enjoyment in the use of these tools, and the level of anxiety when facing the learning of this emerging technology in their academic training and its relevance in their future labor market. A non-experimental ex post facto design was used through surveys with a non-probabilistic sampling by convenience, with a total of 548 teachers and pedagogues in training from faculties of Education Sciences in Spain. Reliability and validity measures were used for the elaboration of the instrument. Regarding reliability, Cronbach's Alpha, Spearman-Brown Coefficient, Guttman's Two Halves and composite reliability were used. Regarding validity, comprehension, construct, convergent and discriminant validity were used. The results showed a highly satisfactory reliability, and in terms of validity, a good model fit was observed. The final version of the instrument consists of 25 items classified in five latent factors.

Keywords: artificial intelligence; teachers in training; pedagogues in training; psychometric instrument.

INTRODUCCIÓN

Desde hace varias décadas, existe un creciente interés entre la población acerca de las posibilidades que las tecnologías digitales pueden ofrecer en el ámbito educativo, destacándose en particular la Inteligencia Artificial (IA) como una de las innovaciones más prometedoras (Şahín, 2024). Autores como Lambert y Stevens (2023) subrayan que la IA se refiere a un grupo de sistemas computacionales diseñados para aprender y realizar predicciones. No obstante, la IA que realmente ha captado la atención en los últimos años es la denominada “inteligencia artificial generativa” (IAG) la cual puede crear contenido nuevo y original como textos, imágenes, presentaciones, audios y videos a partir de prompts (instrucción) (Alenezi et al., 2023; González-Mayorga et al., 2024).

La IAG está demostrando ser una herramienta con un potencial transformador en múltiples ámbitos, incluido el educativo. De hecho, los investigadores Ng et al. (2021) subrayan que la IA “se convierte potencialmente en una de las habilidades tecnológicas más importantes del siglo XXI” (p. 2), principalmente “desde el lanzamiento de ChatGPT en noviembre de 2022, que trajo el concepto de IA generativa a la atención del público y despertó un creciente interés en su impacto potencial en la educación” (Yu y Guo, 2023, p. 1). En este contexto, gracias a la IA, los “educadores pueden aprovechar las experiencias de aprendizaje personalizadas, la generación de contenido adaptativo y el apoyo en tiempo real para los estudiantes” (Ruiz-Rojas et al., 2023, p. 1).

No obstante, la integración de la IA exige una planificación rigurosa y una formación adecuada para garantizar un uso efectivo y responsable (Gocen y Aydemir, 2020; Hwang y Chen, 2023). Sin embargo, si el objetivo de los docentes es preparar a los estudiantes para los retos del mercado laboral del siglo XXI, es imprescindible fomentar el desarrollo de habilidades, conocimientos y actitudes en IA y en concreto en IAG que les permitan adaptarse y prosperar en un mercado altamente dinámico (Farrelly y Baker, 2023). Por lo que las instituciones educativas deben ajustar y adaptar de manera ágil sus programas para incorporar el uso de la inteligencia artificial (IA) en todas las áreas del conocimiento (Bellas et al., 2023). No obstante, surge una serie de interrogantes: ¿Están los docentes y pedagogos en formación adecuadamente formados en IA e IAG para enfrentar los desafíos de su ámbito profesional? Para responder a estas cuestiones es necesario hacerse una pregunta previa, ¿Existen instrumentos de medición que evalúen el uso responsable y posibles factores intrínsecos de los docentes y pedagogos en formación con relación al uso de la IA e IAG y su repercusión en su futuro laboral?

Aunque se han realizado avances en los últimos años sobre la medición educativa en IA, existe una notable escasez de instrumentos psicométricos en la literatura científica enfocados en docentes y pedagogos en formación. Para respaldar esta afirmación, se llevó a cabo una búsqueda bibliográfica de instrumentos de evaluación de IA bajo los siguientes criterios: 1) que los instrumentos hubieran sido

publicados en los últimos 2-3 años ya que es cuando más ha emergido el uso de la IA e IAG; 2) que los instrumentos estén enfocados en docentes y pedagogos en formación o docentes en activo ya que ambos pertenecen al mismo colectivo; y 3) que los instrumentos sean fiables y válidos, ya sea mediante validación de expertos o validación de constructo (análisis factorial exploratorio o confirmatorio, AFE y AFC). La Tabla 1 refleja la existencia de varios instrumentos, cada uno con una estructura y taxonomía concreta. A partir de la revisión de los estudios presentados, se observa que ninguno de estos estudios está enfocado en docentes y pedagogos en formación, excepto uno de ellos (Espinoza-San Juan et al., 2024), el cual carece de propiedades psicométricas.

Tabla 1
Análisis bibliográfico sobre instrumentos basados en IA en educación

Autores	Propósito	Muestra	Fiabilidad	Validez	
				AFE	AFC
Hornberger et al. (2023)	Desarrollar y validar una prueba de alfabetización en IA.	Estudiantes Universitarios	No	No	Sí
Marquina et al. (2024)	Adaptar y validar un instrumento para medir las actitudes de los estudiantes hacia la inteligencia artificial.	Estudiantes universitarios	Sí	Sí	Sí
Nazaretsky et al. (2022)	Medir la confianza de los docentes en la tecnología educativa basada en IA.	Docentes de Ciencia	Sí	Sí	Sí
Saz-Pérez et al. (2024)	Validar un cuestionario sobre el uso de programas de IAG.	Docentes	Sí	Sí	No
Ng et al. (2023)	Examinar cómo los estudiantes desarrollan la alfabetización en IA.	Estudiantes de Educación Secundaria	Sí	Sí	Sí
Grájeda et al. (2024)	Evaluar la adopción y el impacto de las herramientas de IA.	Estudiantes universitarios	Sí	Sí	Sí
Chai et al. (2024)	Desarrollo y validación de una escala sobre la Intención de Aprendizaje de IA.	Estudiantes universitarios	No	No	Sí
Üzümlü et al. (2024)	Desarrollar una escala de percepción de los docentes respecto al uso de la inteligencia artificial en educación.	Docentes de diferentes etapas educativas	Sí	Sí	Sí

Morales-García et al. (2024)	Desarrollar y validar una Escala de Dependencia de IA.	Estudiantes universitarios	Sí	Sí	Sí
Yilmaz et al. (2023)	Diseñar un instrumento sobre la aceptación de las aplicaciones de IAG.	Estudiantes universitarios	Sí	Sí	Sí
Espinoza-San Juan et al. (2024)	Validar un cuestionario para evaluar la percepción y el uso de la Inteligencia Artificial Generativa (IA-Gen).	Estudiantes de pedagogía	Sí	No	No
Cheng et al. (2023)	Diseñar un instrumento para medir las concepciones sobre la IA en la educación.	Estudiantes universitarios	Sí	No	Sí
Jang et al. (2022)	Desarrollar un instrumento para evaluar las actitudes hacia la ética de la IA.	Estudiantes universitarios	Sí	Sí	Sí
Kim y Lee (2022)	Desarrollar una herramienta que mida la alfabetización en IA.	Estudiantes de secundaria	Sí	Sí	Sí
Wang et al. (2024)	Desarrollar una escala para evaluar la interactividad con la IA.	Estudiantes universitarios	Sí	Sí	Sí

Como se observa en la tabla previa, faltan herramientas que evalúen el uso responsable que hacen los docentes y pedagogos en formación, así como factores intrínsecos, como la creatividad para la creación de material didáctico y su planificación en tareas didácticas, el disfrute asociado a la hora de usar aplicaciones tecnológicas emergentes, o en cambio, el propio miedo-stress que pueden sentir en el propio aprendizaje y uso de las aplicaciones de AIG. Este artículo tiene como propósito la creación de un nuevo instrumento psicométrico, con un enfoque único al centrarse en estos aspectos específicos y en un colectivo que ha sido poco investigado hasta el momento.

MARCO TEÓRICO SOBRE LA TAXONOMÍA DE LOS FACTORES DEL INSTRUMENTO

El uso de la IA e IAG por parte de docentes en activo y formación está condicionado por una amplia variedad de factores que desempeñan un papel crucial (Tiwari et al., 2024). En primer lugar, uno de los principales factores es el uso responsable de esta tecnología (Aler et al., 2024), ya que “si bien ofrecen grandes oportunidades, los sistemas de IA también generan ciertos riesgos que deben gestionarse de manera adecuada y proporcionada” (European Commission, 2019, p. 4), por lo que es necesario que “los estudiantes reciban capacitación sobre el uso responsable y las

pautas éticas” (Hasanein y Sobaih, 2023, p. 2609). Estamos de acuerdo con las reflexiones de Brandão et al. (2024) al afirmar que existe un número de estudiantes los cuales probablemente utilizan herramientas con IAG debido a su accesibilidad y uso gratuito en contextos extracurriculares. Por ello, consideramos que el empleo de estas tecnologías exige no solo una comprensión técnica, sino también una reflexión ética por parte de los docentes y pedagogos en formación sobre el uso responsable de la IA e IAG, teniendo en cuenta las exigencias y transformaciones de sus futuros mercados laborales.

En segundo lugar, según Uzumcu y Acilmis (2024), la afirmación “los innovadores son personas creativas y emprendedoras dispuestas a asumir riesgos y abiertas a nuevas ideas” (p. 1112) encaja perfectamente con el rol de los docentes en la integración de la inteligencia artificial en sus prácticas pedagógicas. Los educadores que se destacan por su creatividad y mentalidad emprendedora son fundamentales para aprovechar el potencial de las TIC (Alemany et al., 2021), en nuestro caso, del uso de la IA e IAG ya que estas tecnologías “podría ayudarles a ser creativos en su práctica” (Chounta et al., 2022, p. 735). En lugar de limitarse a usar la IA e IAG como un simple facilitador de tareas, esta tipología de docentes tendrá una oportunidad para diseñar metodologías activas y personalizadas que mejoren la experiencia de aprendizaje (Kaouni et al., 2023), y consecuentemente, la IA podría ayudar a contribuir “al logro del cuarto ODS (objetivos desarrollo sostenible) propuesto por la ONU (Organización de las Naciones Unidas) al promover una educación inclusiva, equitativa y de calidad. Esta apertura hacia nuevas posibilidades y disposición a experimentar creativamente va a permitir que la IA e IAG se conviertan en herramientas dinámicas, capaces de enriquecer el aula y fomentar la autonomía no solo del docente sino también de los estudiantes (Mohamed et al., 2024).

Un tercer concepto es el concepto de *flow digital* traducido al castellano como el disfrute asociado al uso de la tecnología. La teoría del *flow digital* se refiere a un estado de inmersión profunda y concentración total en una actividad digital, donde el usuario experimenta una combinación de disfrute e interés intrínseco (Guillén-Gámez et al., 2023). Este concepto fue adaptado de la teoría del flujo de Mihaly Csikszentmihalyi, que originalmente describía el “flow” como un estado mental en el que una persona se encuentra completamente absorbida por una tarea (Csikszentmihalyi et al., 2014). En otras palabras, ello es “un estado psicológico en el que la persona se siente simultáneamente cognitivamente eficiente, motivada y feliz” (Moneta y Csikszentmihalyi, 1996, p. 277). En el contexto digital, el *flow* se refiere a actividades realizadas con tecnologías o plataformas digitales, en las que las personas sienten una alta satisfacción mientras interactúan con estas herramientas (Zhan et al., 2024). En el contexto de esta investigación, los docentes y pedagogos en formación que disfruten del uso de la IA generativa lograrán dos beneficios clave: primero, fortalecerán sus habilidades digitales, lo que les permitirá sobresalir en un mercado laboral altamente competitivo y demandante en habilidades digitales. Segundo, desarrollarán un estado de disfrute, facilitando su adaptación

a entornos educativos dinámicos e innovadores, alineándose con las expectativas de empleadores que buscan profesionales con capacidades tecnológicas avanzadas y flexibilidad cognitiva.

Por último, estos colectivos docentes podrían experimentar estrés o ansiedad al aprender a utilizar la IA e IAG en contextos educativos, principalmente debido al impacto de estas tecnologías emergentes en las tasas de empleo y la vida social (Hopcan et al., 2024). La ansiedad relacionada con la IAG se puede describir como la inquietud o el miedo que algunas personas experimentan ante los posibles efectos negativos y riesgos que pueden surgir con el uso de estas tecnologías en diferentes sectores sociales (Wang y Wang, 2022). En el contexto educativo, existen varias razones por las que los educadores pueden sentir ansiedad sobre la IA. Por un lado, la idea de que ello pueda reemplazar a los docentes, lo cual provocaría una pérdida de empleos y una disminución en la calidad educativa (Wang et al., 2022). Y, por otro lado, algunos docentes temen que los sistemas de IA no puedan captar aspectos clave de la enseñanza, como la construcción de relaciones con los estudiantes y la personalización del aprendizaje (Ouyang et al., 2022). Sin embargo, hay una tercera razón que ha sido poco explorada hasta el momento en la literatura científica: la posibilidad de que los docentes y pedagogos en formación no posean y adquieran durante el grado educativo las habilidades necesarias para utilizar la IA e IAG, lo que podría limitar su capacidad para competir eficazmente en el mercado laboral y educativo.

Teniendo en consideración este marco teórico previo, el objetivo a alcanzar es validar un instrumento psicométrico para evaluar tanto el uso responsable de la IA e IAG por parte de docentes y pedagogos en formación, como de factores intrínsecos como la actitud, creatividad, disfrute (*flow digital*) y ansiedad hacia la integración de estas tecnologías en tareas relacionadas con su profesión docente.

MÉTODO

Diseño y tipo de muestreo

Para alcanzar los objetivos de este estudio, fue empleado un diseño no experimental (ex post facto) para evaluar las propiedades psicométricas (fiabilidad y validez) del instrumento diseñado. Los datos fueron recolectados a través del contacto directo de los autores con sus propios estudiantes, así como con otros estudiantes de sus universidades y de otras instituciones que ofrecían los grados de Magisterio y/o Pedagogía. La selección de la muestra se realizó de manera no probabilística durante los meses de septiembre y octubre de 2024, siguiendo un criterio de conveniencia, dado el acceso directo a estos grupos. La participación fue voluntaria y anónima, garantizando el consentimiento informado antes de la recolección de los datos. El proceso de recolección se llevó a cabo en horario de clase, previa autorización de los docentes responsables de las asignaturas en las que se encontraban los estudiantes.

Se explicó a los participantes el propósito del estudio y se les proporcionaron instrucciones claras sobre cómo completar el instrumento. La administración del cuestionario se realizó en formato digital mediante una plataforma en línea (*Google Forms*), lo que permitió recopilar respuestas de manera eficiente y sistematizada. La muestra final estuvo formada por un total de 548 docentes y pedagogos en formación. La Tabla 2 muestra la distribución de los participantes según diversas variables demográficas.

Tabla 2

Distribución de la muestra de los docentes en formación

Genero	Femenino (81.60 %, n=447)	Etapas educativas	Educación Infantil (26.50 %, n= 145)
	Masculino (18.40 %, n= 101)		Educación Primaria (29 %, n= 159)
Edad	Femenino (20.76±2.13)		Pedagogía (6.6 %, n= 36)
	Masculino (23.15±6.46)		Doble Grado (30.30 %, n= 166)
			Postgrado (7.7 %, n= 42)

Procedimiento en la elaboración del instrumento

El instrumento se creó tras revisar la literatura sobre el uso responsable de la IA por parte de docentes y pedagogos en formación, como de factores intrínsecos como el nivel de creatividad, disfrute tecnológico (*flow digital*) y estrés hacia la integración de la IA en tareas relacionadas con su profesión docente. Al encontrar pocos instrumentos sobre la temática, se decidió llevar a cabo el diseño y creación de un instrumento psicométrico. Se utilizó una escala Likert de siete puntos, donde el valor uno fue asociado a la etiqueta “totalmente en desacuerdo” mientras que el valor siete se asoció a “totalmente de acuerdo”.

Tras revisar la literatura y desarrollar una batería de 41 ítems, se tomó en cuenta las recomendaciones de Hair et al. (2010), que sugieren recoger una muestra entre cinco y diez veces el número de ítems del cuestionario para analizar sus propiedades psicométricas. En este estudio, los ítems iniciales fueron 41, resultando en una proporción de 13.37, bastante superior a esta recomendación. Para los posteriores procedimientos, se siguieron las pautas recomendadas por Pérez y Carretero-Dios (2005): validez de comprensión, validez de constructo, validez convergente y discriminante y fiabilidad. Para el desarrollo de los análisis estadísticos se utilizó el software SPSS V24 y AMOS V24 perteneciente a la compañía IBM.

RESULTADOS

Validez de comprensión: análisis estadístico de los ítems

Dado que el estudio de la normalidad de los datos a través de Kolmogorov-Smirnov es muy sensible a pequeñas desviaciones en escalas tipo Likert, no se consideró como único criterio de normalidad al no soportar su contraste ($p < .05$). En cambio, se siguió el criterio de Pérez y Medrano (2010) para evaluar la validez de comprensión, quienes sugieren que los ítems con asimetría y curtosis entre ± 1.5 son adecuados, junto con aquellos que tienen una desviación estándar superior a 1 (Meroño et al., 2018).

La Tabla 3 muestra los ítems del instrumento organizados según sus respectivos factores. Los resultados descriptivos indicaron que todos los ítems tienen una adecuada validez de comprensión, ya que sus valores se ubicaron dentro de los límites establecidos, excepto el ítem DIM2.1 el cual sobrepasó los umbrales recomendados según Pérez y Medrano (2010), por lo tanto, fue eliminado.

Tabla 3

Coefficientes de asimetría, curtosis y desviación típica

	DT	A	K
DIM1 “Actitud hacia el uso de la IA para la empleabilidad como futuro docente o pedagogo”			
<i>DIM1.1.</i> Considero que dominar herramientas de IA afectará positivamente mis oportunidades laborales.	1.75	-.90	.02
<i>DIM1.2.</i> Creo que las escuelas o instituciones socio-comunitarias deben contratar agentes educativos con experiencia en el uso de IA.	1.73	-.57	-.49
<i>DIM1.3.</i> Siento que tener formación en IA me hará más competitivo en el mercado laboral docente.	1.76	-1.04	.15
<i>DIM1.4.</i> Pienso que un director/jefe/evaluador verá la habilidad de usar IA como un requisito importante para obtener un puesto de trabajo.	1.73	-.43	-.70
<i>DIM1.5.</i> Considero que ganaré oportunidades de empleo si me capacito adecuadamente en la implementación de IA.	1.69	-.78	-.28
<i>DIM1.6.</i> Creo que, si desarrollo habilidades en IA, tendré más ventajas en la búsqueda de empleo sobre otros compañeros.	1.74	-.81	-.22
<i>DIM1.7.</i> Considero que las futuras ofertas laborales para docentes, educadores y pedagogos están cada vez más centradas en la integración de IA en el proceso educativo de los individuos.	1.76	-.58	-.62
<i>DIM1.8.</i> Siento que, desarrollando habilidades en IA, seré percibido como un docente/formador más innovador por los futuros compañeros de trabajo.	1.73	-.78	-.31

	DT	A	K
DIM2 “Flow digital en el uso responsable y seguro de la IA”			
<i>DIM2.1.</i> Me gustaría aprender aplicaciones de inteligencia artificial que me puedan servir en mi futuro laboral.	1.61	-1.55	-1.50
<i>DIM2.2.</i> Me resulta ameno aprender sobre aplicaciones de inteligencia artificial	1.63	-.79	-.60
<i>DIM2.3.</i> Me siento motivado/a para investigar y familiarizarme con las herramientas de inteligencia artificial disponibles para el mundo educativo.	1.64	-.83	-.50
<i>DIM2.4.</i> Estoy interesado/a en recibir formación adicional sobre aplicaciones de inteligencia artificial que puedan mejorar mis habilidades como futuro docente/educador/pedagogo.	1.63	-1.02	.23
<i>DIM2.5.</i> Me motiva pensar que el dominio de la IA durante mi formación universitaria me permitirá destacarme en el mercado laboral y aumentar mis posibilidades de empleo en instituciones sociales.	1.68	-.90	-.56
<i>DIM2.6.</i> Me resulta gratificante pensar que aprender a utilizar herramientas de IA durante el grado universitario mejorará mis oportunidades de empleo en centros educativos o socio-comunitarios o empresas, ya que se valorará mi capacidad para personalizar el contenido de las clases.	1.64	-1.00	.27
<i>DIM2.7.</i> Me motiva pensar que aprender durante el grado universitario el dominio de la IA aumentará mis posibilidades de ser contratado en instituciones educativas y sociales que busquen innovar en sus programas educativos.	1.65	-.94	.12
<i>DIM2.8.</i> Creo que aprender IA durante el grado universitario será un factor decisivo para destacar en instituciones que promuevan la educación digital y el aprendizaje interactivo.	1.66	-.80	-.26
<i>DIM2.9.</i> Confío en que la habilidad de crear experiencias de aprendizaje inmersivas con IA será un diferenciador importante al postularme para trabajos en asociaciones que apoyan el desarrollo comunitario.	1.63	-.92	.12
<i>DIM2.10.</i> Me gustaría aprender durante el grado universitario el uso de IA para la creación de contenido educativo ya que me ayudará a ser más competitivo en sectores donde se espera que los docentes se adapten a nuevas tecnologías.	1.62	-1.14	.50
DIM3 “Creatividad para usar la IA de manera responsable como futuro docente o pedagogo”			
<i>DIM3.1.</i> Considero que tengo la creatividad necesaria como futuro agente educativo para planificar mis acciones educativas con inteligencia artificial.	1.48	-.76	.07

	DT	A	K
<i>DIM3.2.</i> Creo que tengo bastante creatividad para adaptar con inteligencia artificial los contenidos que estoy aprendiendo en el grado universitario a los colectivos con los que trabajaré en mi futuro mercado laboral.	1.43	-.64	-.20
<i>DIM3.3.</i> Siento que tengo la capacidad de ser creativo al desarrollar materiales educativos innovadores utilizando inteligencia artificial.	1.48	-.70	-.16
<i>DIM3.4.</i> Considero que tengo la habilidad para crear actividades interactivas y atractivas con la ayuda de inteligencia artificial, adaptadas a los colectivos con los que trabajaré una vez que acabe mi grado universitario.	1.50	-.67	-.20
<i>DIM3.5.</i> Considero que poseo la capacidad de crear experiencias de aprendizaje personalizadas e inmersivas utilizando inteligencia artificial.	1.57	-.63	-.32
<i>DIM3.6.</i> Sé que soy capaz de utilizar mi creatividad para adaptar y mejorar mis futuras tareas educativas con inteligencia artificial para atender mejor las necesidades de los colectivos con los que trabajaré.	1.52	-.71	-.19
<i>DIM3.7.</i> Considero que tengo la capacidad para desarrollar actividades innovadoras con inteligencia artificial para poder utilizarlas en instituciones sociales y educativas.	1.52	-.77	-1.06
DIM4 “Ansiedad para usar la IA en mi futuro contexto laboral”			
<i>*DIM4.1.</i> Me resulta estresante considerar las habilidades que debo adquirir sobre inteligencia artificial durante el grado universitario para destacar en mi futuro empleo como docente/educador.	1.83	-.11	-1.06
<i>*DIM4.2.</i> Me agobia la idea de que la falta de conocimientos impartidos en la universidad sobre inteligencia artificial pueda limitar mis oportunidades laborales en el ámbito educativo.	1.89	.26	-1.16
<i>*DIM4.3.</i> Me siento frustrado al enfrentar la presión de aprender a utilizar herramientas de inteligencia artificial para la creación de recursos educativos.	1.88	-.06	-1.14
<i>*DIM4.4.</i> Me inquieta no tener suficiente tiempo para desarrollar las competencias necesarias en inteligencia artificial antes de ingresar al mercado laboral.	1.85	.16	-1.11
<i>*DIM4.5.</i> Me resulta difícil imaginar cómo integrar la inteligencia artificial en mis clases sin sentirme abrumado por la carga de aprendizaje.	1.84	-.15	-1.09
<i>*DIM4.6.</i> Me cansa pensar en la complejidad de aprender durante el grado universitario a crear materiales didácticos con inteligencia artificial mientras finalizo mis estudios.	1.91	-.23	-1.15
<i>*DIM4.7.</i> Me inquieta la posibilidad de que mis habilidades en inteligencia artificial no sean suficientes para competir en el mercado laboral y educativo.	1.91	.12	-1.20

	DT	A	K
DIM5 “Uso responsable de la IA para la empleabilidad como futuro docente o pedagogo”			
<i>DIM5.1.</i> Tengo en cuenta la privacidad y el uso responsable de datos cuando implemento herramientas de inteligencia artificial.	1.67	-.88	-.92
<i>DIM5.2.</i> Soy consciente de las posibles consecuencias negativas del mal uso de la IA y procuro evitarlas en mi actividad.	1.62	-.99	-.93
<i>DIM5.3.</i> Como futuro profesional, me informo sobre las normativas y regulaciones para el uso responsable de la inteligencia artificial enfocado en las labores de mi futuro mercado laboral.	1.85	-.35	-.93
<i>DIM5.4.</i> Me preocupo por las implicaciones sociales, éticas, y de uso responsable de la IA, las cuales pueden afectar a los colectivos con los que trabaje en mi futuro mercado laboral.	1.64	-.75	-.22
<i>DIM5.5.</i> Creo que los futuros profesionales debemos ser conscientes del papel que jugará la inteligencia artificial en la educación y estar preparados para gestionarla de forma responsable en nuestro ejercicio profesional.	1.39	-1.32	1.18
<i>DIM5.6.</i> Creo que es fundamental que los futuros profesionales estemos capacitados en el uso responsable de la inteligencia artificial en nuestras futuras prácticas laborales.	1.45	-1.24	.82
<i>DIM5.7.</i> Me siento encargado de aprender sobre el uso responsable de la IA para prepararme mejor de cara a mi futuro mercado laboral.	1.46	-1.04	.43
<i>DIM5.8.</i> Me preocupo por las implicaciones sociales y de uso responsable de la IA en el diseño y creación de material didáctico con IA.	1.57	-.89	.02
<i>DIM5.9.</i> Pienso que el uso responsable de la inteligencia artificial será una competencia clave en el entorno laboral.	1.44	-1.27	.98

Nota: elaboración propia. DT= desviación típica; A= asimetría; K= curtosis. * Ítems en orden inverso

Tras verificar los valores de dispersión, de asimetría y curtosis, se comprobó si la fiabilidad de cada ítem aumentaba o disminuía en comparación con el Alfa de Cronbach global al eliminarlo. También se revisó el índice de homogeneidad (correlación elemento-total corregida) para descartar ítems con coeficientes inferiores a 0.4 (Shaffer y Kipp, 2010). Por ello, se eliminaron para los posteriores procedimientos los ítems DIM4.1, DIM4.3, DIM4.5, DIM4.6. En la Tabla 4 se observa los estadísticos de cada ítem una vez eliminados los ítems con poca carga de fiabilidad.

Tabla 4
Análisis del índice de discriminación de escala

	Media de escala si el elemento se ha suprimido	Varianza de escala si el elemento se ha suprimido	Correlación total de elementos corregida	Alfa de Cronbach si el elemento se ha suprimido
DIM1				
DIM1.1	174.2755	1379.630	.670	.962
DIM1.2	174.7920	1381.269	.667	.962
DIM1.3	174.2372	1373.995	.712	.961
DIM1.4	174.9489	1379.683	.681	.962
DIM1.5	174.4562	1373.645	.746	.961
DIM1.6	174.5365	1376.958	.700	.961
DIM1.7	174.8522	1381.984	.650	.962
DIM1.8	174.4982	1376.503	.704	.961
DIM2				
DIM2.2	174.5456	1387.930	.654	.962
DIM2.3	174.4635	1382.026	.698	.961
DIM2.4	174.1898	1378.600	.732	.961
DIM2.5	174.3832	1369.988	.785	.961
DIM2.6	174.3047	1369.130	.812	.961
DIM2.7	174.3376	1367.438	.820	.961
DIM2.8	174.5018	1374.580	.752	.961
DIM2.9	174.3759	1374.209	.773	.961
DIM2.10	174.0949	1374.126	.776	.961
DIM3				
DIM3.1	174.4069	1394.589	.665	.962
DIM3.2	174.4507	1392.577	.706	.961
DIM3.3	174.4124	1392.396	.684	.962
DIM3.4	174.5128	1391.950	.679	.962
DIM3.5	174.6788	1390.866	.654	.962
DIM3.6	174.4872	1390.861	.677	.962
DIM3.7	174.5255	1391.252	.675	.962
DIM4				
DIM4.2	175.2263	1414.307	.466	.964
DIM4.4	175.3339	1414.837	.573	.964
DIM4.7	175.3522	1419.026	.529	.964

	Media de escala si el elemento se ha suprimido	Varianza de escala si el elemento se ha suprimido	Correlación total de elementos corregida	Alfa de Cronbach si el elemento se ha suprimido
DIM5				
DIM5_1	174.2938	1403.780	.509	.963
DIM5_2	174.1478	1400.909	.548	.962
DIM5_3	175.0310	1409.752	.409	.963
DIM5_4	174.4325	1405.332	.505	.963
DIM5_5	173.7391	1401.034	.643	.962
DIM5_6	173.8139	1399.421	.631	.962
DIM5_7	174.0201	1396.389	.656	.962
DIM5_8	174.1880	1397.703	.597	.962
DIM5_9	177.7591	1231.130	.669	.950

Nota: elaboración propia

Por último, en el contexto de este tipo de validez, Asencio et al. (2017) recomiendan verificar la unidimensionalidad del instrumento, es decir, el nivel de varianza común entre los ítems de la prueba, mediante el análisis de la correlación entre sus distintas dimensiones. La Tabla 5 muestra la matriz de correlación entre los factores latentes del instrumento tras aplicar la rotación *Oblimin*, lo que indica que los factores están correlacionados. Este resultado sugiere que el instrumento presenta una estructura unidimensional, basada en cinco factores latentes.

Tabla 5
Matriz de correlación factorial

Factores latentes	DIM2	DIM5	DIM3	DIM4	DIM1
DIM2	1.000				
DIM5	.421	1.000			
DIM3	.544	.501	1.000		
DIM4	-.267	-.244	-.207	1.000	
DIM1	.683	.381	.513	-.306	1.000

Nota: elaboración propia

Validez de constructo: análisis factorial exploratorio

Después de comprobar la validez de comprensión, se evaluó la unidimensionalidad del instrumento mediante el AFE. Este tipo de validez fue realizada bajo la rotación *Oblimin* y el método de Factorización de Ejes Principales para explicar la mayor parte de la varianza común. Este enfoque es adecuado incluso cuando el supuesto de normalidad no se cumple plenamente (Fabrigar et al., 1999).

En primer lugar, se comprobó la adecuación de los ítems respecto a la pertenencia a sus factores latentes, con la prueba de esfericidad de Bartlett ($\chi^2=22279.671$; $gl=666$; $p. < 0.05$) y el índice de esfericidad de Kaiser-Meyer-Olkin ($KMO= 0.960$). Ambos coeficientes fueron satisfactorios. Según el criterio de Kaiser de Cattell (1966), solo los factores con autovalores superiores a 1 deben extraerse; en este caso, los cinco primeros factores latentes cumplieron con este criterio, como se muestra en la Tabla 6. Así, la escala creada consta de cinco factores, explicando una varianza total del 76.89 %: DIM2 (47.13 %), DIM5 (11.07 %), DIM3 (7.32 %), DIM4 (6.05 %), Y DIM1 (5.32 %).

Tabla 6
Autovalores y varianza explicada

Factores latentes	Autovalores	% de varianza	% Acumulado
DIM2	17.435	47.123	47.123
DIM5	4.095	11.069	58.192
DIM3	2.708	7.320	65.511
DIM4	2.237	6.046	71.557
DIM1	1.970	5.324	76.881

Nota: elaboración propia

En la Tabla 7 se observa que el mayor porcentaje de varianza verdadera en la unidimensionalidad del instrumento fue la dimensión 2 (disfrute tecnológico en el uso responsable y seguro de la IA), incluyendo los ítems DIM2.4, DIM2.3, DIM2.10, DIM2.8, DIM2.5, DIM2.2, DIM2.7, DIM2.6, y DIM2.9. El segundo factor con mayor carga fue la dimensión 5 (Uso responsable de la IA para la empleabilidad como futuro docente o pedagogo), incluyendo los ítems DIM5_2, DIM5_8, DIM5_4, DIM5_7, DIM5_1, DIM5_6, DIM5_5, DIM5_9, DIM5_3. El tercer factor con mayor saturación fue la dimensión 3 (Creatividad para usar la IA de manera responsable como futuro docente o pedagogo), incluyendo los ítems DIM3.4, DIM3.5, DIM3.7, DIM3.6, DIM3.3, DIM3.2 y DIM3.1. El cuarto factor en orden de saturación fue la dimensión 4 (Ansiedad para usar la IA en mi futuro contexto laboral), con los ítems DIM4.7, DIM4.2 y DIM4.4. El factor con menor porcentaje fue la dimensión 1

(Actitud hacia el uso de la IA para la empleabilidad como futuro docente o pedagogo), con los ítems DIM1.6, DIM1.5, DIM1.4, DIM1.7, DIM1.3, DIM1.8, DIM1.2 y DIM1.1.

Tabla 7
Cargas factoriales rotadas

Dimensiones	Factores latentes				
	1	2	3	4	5
DIM2.4	.892				
DIM2.3	.858				
DIM2.10	.832				
DIM2.8	.831				
DIM2.5	.816				
DIM2.2	.812				
DIM2.7	.809				
DIM2.6	.794				
DIM2.9	.777				
DIM5.2		.872			
DIM5.8		.852			
DIM5.4		.829			
DIM5.7		.809			
DIM5.1		.799			
DIM5.6		.782			
DIM5.5		.730			
DIM5.9		.694			
DIM5.3		.619			
DIM3.4			.897		
DIM3.5			.888		
DIM3.7			.880		
DIM3.6			.872		
DIM3.3			.859		
DIM3.1			.795		
DIM3.2			.794		
DIM4.7				.881	
DIM4.2				.867	
DIM4.4				.845	
DIM1.6					.921

Dimensiones	Factores latentes				
	1	2	3	4	5
DIM1.5					.908
DIM1.4					.866
DIM1.7					.850
DIM1.3					.803
DIM1.8					.792
DIM1.2					.735
DIM1.1					.662

Nota: elaboración propia

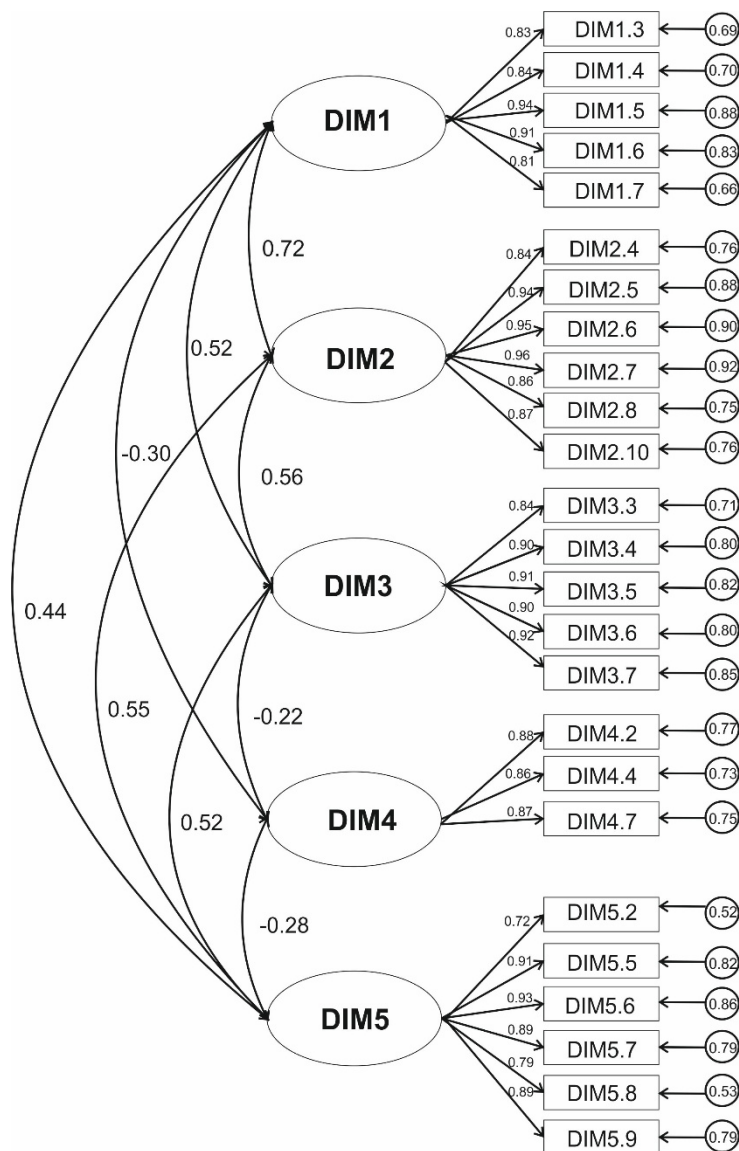
Validez de constructo: análisis factorial confirmatorio

Después del AFE, se realizó el AFC para verificar el ajuste de los datos mediante un modelo de ecuaciones estructurales, con el propósito de evaluar el ajuste del modelo teórico identificado en el AFE, según las recomendaciones de Thompson (2004) y otros autores con la creación de sus instrumentos psicométricos (Guillén Gámez et al., 2024; Soriano-Alcantara et al., 2024). Para interpretar los índices del AFC, se siguieron las recomendaciones de Bentler (1989) y Hu y Bentler (1999): discrepancia mínima/grados de libertad (CMIN/DF), donde valores menores a 5 indican un ajuste razonable; error cuadrático medio de aproximación (RMSEA), con valores inferiores a 0.07 considerados óptimos; y los índices de bondad de ajuste (GFI), ajuste comparativo (CFI) y ajuste normalizado (NFI), considerados adecuados cuando son iguales o superiores a 0.9.

Se llevaron a cabo diferentes modelos psicométricos donde fueron eliminados los siguientes ítems hasta encontrar el modelo que mejores propiedades psicométricas tenía: DIM1.1, DIM1.2, DIM1.8, DIM2.2, DIM2.3, DIM2.9, DIM3.1, DIM3.2, DIM5.1, DIM5.3, DIM5.4. Los coeficientes obtenidos fueron los siguientes: CMIN/DF (2.487) con un mayor inferior a 5; RMSEA (.052), siendo inferior a .07; GFI (.910), CFI (.973) y NFI (.955) con valores superiores a .90. La Figura 1 presenta el modelo factorial final obtenido del AFC. En esta figura también se muestran los valores de correlación estandarizados derivados del AFC.

Figura 1

Modelo de ecuaciones estructurales. Instrumento definitivo



Nota: elaboración propia

Validez convergente y discriminante

Una vez llevada a cabo la validez de constructo bajo el AFE y AFC, se verificaron dos tipos más de validez más. Por un lado, la validez convergente, que se refiere a la confianza en que los ítems evaluados miden el mismo factor latente (Cheung y Wang, 2017), utilizando los valores de varianza promedio extraída (AVE) los cuales han de ser superiores a .50, tal y como recomiendan Hair et al. (2010). Además, el valor de la raíz cuadrada del AVE en la diagonal debe ser mayor que las correlaciones entre factores (Hair et al., 2010). En la Tabla 8 se observa que los valores AVE superan .50 y que las raíces cuadradas de los AVE (en la diagonal y en negrita) son mayores que las correlaciones entre los factores latentes. Y, por otro lado, la validez discriminante, la cual evalúa hasta qué punto un constructo es verdaderamente distinto de otros constructos dentro de un modelo de investigación, y para ello es utilizado el índice MSV (máxima varianza compartida al cuadrado), cuyo requisito es que su valor sea inferior al AVE de cada factor latente (Fornell y Larcker, 1981). Según los resultados en la Tabla 8 se conserva la validez discriminante entre los factores latentes del instrumento.

Tabla 8
Coefficientes validez convergente y discriminante

Factores	AVE	MSV	Raíz cuadrada AVE (diagonal) y Correlaciones entre factores				
			DIM2	DIM5	DIM3	DIM4	DIM1
DIM2	.819	.521	.905				
DIM5	.735	.300	.548	.857			
DIM3	.799	.319	.565	.520	.894		
DIM4	.750	.097	-.312	-.280	-.222	.866	
DIM1	.751	.521	.722	.445	.518	-.301	.866

Nota: elaboración propia

Análisis de fiabilidad

Finalmente, se calculó la fiabilidad de cada factor latente del instrumento, así como la consistencia interna global. Para ello, se emplearon los coeficientes alfa de Cronbach, Spearman-Brown, Guttman y confiabilidad compuesta (CR), teniendo en consideración que los valores recomendados han de ser superiores a 0.7 (Nunally, 1978; Heinzl et al., 2011). Los resultados obtenidos en los cuatro índices fueron muy satisfactorios (Tabla 9), indicando que la consistencia interna del instrumento es adecuada.

Tabla 9
Coeeficientes de fiabilidad

Dimensión	DIM1	DIM2	DIM3	DIM4	DIM5	TOTAL
Alfa de Cronbach	.937	.964	.952	.900	.941	.924
Spearman-Brown	.929	.954	.945	.902	.903	.792
Dos mitades de Guttman	.888	.963	.899	.808	.903	.770
CR	.938	.964	.952	.900	.943	-

Nota: elaboración propia

CONCLUSIONES Y DISCUSIONES

En el contexto educativo actual, marcado por la rápida integración de la IA e IAG (Şahín, 2024), este estudio se propuso desarrollar y validar un instrumento que permita medir las autopercepciones de docentes y pedagogos en formación sobre el uso de la IA y su relevancia en el mercado laboral. Dada la creciente importancia de esta tecnología emergente en el ámbito formativo y profesional, contar con una herramienta que evalúe múltiples dimensiones del uso de la IA e IAG resulta fundamental (Ng et al., 2021), ya que los educadores van a poder aprovechar estas herramientas para personalizar el aprendizaje en el estudiantado, generar contenidos multimedia como imágenes, videos o texto, así como analizar el aprendizaje en tiempo real (Ruiz-Rojas et al., 2023).

La mayoría de los instrumentos de medición creados hasta ahora han utilizado muestras de estudiantes a nivel general (Hornberger et al., 2023; Marquina et al., 2024; Nazaretsky et al., 2022; Saz-Pérez et al., 2024; Ng et al., 2023; Grájeda et al., 2024; Chai et al., 2024; Üzüml et al., 2024; Morales-García et al., 2024; Yilmaz et al., 2023), pero muy pocos han sido enfocados en docentes y pedagogos en formación (Espinoza-San Juan et al., 2024), dándole un valor añadido a este estudio. Este instrumento permite medir no solo el compromiso ético y responsable hacia la tecnología en términos de empleabilidad, sino también la actitud de los docentes y pedagogos en formación respecto a cómo el uso de la IA e IAG podría influir en sus oportunidades laborales. Además, incluye la evaluación de factores como el disfrute tecnológico (*flow digital*), la creatividad en el uso de la IA y el grado de ansiedad asociado a su aprendizaje y aplicación, aspectos que son claves para su desarrollo profesional.

Para la creación del instrumento psicométrico se siguieron los pasos recomendados por diferentes estudios realizados por Pérez y Carretero-Dios (2005), Guillén Gámez et al. (2024) o Soriano-Alcántara et al. (2024). Para tal fin se elaboró una versión inicial del instrumento con un total de 41 ítems repartidos en cinco factores latentes. Los ítems fueron creados para una escala Likert de siete puntos. En el proceso de validación psicométrica, se comprobó que la muestra era adecuada,

superando ampliamente la recomendación de Hair et al. (2010), al tener un número de participantes que multiplicaba por 10 la cantidad de ítems (con una proporción inicial de 13.37, y en la versión final del instrumento de 21.92).

Para asegurar la validez de comprensión, se eliminaron los ítems que no cumplían con los rangos establecidos, basándose en los coeficientes de asimetría y curtosis y en el índice de discriminación de la escala, tal y como recomiendan Meroño et al. (2018) y Pérez y Medrano (2010). En relación con la validez del constructo, durante el AFE no se descartó ningún ítem ya que todos alcanzaron la saturación mínima de carga de .40, siguiendo las recomendaciones de Cattell (1966). Este proceso dio lugar a los cinco factores latentes descritos previamente los cuales explicaron el 76.89 % de la varianza verdadera en las puntuaciones de los participantes.

En el AFC se realizó varios ajustes con la eliminación de los ítems con peor saturación en sus correspondientes factores latentes, hasta identificar un buen ajuste, según los criterios de Bentler (1989) y Hu y Bentler (1999). En esta versión final, los coeficientes encontrados de los índices CFI, NFI, RMSEA o CMIN fueron realmente satisfactorios. Además, se comprobó la validez convergente y discriminante de esta versión final del instrumento, basándose en las pautas recomendadas y seguidas por Cheung y Wang (2017), Guillén Gámez et al. (2024) o Soriano-Alcantara et al. (2024). Concretamente, encontramos valores satisfactorios tanto en el índice AVE como en el índice MSV, tal y como recomiendan autores con una gran relevancia en la comunidad científica como Hair et al. (2010) y Fornell y Larcker (1981).

En cuanto a la fiabilidad del instrumento, se obtuvieron excelentes propiedades psicométricas medidos con el índice Alpha de Cronbach, tanto en los cinco factores latentes que conforman el instrumento, así como en la valoración global del mismo ($\alpha=.924$). Los demás índices de ajuste utilizados para comprobar la consistencia interna del instrumento, como Spearman-Brown, dos mitades de Guttman y la fiabilidad compuesta (CR), también respaldaron la fiabilidad del instrumento. Los coeficientes se ubicaron dentro de los intervalos recomendados por Heinzl et al. (2011) y Nunally (1978).

Tras los diferentes análisis estadísticos llevados a cabo, la versión final del instrumento fue de 25 ítems, clasificados en cinco factores. El primer factor, llamado “Actitud hacia el uso de la IA para la empleabilidad como futuro docente o pedagogo” estuvo compuesto finalmente por cinco ítems; el segundo factor llamado “*flow digital* en el uso responsable y seguro de la IA” estuvo compuesto por un total de seis ítems; el tercer factor se tituló “Creatividad para usar la IA de manera responsable como futuro docente o pedagogo” con un total de cinco ítems; el cuarto factor fue “Ansiedad para usar la IA en mi futuro contexto laboral” con un total de tres ítems; y por último, la quinta dimensión fue “Uso responsable de la IA para la empleabilidad como futuro docente o pedagogo” con un total de seis ítems.

Aunque los resultados obtenidos muestran una fiabilidad y validez satisfactoria, existen ciertas limitaciones que deben reconocerse. La muestra utilizada en este estudio, compuesta exclusivamente por docentes en formación en España y

seleccionada mediante un muestreo no probabilístico, representa una limitación que restringe la generalización de los resultados a otros contextos culturales y educativos. La ausencia de una selección aleatoria y la concentración en un único país pueden influir en la aplicabilidad de los hallazgos a diferentes realidades educativas. Para mitigar esta limitación, futuros estudios podrían ampliar la muestra mediante un muestreo aleatorio y estratificado que incluya docentes y pedagogos en formación de distintos niveles educativos y de diversas regiones geográficas, tanto dentro de España como en otros países. Asimismo, la replicación del estudio en contextos internacionales permitiría evaluar la validez del instrumento en distintos entornos educativos y socioculturales, lo que contribuiría a incrementar su aplicabilidad y robustez. Asimismo, sería valioso realizar estudios longitudinales para analizar cómo evoluciona la percepción y el uso de la IAG a lo largo del tiempo y de acuerdo con el avance en la formación docente.

También sería útil adaptar y validar el instrumento en entornos multiculturales para realizar estudios mixtos que involucren tanto a los docentes y pedagogos en formación como a equipos directivos de instituciones educativas y laborales, con el fin de comprender cómo afecta el uso de la IA e IAG a la práctica docente en el aula y en la gestión laboral. Es decir, como trabajo futuro, se sugiere complementar el enfoque cuantitativo con análisis cualitativos que permitan una comprensión más profunda de las percepciones y experiencias de los participantes. La integración de entrevistas o grupos focales podría aportar una visión más detallada sobre el uso de la IA. Además, la rápida evolución de esta tecnología y la variabilidad en la autopercepción de la IA e IAG por parte del colectivo analizado, podría requerir de futuras adaptaciones del instrumento que se ha creado, con el propósito de mantener su relevancia y la pertinencia de los ítems al propio instrumento.

Respecto a las implicaciones teóricas y prácticas de este estudio científico, la creación de este instrumento contribuye al campo de la educación al proporcionar un marco de referencia para evaluar el uso responsable de la IA e IAG y su relevancia en el mercado laboral. La creación de este instrumento ofrece una guía valiosa para el diseño de programas formativos que integren la IA e IAG de manera ética y efectiva, analizando la autopercepción en creencias, creatividad, disfrute tecnológico (*flow digital*) o niveles de ansiedad al tener que aprender el uso seguro de estas herramientas emergentes de cara a incluirlas en su futura práctica profesional.

Además, este instrumento permite a las instituciones identificar áreas específicas de formación en IA e IAG, con el fin de fortalecer la empleabilidad de los egresados no solo en el contexto educativo de docentes en formación de las etapas de Educación Primaria y Educación Infantil, donde muchos trabajarán en colegios, sino también en otros entornos laborales y socio comunitarios en los que trabajarán los pedagogos. De esta forma, la formación en IA e IAG no solo mejorará las perspectivas profesionales de este colectivo, sino que también amplía sus oportunidades en sectores que valoran habilidades tecnológicas avanzadas, respondiendo así a las demandas de un mercado laboral en constante evolución.

Financiación

Este estudio ha sido patrocinado por la *Cátedra ERIA - UAM-Founderz-Microsoft* (Empleabilidad y uso responsable de la IA <https://catedraeria.com>) y financiado por el proyecto *Recursos para el uso responsable de la Inteligencia Artificial (RAI) en el desarrollo profesional y de habilidades docentes* (Ref. FPYE_026.24_INN)

REFERENCIAS

- Alemany Díaz, M. D. M., Vallés Lluch, A., Villanueva López, J. F. y García-Serra García, J. (2021). E-learning in “innovation, creativity and entrepreneurship”: Exploring the new opportunities and challenges of technologies. *Journal of Small Business Strategy (Online)*, 31(1), 39-50. <https://doi.org/10.21125/inted.2020.0686>
- Alenezi, M. A. K., Mohamed, A. M. y Shaaban, T. S. (2023). Revolutionizing EFL special education: how ChatGPT is transforming the way teachers approach language learning. *Innoeduca. International Journal of Technology and Educational Innovation*, 9(2), 5-23. <https://doi.org/10.24310/innoeduca.2023.v9i2.16774>
- Aler Tubella, A., Mora-Cantalops, M. y Nieves, J. C. (2024). How to teach responsible AI in Higher Education: challenges and opportunities. *Ethics and Information Technology*, 26(1), 1-14. <https://doi.org/10.1007/s10676-023-09733-7>
- Asencio, E. N., García, E. J., Redondo, S. R. y Ruano, B. T. (2017). *Fundamentos de la investigación y la innovación educativa*. UNIR editorial.
- Bellas, F., Guerreiro-Santalla, S., Naya, M. y Duro, R. J. (2023). AI curriculum for European high schools: An embedded intelligence approach. *International Journal of Artificial Intelligence in Education*, 33(2), 399-426. <https://doi.org/10.1007/s40593-022-00315-0>
- Bentler, P. M. (1989). *EQS structural equations program manual*. BMDP Statistical Software.
- Brandão, A., Pedro, L. y Zagalo, N. (2024). Teacher professional development for a future with generative artificial intelligence—an integrative literature review. *Digital Education Review*, (45), 151-157. <https://doi.org/10.1344/der.2024.45.151-157>
- Cattell, R. B. (1966). The screen test for the number of factors. *Multivariate Behavioral Research*, 1(2), 245-276. https://doi.org/10.1207/s15327906mbro102_10
- Chai, C. S., Yu, D., King, R. B. y Zhou, Y. (2024). Development and validation of the Artificial Intelligence Learning Intention Scale (AILIS) for University Students. *SAGE Open*, 14(2), 21582440241242188. <https://doi.org/10.1177/21582440241242188>
- Cheng, L., Umapathy, K., Rehman, M., Ritzhaupt, A., Antonyan, K., Shidfar, P., Nichols, J., Lee, M. y Abramowitz, B. (2023). Designing, developing, and validating a measure of undergraduate students' conceptions of artificial intelligence in education. *Journal of Interactive Learning Research*, 34(2), 275-311. <https://doi.org/10.1037/t93665-000>
- Cheung, G. W. y Wang, C. (2017). Current approaches for assessing convergent and discriminant validity with SEM: issues and solutions. *Academy of Management*

- Proceedings*, 2017(1), 12706. <https://doi.org/10.5465/AMBPP.2017.12706abstract>
- Chounta, I. A., Bardone, E., Raudsep, A. y Pedaste, M. (2022). Exploring teachers' perceptions of artificial intelligence as a tool to support their practice in Estonian K-12 education. *International Journal of Artificial Intelligence in Education*, 32(3), 725-755. <https://doi.org/10.1007/s40593-021-00243-5>
- Csikszentmihalyi, M., Csikszentmihalyi, M., Abuhamdeh, S. y Nakamura, J. (2014). Flow. *Flow and the foundations of positive psychology: The Collected Works of Mihaly Csikszentmihalyi*, 227-238. https://doi.org/10.1007/978-94-017-9088-8_15
- Espinoza-San Juan, J., Raby, M. D. y Sagredo-Lillo, E. (2024). Validación de un cuestionario sobre las percepciones y usos de la IA-Gen entre estudiantes de pedagogía. *Revista Ibérica de Sistemas e Tecnologías de Informação*, (E70), 574-585.
- European Commission. (2019). Directorate-General for communications networks content and technology: Ethics guidelines for trustworthy AI. Publications Office. <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C. y Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4(3), 272-299. <https://doi.org/10.1037/1082-989X.4.3.272>
- Farrelly, T. y Baker, N. (2023). Generative artificial intelligence: Implications and considerations for higher education practice. *Education Sciences*, 13(11), 1109. <https://doi.org/10.3390/educsci13111109>
- Fornell, C. y Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39-50. <https://doi.org/10.1177/002224378101800104>
- Gocen, A. y Aydemir, F. (2020). Artificial intelligence in education and schools. *Research on Education and Media*, 12(1), 13-21. <https://doi.org/10.2478/rem-2020-0003>
- González-Mayorga, H., Rodríguez-Esteban, A. y Vidal, J. (2024). El uso del modelo GPT de OpenAI para el análisis de textos abiertos en investigación educativa [Using OpenAI's GPT Model to Analyse Open Texts in Educational Research]. *Pixel-Bit. Revista de Medios y Educación*, 69, 227-253. <https://doi.org/10.12795/pixelbit.102032>
- Grájeda, A., Burgos, J., Córdova, P. y Sanjinés, A. (2024). Assessing student-perceived impact of using artificial intelligence tools: Construction of a synthetic index of application in higher education. *Cogent Education*, 11(1), 2287917. <https://doi.org/10.1080/2331186X.2023.2287917>
- Guillén-Gámez, F. D., Ruiz-Palmero, J. y García, M. G. (2023). Digital competence of teachers in the use of ICT for research work: development of an instrument from a PLS-SEM approach. *Education and Information Technologies*, 28(12), 16509-16529. <https://doi.org/10.1007/s10639-023-11895-2>
- Guillén-Gámez, F. D., Tomczyk, Ł., Colomo-Magaña, E. y Mascia, M. L. (2024). Digital competence of Higher Education teachers in research work: validation of an explanatory and confirmatory model. *Journal of E-Learning and Knowledge Society*, 20(3), 1-12. <https://doi.org/10.20368/1971-8829/1135963>
- Hair Jr, J. F., Black, W. C., Babin, B. J. y Anderson, R. E. (2010). *Multivariate data analysis* (pp. 785-785). Prentice Hall.
- Hasanein, A. M. y Sobaih, A. E. E. (2023). Drivers and consequences of ChatGPT use in higher education: Key stakeholder perspectives. *European Journal of Investigation in Health, Psychology and Education*, 13(11), 2599-2614. <https://doi.org/10.3390/ejihpe13110181>

- Heinzl, A., Buxmann, P., Wendt, O. y Weitzel, T. (Eds.). (2011). *Theory-guided modeling and empiricism in information systems research*. Springer Science & Business Media. <https://doi.org/10.1007/978-3-7908-2781-1>
- Hopcan, S., Türkmen, G. y Polat, E. (2024). Exploring the artificial intelligence anxiety and machine learning attitudes of teacher candidates. *Education and Information Technologies*, 29(6), 7281-7301. <https://doi.org/10.1007/s10639-023-12086-9>
- Hornberger, M., Bowersdorff, A. y Nerdel, C. (2023). What do university students know about Artificial Intelligence? Development and validation of an AI literacy test. *Computers and Education: Artificial Intelligence*, 5, 1-12. <https://doi.org/10.1016/j.caeai.2023.100165>
- Hu, L. y Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1-55. <https://doi.org/10.1080/10705519909540118>
- Hwang, G. J. y Chen, N. S. (2023). Exploring the potential of generative artificial intelligence in education: applications, challenges, and future research directions. *Journal of Educational Technology & Society*, 26(2), 1-19. [https://doi.org/10.30191/ETS.202304_26\(2\).0014](https://doi.org/10.30191/ETS.202304_26(2).0014)
- Jang, Y., Choi, S. y Kim, H. (2022). Development and validation of an instrument to measure undergraduate students' attitudes toward the ethics of artificial intelligence (AT-EAI) and analysis of its difference by gender and experience of AI education. *Education and Information Technologies*, 27(8), 11635-11667. <https://doi.org/10.1007/s10639-022-11086-5>
- Kaouni, M., Lakrami, F. y Laboudiya, O. (2023). The design of an adaptive E-learning model based on Artificial Intelligence for enhancing online teaching. *International Journal of Emerging Technologies in Learning (Online)*, 18(6), 202-219. <https://doi.org/10.3991/ijet.v18i06.35839>
- Kim, S. W. y Lee, Y. (2022). The artificial intelligence literacy scale for middle school students. *Journal of the Korea Society of Computer and Information*, 27(3), 225-238. <https://doi.org/10.9708/jksci.2022.27.03.225>
- Lambert, J. y Stevens, M. (2023). ChatGPT and generative AI technology: a mixed bag of concerns and new opportunities. *Computers in the Schools*, 1-25. <https://doi.org/10.1080/07380569.2023.2256710>
- Marquina, M. C. G., Pinto-Villar, Y. M., Aranzamendi, J. A. M. y Gutiérrez, B. J. A. (2024). Adaptación y validación de un instrumento para medir las actitudes de los universitarios hacia la inteligencia artificial. *Revista de Comunicación*, 23(2), 125-142. <https://doi.org/10.26441/RC23.2-2024-3493>
- Meroño, L., Calderón Luquin, A., Arias Estero, J. L. y Méndez Giménez, A. (2018). Diseño y validación del cuestionario de percepción del profesorado de Educación Primaria sobre el aprendizaje del alumnado basado en competencias (#ICOMPri2). *Revista Complutense de Educación*, 29(1), 215-235. <https://doi.org/10.5209/RCED.52200>
- Mohamed, A. M., Shaaban, T. S., Bakry, S. H., Guillén-Gámez, F. D. y Strzelecki, A. (2024). Empowering the Faculty of Education Students: Applying AI's Potential for Motivating and Enhancing Learning. *Innovative Higher Education*, 1-23. <https://doi.org/10.1007/s10755-024-09747-z>
- Moneta, G. B. y Csikszentmihalyi, M. (1996). The effect of perceived challenges and skills on the quality of subjective experience. *Journal of Personality*, 64(2), 275-310. <https://doi.org/10.1111/j.1467-6494.1996.tb00512.x>

- Morales-García, W. C., Sairitupa-Sanchez, L. Z., Morales-García, S. B. y Morales-García, M. (2024). Development and validation of a scale for dependence on artificial intelligence in university students. *Frontiers in Education*, 9, Article 1323898. <https://doi.org/10.3389/feduc.2024.1323898>
- Nazaretsky, T., Cukurova, M. y Alexandron, G. (2022, March). An instrument for measuring teachers' trust in AI-based educational technology. En *LAK22: 12th International Learning Analytics and Knowledge conference* (pp. 56-66). <https://doi.org/10.1145/3506860.3506866>
- Ng, D. T. K., Leung, J. K. L., Chu, S. K. W. y Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, 2, 100041. <https://doi.org/10.1016/j.caeai.2021.100041>
- Ng, D. T. K., Wu, W., Leung, J. K. L. y Chu, S. K. W. (2023). Artificial Intelligence (AI) literacy questionnaire with confirmatory factor analysis. En *2023 IEEE International Conference on Advanced Learning Technologies (ICALT)* (pp. 233-235). IEEE. <https://doi.org/10.1109/ICALT58122.2023.00074>
- Nunnally, J. C. (1978). *An overview of psychological measurement. Clinical diagnosis of mental disorders: A handbook* (pp. 97-146). https://doi.org/10.1007/978-1-4684-2490-4_4
- Ouyang, F., Zheng, L. y Jiao, P. (2022). Artificial intelligence in online higher education: A systematic review of empirical research from 2011 to 2020. *Education and Information Technologies*, 27(6), 7893-7925. <https://doi.org/10.1007/s10639-022-10925-9>
- Pérez, C. y Carretero-Dios, H. (2005). Normas para el desarrollo y revisión de estudios instrumentales. *International Journal of Clinical and Health Psychology*, 5(3), 521-551.
- Pérez, E. R. y Medrano, L. A. (2010). Análisis factorial exploratorio: bases conceptuales y metodológicas. *Revista Argentina de Ciencias del Comportamiento (RACC)*, 2(1), 58-66.
- Ruiz-Rojas, L. I., Acosta-Vargas, P., De-Moreta-Llovet, J. y Gonzalez-Rodriguez, M. (2023). Empowering education with generative artificial intelligence tools: Approach with an instructional design matrix. *Sustainability*, 15(15), 11524. <https://doi.org/10.3390/su151511524>
- Şahin Kölemen, C. (2024). Artificial intelligence technologies and ethics in educational processes: solution suggestions and results. *Innoeduca. International Journal of Technology and Educational Innovation*, 10(2), 1-18. <https://doi.org/10.24310/ijtei.102.2024.19806>
- Saz-Pérez, F., Mir, B. P. y Carrió, A. L. (2024). Validación y estructura factorial de un cuestionario TPACK en el contexto de Inteligencia Artificial Generativa (IAG). *Hachetetepe: Revista Científica de Educación y Comunicación*, (28), 1-14. <https://doi.org/10.25267/Hachetetepe.2024.i28.1101>
- Shaffer, D. R. y Kipp, K. (2010). *Developmental Psychology: Childhood and Adolescence*. Wadsworth.
- Soriano-Alcantara, J. M., Guillén-Gámez, F. D. y Ruiz-Palmero, J. (2024). Exploring Digital Competencies: Validation and Reliability of an Instrument for the Educational Community and for all Educational Stages. *Technology, Knowledge and Learning*, 1-20. <https://doi.org/10.1007/s10758-024-09741-6>
- Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications*. American Psychological Association. <https://doi.org/10.1037/10694-000>
- Tiwari, C. K., Bhat, M. A., Khan, S. T., Subramaniam, R. y Khan, M. A. I. (2024). What drives students toward ChatGPT?

- An investigation of the factors influencing adoption and usage of ChatGPT. *Interactive Technology and Smart Education*, 21(3), 333-355. <https://doi.org/10.1108/ITSE-04-2023-0061>
- Üzüm, B., Elçiçek, M. y Pesen, A. (2024). Development of Teachers' Perception Scale Regarding Artificial Intelligence Use in Education: Validity and Reliability Study. *International Journal of Human-Computer Interaction*, 1-12. <https://doi.org/10.1080/10447318.2024.2385518>
- Uzumcu, O. y Acilmis, H. (2024). Do innovative teachers use AI-powered tools more interactively? a study in the context of diffusion of innovation theory. *Technology, Knowledge and Learning*, 29(2), 1109-1128. <https://doi.org/10.1007/s10758-023-09687-1>
- Wang, F., Cheung, A. C., Chai, C. S. y Liu, J. (2024). Development and validation of the perceived interactivity of learner-AI interaction scale. *Education and Information Technologies*, 1-32. <https://doi.org/10.1007/s10639-024-12963-x>
- Wang, Y. M., Wei, C. L., Lin, H. H., Wang, S. C. y Wang, Y. S. (2022). What drives students' AI learning behavior: A perspective of AI anxiety. *Interactive Learning Environments*, 1-17. <https://doi.org/10.1080/10494820.2022.2153147>
- Wang, Y. Y. y Wang, Y. S. (2022). Development and validation of an artificial intelligence anxiety scale: An initial application in predicting motivated learning behavior. *Interactive Learning Environments*, 30(4), 619-634. <https://doi.org/10.1080/10494820.2019.1674887>
- Yilmaz, F. G. K., Yilmaz, R. y Ceylan, M. (2023). Generative artificial intelligence acceptance scale: A validity and reliability study. *International Journal of Human-Computer Interaction*, 1-13. <https://doi.org/10.1080/10447318.2023.2288730>
- Yu, H. y Guo, Y. (2023, June). Generative artificial intelligence empowers educational reform: current status, issues, and prospects. *Frontiers in Education*, 8, Article 1183162. <https://doi.org/10.3389/educ.2023.1183162>
- Zhan, Y., Qiu, Z., Li, X. y Zhao, Z. (2024). Ease of Use or Fun Perception? Factors Affecting Retention of Newly Registered Mobile Game Players Based on Flow Theory and The Technology Acceptance Model. *Journal of Internet Technology*, 25(4), 497-505. <https://doi.org/10.70003/160792642024072504001>

Fecha de recepción del artículo: 1 de diciembre de 2024

Fecha de aceptación del artículo: 18 de febrero de 2025

Fecha de aprobación para maquetación: 12 de marzo de 2025

Fecha de publicación en OnlineFirst: 3 de abril de 2025

Fecha de publicación: 1 de julio de 2025

Del optimismo a la confianza: el impacto de ChatGPT en la confianza de los estudiantes en el aprendizaje asistido por IA

From optimism to trust: how ChatGPT is reshaping student confidence in AI-driven learning



- Frederic Marimon - *Universidad Internacional de Cataluña, UIC (España)*
- María Belén Arias-Valle - *Universidad Católica de Cuyo, UCCuyo (Argentina)*
- César Javier Coria-Augusto - *Universidad Católica de Cuyo, UCCuyo (Argentina)*
- Claudio Marcelo Larrea Arnau - *Universidad Católica de Cuyo, UCCuyo (Argentina)*

RESUMEN

La inteligencia artificial generativa, como ChatGPT, está transformando el ámbito de la educación superior, especialmente en el apoyo a tareas académicas. Sin embargo, su adopción efectiva depende de factores como la confianza de los estudiantes y su percepción sobre la utilidad de la herramienta. Este estudio analiza cómo el optimismo tecnológico, la utilidad percibida y la facilidad de uso influyen en la confianza de los estudiantes universitarios en ChatGPT. Se empleó una metodología cuantitativa, utilizando un modelo de ecuaciones estructurales (SEM) basado en datos recopilados de 316 estudiantes universitarios en Argentina. La encuesta incluyó preguntas sobre optimismo, utilidad percibida, facilidad de uso y confianza en ChatGPT, y se realizaron análisis factoriales para validar los constructos y examinar las relaciones propuestas en el modelo. Los resultados muestran que el optimismo de los estudiantes impacta significativamente en su percepción de utilidad y confianza en ChatGPT, mientras que la facilidad de uso no tiene un efecto directo significativo sobre la confianza. Esto sugiere que los estudiantes valoran los beneficios prácticos que la herramienta aporta a su proceso de aprendizaje más que su facilidad de uso. Estos hallazgos sugieren que las universidades deberían centrarse en resaltar el valor práctico de ChatGPT mediante programas de capacitación específicos y en fomentar un uso crítico de la herramienta. También se recomienda implementar estrategias que mejoren la interacción entre estudiantes y docentes y evaluar el potencial de ChatGPT para mejorar el rendimiento académico de los estudiantes.

Palabras clave: inteligencia artificial; tecnología educativa; aprendizaje automático; innovación educativa; teorías de aprendizaje; confianza.

ABSTRACT

Generative artificial intelligence, such as ChatGPT, is transforming the field of higher education, especially in supporting academic tasks. However, its effective adoption depends on factors such as students' trust and their perception of the tool's usefulness. This study analyzes how technological optimism, perceived usefulness, and ease of use influence university students' trust in ChatGPT. A quantitative methodology was employed, using structural equation modeling (SEM) based on data collected from 316 university students in Argentina. The survey included questions about optimism, perceived usefulness, ease of use, and trust in ChatGPT, with factorial analyses conducted to validate the constructs and examine the proposed relationships in the model. The results show that students' optimism significantly impacts their perception of usefulness and trust in ChatGPT, while ease of use does not have a direct significant effect on trust. This suggests that students value the practical benefits the tool brings to their learning process more than its ease of use. These findings suggest that universities should focus on highlighting the practical value of ChatGPT through specific training programs and fostering a critical use of the tool. It is also recommended to implement strategies that enhance the interaction between students and teachers and assess the potential of ChatGPT to improve students' academic performance.

Keywords: artificial intelligence; educational technology; machine learning; educational innovation; learning theories; trust.

INTRODUCCIÓN

La introducción de la inteligencia artificial generativa, en particular ChatGPT, ha transformado significativamente la educación superior. Esta herramienta, reconocida por su facilidad de uso y accesibilidad, ha sido adoptada rápidamente por los estudiantes universitarios (García Peñalvo et al., 2024). Al igual que otras tecnologías emergentes, su adopción sigue un patrón de difusión en el que los primeros usuarios incorporan la innovación con mayor rapidez (Rogers, 1995).

Las investigaciones sugieren que quienes adoptan tempranamente herramientas de inteligencia artificial, incluido ChatGPT, tienden a presentar un mayor optimismo tecnológico y una disposición innovadora (Parasuraman y Colby, 2014; Parasuraman, 2000). Este optimismo influye en su percepción de utilidad, ya que identifican con rapidez los beneficios prácticos que la herramienta aporta a sus tareas académicas (Davis, 1989).

Sin embargo, la naturaleza generativa de ChatGPT introduce un grado de imprevisibilidad en sus respuestas, lo que puede ocasionar inexactitudes o “alucinaciones”. Esto resalta la necesidad de que los usuarios desarrollen criterios de discernimiento para evaluar su fiabilidad. La confianza en las herramientas de inteligencia artificial, entendida como la disposición a depender de sus resultados a pesar de la incertidumbre (Mayer et al., 1995), es un factor clave en su adopción. En el caso de los estudiantes, la interacción reiterada con ChatGPT les permite comprender sus fortalezas y limitaciones, lo que moldea su confianza en su aplicación académica.

El modelo teórico de este estudio se basa en el enfoque propuesto por Marimon et al. (2024), que analiza la relación entre la utilidad percibida, la facilidad de uso y la confianza en herramientas de inteligencia artificial generativa en el ámbito laboral. Este estudio adapta dicho modelo al contexto universitario, considerando que la utilidad percibida y la confianza en herramientas como ChatGPT son factores determinantes para su adopción por parte de los estudiantes. Al igual que en el entorno laboral, la confianza se identifica como un mediador esencial que puede influir en la disposición de los estudiantes a integrar estas herramientas en su proceso de aprendizaje.

El presente estudio tiene como objetivo explorar cómo el optimismo de los estudiantes hacia las nuevas tecnologías, en particular ChatGPT, influye en su percepción de utilidad y facilidad de uso. Asimismo, se investiga cómo estos factores impactan, a su vez, en la confianza depositada en ChatGPT. Adicionalmente, se examina la relación directa entre el optimismo y la confianza. A través del análisis de estas relaciones, este estudio busca aportar evidencia sobre el papel del optimismo en la construcción de la confianza en herramientas de inteligencia artificial, lo que resulta fundamental para su uso efectivo en entornos académicos.

Revisión de la literatura y modelo de investigación

ChatGPT se ha consolidado como una herramienta valiosa para los estudiantes universitarios, ya que facilita la generación de contenido, el desarrollo de ideas y la resolución de problemas (Tang et al., 2023; Zawacki-Richter et al., 2019). No obstante, su adopción depende de la confianza que los estudiantes depositen en la herramienta, la cual está determinada por la utilidad percibida y la facilidad de uso (Venkatesh et al., 2003).

La incorporación de ChatGPT en la educación superior ha transformado la manera en que los estudiantes interactúan con el conocimiento y gestionan sus tareas académicas (Forero-Corba y Negre Bennasar, 2024). Desde la redacción de ensayos hasta la realización de investigaciones, esta herramienta ofrece un apoyo flexible y accesible que complementa las metodologías de enseñanza tradicionales. Sin embargo, para que los estudiantes adopten de manera efectiva estas tecnologías, es fundamental comprender los factores que influyen en la confianza que desarrollan hacia ellas, así como su percepción sobre la utilidad y la facilidad de uso de ChatGPT (García Peñalvo et al., 2024).

Si bien investigaciones previas han explorado la confianza en la inteligencia artificial en entornos laborales (Marimon et al., 2024), aún existe un conocimiento limitado sobre su papel en el ámbito académico. Con el propósito de abordar esta brecha, el presente estudio examina cómo los estudiantes desarrollan confianza en ChatGPT, un factor clave para maximizar sus beneficios educativos.

Acemoglu et al. (2023) advierten sobre los riesgos del optimismo ciego hacia la inteligencia artificial, argumentando que, sin una perspectiva crítica y una implementación adecuada, la inteligencia artificial podría exacerbar las desigualdades sociales, al igual que lo han hecho revoluciones tecnológicas previas. Estos autores sostienen que, si bien el optimismo tecnológico puede acelerar la adopción de innovaciones, también puede generar una sobrevaloración de sus beneficios y una subestimación de sus limitaciones, especialmente en contextos donde la calidad de los resultados es un aspecto crítico, como en la educación. Esto sugiere la necesidad de mantener un equilibrio entre el optimismo y el pensamiento crítico respecto a la inteligencia artificial, evitando así una dependencia excesiva de estas tecnologías.

Acemoglu et al. (2023) enfatizan que, aunque las innovaciones tecnológicas tienen el potencial de aumentar la productividad y mejorar la calidad de vida, su impacto puede ser perjudicial si no se gestionan de manera equitativa. En esta misma línea, el ganador del Premio Nobel de Economía 2024 destacó la importancia de evaluar cuidadosamente las implicaciones socioeconómicas de la automatización y la inteligencia artificial. A pesar de que algunos estudios han demostrado que el optimismo hacia la inteligencia artificial influye en la confianza en su uso (Marimon et al., 2024), la perspectiva de Acemoglu subraya la importancia de que dicho optimismo vaya acompañado de una reflexión crítica, con el fin de evitar la perpetuación o el agravamiento de desigualdades estructurales.

El Modelo de Aceptación Tecnológica (TAM, por sus siglas en inglés) sostiene que la utilidad percibida y la facilidad de uso influyen en la adopción de tecnologías (Davis, 1989). En el contexto de la inteligencia artificial aplicada a la educación, diversos estudios han confirmado que la utilidad percibida es un predictor más sólido de la confianza y de la intención de uso que la facilidad de uso (Al-Abdullatif y Alsubaie, 2024; Huang et al., 2022).

Asimismo, la confianza en las herramientas de inteligencia artificial ha sido identificada como un factor clave para su adopción (Marimon et al., 2024; Mayer et al., 1995). La confianza se define como la disposición del usuario a ser vulnerable ante las acciones de la herramienta, esperando obtener resultados confiables y consistentes. Esto cobra especial relevancia en el caso de la inteligencia artificial generativa, cuyos resultados no siempre son deterministas y pueden presentar variaciones o “alucinaciones” (Bender et al., 2021). Al-kfairy et al. (2024) destacan que la confianza en la precisión y fiabilidad de los resultados generados por inteligencia artificial es un factor determinante para su uso continuado, especialmente en tareas que requieren altos niveles de rigor académico.

El presente estudio hace referencia al modelo propuesto por Marimon et al. (2024), adaptándolo al contexto de los estudiantes universitarios que utilizan herramientas de inteligencia artificial generativa como ChatGPT. Marimon et al. (2024) desarrollaron un modelo que analiza el impacto de la disposición hacia la adopción de tecnologías de inteligencia artificial generativa y cómo la confianza en estas herramientas influye en el compromiso y el desempeño en el entorno laboral. Su investigación enfatiza el papel fundamental de la confianza como mediador clave entre la experiencia del usuario y la implicación en el trabajo, demostrando que, si bien la facilidad de uso es importante, la utilidad percibida y la confianza en la tecnología son los factores esenciales para generar un impacto positivo.

Este estudio adapta el modelo de Marimon et al. (2024) al contexto de los estudiantes universitarios y su interacción con herramientas de inteligencia artificial generativa como ChatGPT. Aunque el modelo original de Marimon fue diseñado para analizar la influencia de la utilidad percibida, la facilidad de uso y la confianza en el ámbito laboral, su marco teórico ofrece una base sólida para explorar la adopción de tecnologías emergentes en el ámbito educativo. No obstante, las diferencias fundamentales entre los entornos profesionales y académicos justifican la necesidad de realizar ajustes específicos en el modelo.

En el contexto educativo, la disposición de los estudiantes para adoptar nuevas tecnologías puede estar influenciada por factores como el apoyo institucional, la formación digital previa y las experiencias personales con herramientas similares. A diferencia del entorno laboral, donde la facilidad de uso puede ser un factor crítico para la integración rápida en el flujo de trabajo (Marimon et al., 2024), en el ámbito académico los estudiantes tienden a valorar más la utilidad práctica de una herramienta. Esto se debe a que enfrentan tareas académicas específicas que

requieren resultados relevantes y precisos, más que una interfaz intuitiva (Venkatesh et al., 2003).

Además, el optimismo tecnológico de los estudiantes juega un papel determinante en la adopción de herramientas de inteligencia artificial, como lo sugiere la literatura sobre la adopción de innovaciones (Parasuraman y Colby, 2014). La utilidad percibida de ChatGPT, es decir, su capacidad para facilitar la generación de contenido académico y la resolución de problemas complejos influye directamente en la confianza que los estudiantes depositan en esta herramienta. Esta relación entre optimismo, utilidad y confianza ha sido adaptada para explorar cómo los estudiantes, en lugar de los empleados, integran ChatGPT en su proceso de aprendizaje.

Al adaptar el modelo de Marimon et al. (2024), este estudio plantea que la confianza en ChatGPT depende no solo de la facilidad de uso percibida, sino principalmente de su utilidad en contextos de aprendizaje específicos, como la redacción de ensayos y la investigación académica. Esto resalta la importancia de que los estudiantes perciban un valor tangible en el uso de la herramienta para que puedan confiar en sus capacidades y adoptarla de manera continua. La adaptación de este modelo busca capturar las particularidades del entorno educativo y los factores que facilitan la adopción de herramientas de inteligencia artificial generativa en la educación superior.

Este modelo de estudio se adapta para explorar cómo los estudiantes universitarios desarrollan confianza en ChatGPT, considerando los constructos de optimismo, utilidad percibida y facilidad de uso, similares a los propuestos por Marimon et al. (2024). Al igual que en el ámbito profesional, se espera que la disposición de los estudiantes hacia la tecnología y su percepción sobre la utilidad de ChatGPT influyan en el nivel de confianza depositado en la herramienta. Este enfoque ofrece una comprensión más integral de cómo la confianza en ChatGPT puede facilitar su integración en el entorno académico y cómo dicha confianza podría, eventualmente, mejorar la utilidad percibida de la herramienta en el proceso de aprendizaje.

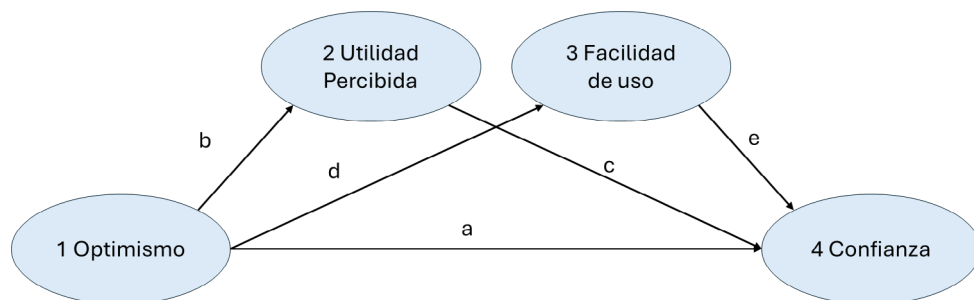
Sin embargo, existen diferencias clave entre los contextos laboral y educativo que justifican la adaptación del modelo de Marimon et al. (2024). En el ámbito educativo, la disposición de los estudiantes para aceptar y utilizar nuevas tecnologías puede estar influida por factores adicionales, como el apoyo institucional, la formación digital previa y las experiencias anteriores de aprendizaje con tecnologías similares. Además, estudios como el de Ly et al. (2024) sobre la adopción de plataformas digitales de aprendizaje en universidades sugieren que la motivación intrínseca desempeña un papel particularmente relevante en el contexto académico, ya que los estudiantes que perciben que una tecnología contribuye a su desarrollo académico personal están más dispuestos a integrarla en su rutina de estudio.

Por tanto, este estudio se diferencia de investigaciones previas al adaptar el enfoque de Marimon et al. (2024) a las particularidades del contexto universitario, incorporando el papel del optimismo de los estudiantes hacia la tecnología y explorando su efecto sobre la utilidad percibida y la confianza en ChatGPT. A través

de esta adaptación, el estudio busca arrojar luz sobre los mecanismos que facilitan la integración de herramientas de inteligencia artificial generativa en contextos educativos, contribuyendo a una comprensión más completa de los factores que influyen en la adopción de estas tecnologías por parte de los estudiantes.

En conclusión, el modelo de investigación propuesto se basa en la adaptación del enfoque de Marimon et al. (2024) al contexto de los estudiantes universitarios y su interacción con herramientas de inteligencia artificial generativa como ChatGPT. Este modelo integra los constructos de optimismo, utilidad percibida, facilidad de uso y confianza, con el objetivo de comprender cómo estas variables influyen en la adopción y aceptación de ChatGPT por parte de los estudiantes. La hipótesis central plantea que el optimismo de los estudiantes hacia la tecnología y su percepción sobre la utilidad de ChatGPT desempeñan un papel mediador clave en el desarrollo de confianza en la herramienta, lo que a su vez afecta su disposición a utilizarla de manera efectiva en entornos académicos. De esta manera, el modelo permite evaluar no solo la adopción de la herramienta, sino también cómo la confianza generada influye en su uso continuo y en la integración de la inteligencia artificial en los procesos de aprendizaje.

Figura 1
Modelo de investigación



Fuente: elaboración propia.

METODOLOGÍA

Esta sección describe en detalle los pasos seguidos en nuestra investigación, mientras que en la siguiente sección se presentarán, resumirán y comentarán los resultados.

Diseño del cuestionario

El cuestionario fue diseñado meticulosamente para abarcar diversos aspectos relevantes del estudio. Se inició con ítems que evalúan la disposición de los encuestados para adoptar nuevas tecnologías y continuó con preguntas dirigidas a medir el nivel de aceptación, la frecuencia de uso y la confianza en estas herramientas.

Para garantizar tanto la validez como la pertinencia del instrumento, los ítems de cada constructo fueron cuidadosamente seleccionados o adaptados de Marimon et al. (2024) y otras fuentes reconocidas y validadas dentro del campo de estudio. La Tabla 1 presenta un desglose detallado de los constructos y los ítems específicos incluidos en la investigación. Desde la introducción de la inteligencia artificial generativa, en particular ChatGPT en 2023, el panorama de la educación superior ha experimentado una transformación profunda.

El cuestionario fue diseñado para evaluar los constructos clave de este estudio, alineándose con el Modelo de Aceptación Tecnológica (TAM) (Davis, 1989). En primer lugar, se midió el optimismo tecnológico de los participantes y su percepción sobre la utilidad de ChatGPT, dado que investigaciones previas han demostrado que la utilidad percibida influye significativamente en la confianza y en la intención de uso de herramientas de inteligencia artificial (Al-Abdullatif y Alsubaie, 2024; Huang et al., 2022).

Asimismo, el cuestionario incluyó ítems para evaluar la facilidad de uso percibida y la confianza en herramientas de inteligencia artificial, utilizando escalas previamente validadas en estudios anteriores (Marimon et al., 2024). Se realizó un análisis factorial para confirmar la validez de los constructos, garantizando que cada ítem capturara con precisión la dimensión correspondiente dentro del marco teórico.

ChatGPT, una herramienta de propósito general reconocida por su facilidad de aprendizaje y accesibilidad, ha ganado rápidamente aceptación en la educación superior, proporcionando apoyo a los estudiantes en tareas de redacción académica, investigación y resolución de problemas. En este contexto, el diseño del cuestionario se orientó a comprender cómo los estudiantes perciben la utilidad y confiabilidad de esta tecnología, así como los factores que influyen en su adopción.

Tabla 1
Ítems propuestos para el cuestionario

Optimismo	OPT1	ChatGPT contribuye a mejorar mi aprendizaje en la universidad.
	OPT2	ChatGPT me proporciona mayor libertad y flexibilidad en mis estudios.
	OPT3	ChatGPT me permite tener más control sobre mis tareas académicas.

Innovación Personal	INN1	Mis compañeros suelen pedirme consejos sobre el uso de ChatGPT.
	INN2	Generalmente, soy de los primeros en mi grupo de amigos en probar nuevas tecnologías como ChatGPT.
	INN3	Puedo aprender a usar nuevas herramientas como ChatGPT sin ayuda de otros.
	INN4	Me mantengo actualizado con los últimos desarrollos de ChatGPT y tecnologías similares en mis áreas de interés.
Utilidad percibida	USE1	Considero que ChatGPT es útil para mis tareas académicas.
	USE2	Usar ChatGPT facilita la realización de mis tareas.
	USE3	Utilizar ChatGPT me permite completar mi trabajo más rápido.
Facilidad de uso	EAS1	Encuentro que ChatGPT es fácil de usar.
	EAS2	Aprender a usar ChatGPT me resultó sencillo.
	EAS3	Me resulta fácil lograr que ChatGPT haga lo que necesito.
Confianza en ChatGPT	TRU1	Me siento cómodo con la información que ChatGPT me proporciona para mis estudios.
	TRU2	Confío en ChatGPT para ayudarme con mis tareas académicas.
	TRU3	ChatGPT es una herramienta útil para mis actividades universitarias.
	TRU4	Si enfrento un problema desafiante en mis estudios, recorro al ChatGPT.
	TRU5	Me siento seguro respecto a la protección de datos con las herramientas de ChatGPT.
	TRU6	Siento que estoy adecuadamente protegido contra problemas al utilizar ChatGPT en mis estudios.
	TRU7	Confío en que las herramientas de ChatGPT cumplen con los estándares legales establecidos en el ámbito académico.

Fuente: elaboración propia.

Administración de la encuesta y perfil demográfico de los participantes

La encuesta fue distribuida y completada en octubre de 2024, obteniéndose un total de 316 respuestas completas. Los encuestados fueron estudiantes argentinos con experiencia en el uso de herramientas de inteligencia artificial generativa (IA-Gen) en sus actividades académicas. La mayoría de los participantes fueron mujeres (60,8 %) y, en su mayoría, adultos jóvenes, con un 67,4 % en el rango etario de entre 18 y 25 años. En cuanto al nivel educativo, el 52,8 % había completado la educación secundaria, mientras que el 25,6 % poseía un título universitario. La muestra incluyó estudiantes de universidades tanto públicas (26,6 %) como privadas (73,4 %), siendo la mayoría asistentes a clases con una matrícula de entre 10 y 50 estudiantes.

Respecto al uso de herramientas de IA generativa, el 35,4 % indicó utilizarlas de forma esporádica, mientras que un 10,4 % reportó un uso diario. Este estudio ofrece información relevante sobre la confianza de los estudiantes en la inteligencia artificial en el contexto del sistema de educación superior argentino, aunque los resultados podrían variar en otros entornos culturales e institucionales.

Tabla 2

Características demográficas de la muestra

	Número	%
Género		
Hombre	124	39.2
Mujer	192	60.8
Total	316	100.0
Edad		
Entre 18 y 20 años	81	25.6
Entre 21 y 25 años	132	41.8
Entre 26 y 35 años	47	14.9
Entre 36 y 45 años	34	10.8
Entre 46 y 55 años	20	6.3
> 55	2	0.6
Total	81	25.6
Nivel Educativo		
Educación Secundaria	167	52.8
Educación técnica o profesional	49	15.5
Título Universitario	81	25.6
Posgrado (Maestría Doctorado)	10	3.2
Prefiere no responder	9	2.8
Total	316	100.0
Frecuencia de uso de IA generativa		
No utilizo inteligencia artificial	68	21.5
Muy raramente	112	35.4
Una vez por semana	52	16.5
Una vez al día	51	16.1
Intensivamente, todos los días	33	10.4
Total	316	100.0
Universidad		
Universidad Católica de Cuyo	210	66.5

	Número	%
Universidad Nacional de San Juan	74	23.4
Universidad de Congreso	18	5.7
Otras	14	4.4
Total	316	100.0
Titularidad de la Universidad		
Pública	84	26.6
Privada	232	73.4
Total	316	100.0
Tamaño de la Clase		
Menos de 10 estudiantes	18	5.7
Entre 10 y 50 estudiantes	209	66.1
Entre 50 y 100 estudiantes	66	20.9
Más de 100 estudiantes	23	7.3
Total	316	100.0

Fuente: elaboración propia.

Evaluación del modelo de investigación

Se llevaron a cabo tres Análisis Factoriales Exploratorios (EFA) independientes utilizando el método de análisis de componentes principales con rotación varimax. Este procedimiento permitió identificar y refinar los ítems que representaban con mayor precisión cada constructo del modelo de investigación. Con este enfoque, se garantizó que los constructos fueran estadísticamente sólidos y estuvieran alineados con los objetivos del estudio.

El primer EFA incluyó ítems correspondientes a las dimensiones de “Optimismo” e “Innovación”, adaptadas de la escala TRI. El segundo EFA se aplicó a los ítems adaptados del Modelo de Aceptación Tecnológica (TAM), centrándose en las dimensiones de “Utilidad Percibida” y “Facilidad de Uso”. El tercer y último EFA se enfocó en los ítems seleccionados para la dimensión “Confianza”.

Una vez completados los análisis exploratorios, el modelo de investigación, representado en la Figura 1, fue probado mediante Modelado de Ecuaciones Estructurales (SEM). Para ello, se empleó el método de máxima verosimilitud robusta basado en la matriz de varianza-covarianza asintótica, con la asistencia del software EQS. Se evaluaron rigurosamente las propiedades psicométricas de cada constructo, incluyendo la fiabilidad (medida mediante el alfa de Cronbach y la fiabilidad compuesta (CR)) y la validez convergente (evaluada a través de la varianza extraída media (AVE)). Además, se realizó un análisis discriminante para garantizar

la independencia de los constructos, comparando la raíz cuadrada del AVE de cada dimensión con sus correlaciones con otros constructos.

El ajuste general del modelo fue evaluado utilizando distintos índices de bondad de ajuste, incluyendo el chi-cuadrado de Bentler-Satorra, su coeficiente y grados de libertad, así como otros índices como el Índice de Ajuste Comparativo (CFI) y el Error Cuadrático Medio de Aproximación (RMSEA). Una vez confirmado el ajuste del modelo, se analizaron e interpretaron los coeficientes estandarizados para obtener información detallada sobre las relaciones entre los constructos.

La validación del cuestionario se llevó a cabo antes de la recolección de datos, siguiendo las pautas metodológicas establecidas por Marimon et al. (2024). El instrumento fue desarrollado mediante la adaptación de ítems validados en estudios previos, asegurando su alineación conceptual con los constructos en análisis. La selección y refinamiento de los ítems se realizó con el juicio de expertos, a través de un panel de especialistas en tecnología educativa y psicometría, quienes revisaron el cuestionario para evaluar su claridad, pertinencia y validez de constructo. Además, se llevó a cabo una prueba piloto preliminar con un grupo reducido de estudiantes universitarios para evaluar la comprensión de los ítems y la coherencia en las respuestas.

El cuestionario fue administrado en español, ya que la población objetivo estaba compuesta por estudiantes universitarios de Argentina. Para asegurar la equivalencia semántica y conceptual de los ítems, se realizó un proceso de validación lingüística, adaptando los términos al contexto local. Se empleó una escala Likert de 5 puntos, que oscilaba entre 1 ("Totalmente en desacuerdo") y 5 ("Totalmente de acuerdo"), permitiendo una medición precisa de las percepciones de los participantes sobre las variables del estudio. La administración del cuestionario se realizó en línea a través de una plataforma de encuestas, asegurando accesibilidad y eficiencia en la recolección de datos.

La muestra del estudio estuvo compuesta por 316 estudiantes universitarios argentinos, representando una diversidad de trayectorias académicas y experiencias previas con herramientas de IA. En términos de nivel educativo, el 52.8 % de los participantes había completado la educación secundaria, el 25.6 % poseía un título universitario, y el 15.5 % contaba con formación técnica o profesional. En relación con la familiaridad con la IA, el 34.2 % reportó experiencia mínima previa, el 42.7 % indicó exposición moderada, y el 23.1 % mencionó un uso frecuente en contextos académicos. La muestra incluyó estudiantes de universidades públicas (26.6 %) y privadas (73.4 %), lo que aseguró una representación heterogénea del panorama de la educación superior en Argentina. Estos datos demográficos proporcionan un marco de referencia integral para comprender la composición de la muestra y contextualizar los hallazgos del estudio dentro de la adopción de la inteligencia artificial en la educación superior.

Para garantizar la transparencia metodológica y el rigor analítico, se emplearon técnicas estadísticas tanto descriptivas como inferenciales. Se realizó un análisis de varianza (ANOVA) para identificar diferencias significativas entre los grupos en

relación con sus percepciones sobre la utilidad percibida, la facilidad de uso y la confianza en ChatGPT. La elección del ANOVA se justificó por su capacidad para comparar diferencias de medias entre múltiples grupos independientes, controlando la varianza dentro de la muestra. Además, se generaron box plots para visualizar la distribución de las respuestas de las variables clave, permitiendo identificar posibles valores atípicos y patrones de variabilidad en los datos. Estas técnicas permitieron realizar un análisis sólido de las relaciones entre los constructos del estudio.

Se realizó un análisis adicional para examinar el papel de variables no determinantes, como el número de estudiantes en clase o el uso del correo electrónico. Si bien estas variables no mostraron efectos directos significativos, se llevaron a cabo análisis exploratorios adicionales para evaluar posibles influencias indirectas. Los resultados sugieren que su falta de impacto podría atribuirse a factores contextuales, como diferencias institucionales o niveles variables de adopción tecnológica entre los encuestados. Además, limitaciones en la medición podrían haber reducido su significancia estadística. No obstante, la exclusión de estas variables no comprometió la robustez del modelo, lo cual fue confirmado mediante validaciones estadísticas complementarias. Esta discusión se integró en la sección de resultados para proporcionar una interpretación más matizada de los hallazgos.

Todos los análisis estadísticos fueron realizados utilizando SPSS versión 27 y el software R (versión 4.2.1). El procesamiento de datos incluyó estadística descriptiva, análisis de fiabilidad (alfa de Cronbach) y pruebas inferenciales para evaluar las relaciones entre las variables. Se utilizó ANOVA para comparar diferencias de medias entre grupos, mientras que el Análisis Factorial Confirmatorio (CFA) fue realizado con el paquete lavaan en R para validar la estructura del modelo de medición. Además, se generaron box plots para visualizar la distribución de los datos y detectar posibles valores atípicos. El enfoque metodológico siguió las mejores prácticas en análisis estadístico, asegurando la replicabilidad de los procedimientos utilizados. Se detallaron todos los pasos de preprocesamiento de datos, codificación y transformaciones para facilitar su reproducibilidad en futuras investigaciones.

Estos resultados respaldan la validez y fiabilidad del modelo de investigación, confirmando que la confianza en ChatGPT es un factor clave en su adopción por parte de los estudiantes universitarios. La aplicación rigurosa de estos métodos analíticos garantiza la solidez de los hallazgos y su contribución a la comprensión de la integración de herramientas de inteligencia artificial en la educación superior.

Doble mediación del rol de la “Utilidad percibida” y la “Facilidad de uso”

Un aspecto crítico de este estudio, como se ilustra en la Figura 1, es el papel mediador de la “Utilidad percibida” y la “Facilidad de uso” en la relación entre el “Optimismo tecnológico” y la “Confianza en ChatGPT”. Esta sección se dedica a un análisis detallado de este efecto de mediación, el cual representa el objetivo secundario del estudio.

La metodología utilizada para evaluar el papel mediador de ambos constructos se basó en los trabajos seminales de Baron y Kenny (1986), Hayes (2009) y Zhao et al. (2010). Estos estudios proporcionan un marco teórico sólido para comprender los mecanismos a través de los cuales la “Utilidad percibida” y la “Facilidad de uso” facilitan la transición del optimismo tecnológico hacia la confianza en la herramienta.

RESULTADOS

La estructura delineada en la sección anterior se siguió para presentar los principales resultados empíricos basados en los datos recopilados. Antes de proceder con el análisis, se realizó la prueba de un solo factor de Harman para verificar la posible presencia de sesgo por método común (CMB, por sus siglas en inglés). Se llevó a cabo un Análisis Factorial Exploratorio (EFA) con los 46 ítems del cuestionario, identificándose ocho factores con valores propios mayores a uno, siendo el primer factor responsable de solo 17.92 % de la varianza total. Este resultado sugiere que el sesgo por método común no representó un problema significativo en este estudio.

Como se mencionó anteriormente, el análisis comenzó con tres EFAs. Los criterios utilizados para la retención de ítems siguieron las directrices de Ladhari (2012) y Wolfenbarger y Gilly (2003), que establecen que: (i) los ítems deben presentar cargas factoriales de 0.7 o superiores en un solo factor, (ii) los ítems no deben cargar más de 0.50 en dos factores diferentes, y (iii) los ítems deben tener una correlación ítem-total superior a 0.50.

El primer EFA, centrado en las dimensiones “Optimismo” e “Innovación” de la escala TRI, confirmó la distinción entre ambas dimensiones originales. Se calculó el alfa de Cronbach para ambos constructos, obteniéndose un valor superior a 0.7 para “Optimismo”, lo que confirmó su fiabilidad. Sin embargo, “Innovación” no alcanzó este umbral, por lo que fue excluido del análisis posterior.

El segundo EFA, que examinó conjuntamente las dimensiones “Utilidad percibida” y “Facilidad de uso”, reveló que dos ítems en cada dimensión presentaban cargas factoriales superiores a 0.7. Ambas dimensiones obtuvieron valores favorables en el alfa de Cronbach, lo que respaldó su fiabilidad y permitió su inclusión en el modelo final.

El tercer y último EFA, aplicado a los siete ítems originales de “Confianza”, retuvo seis ítems, con un alfa de Cronbach superior a 0.7, asegurando su validez y fiabilidad dentro del modelo.

La Tabla 3 presenta el análisis de fiabilidad de los cuatro constructos incluidos en el modelo final. La consistencia interna fue confirmada mediante el alfa de Cronbach y la fiabilidad compuesta (CR), cuyos valores superaron el umbral recomendado de 0.7 (Hair Jr et al., 2010). Adicionalmente, la varianza extraída promedio (AVE) superó el punto de corte de 0.5 (Nunnally y Bernstein, 1994), validando la validez convergente de los constructos.

Tabla 3
Cargas factoriales de los cuatro constructos y análisis de fiabilidad

	1		2		3		4	
	Optimismo		Utilidad percibida		Facilidad de uso		Confianza en Chat	
	OPT2	0.746	USE3	0.793	EAS2	0.789	TRU6	0.802
	OPT1	0.724	USE3	0.771	EAS3	0.757	TRU5	0.784
							TRU2	0.777
							TRU1	0.776
							TRU7	0.705
Alpha Cronbach	0.825		0.814		0.690		0.861	
Fiabilidad Compuesta	0.702		0.759		0.748		0.879	
Varianza Extraída Promedio	0.540		0.612		0.598		0.592	

Fuente: elaboración propia.

La Tabla 4 presenta los resultados del análisis de validez discriminante, el cual se llevó a cabo utilizando correlaciones lineales o covarianzas estandarizadas entre los factores latentes. Para evaluar la validez discriminante, se verificó que las correlaciones entre los constructos fueran menores que la raíz cuadrada de la AVE de cada uno de ellos (Fornell y Larcker, 1981).

La Tabla 4 muestra que la raíz cuadrada de la AVE de cada constructo es mayor que los elementos fuera de la diagonal, lo que respalda la validez discriminante. Esto confirma que cada constructo mide un concepto distinto y no existe colinealidad excesiva entre los factores.

Tabla 4
Matriz de correlaciones entre factores latentes

	1	2	3	4
1 Optimismo	0.735			
2 Utilidad Percibida	0.514	0.782		
3 Facilidad de uso	0.425	0.470	0.773	
4 Confianza	0.591	0.633	0.469	0.770

Fuente: elaboración propia.

Una vez seleccionados los ítems que se utilizarían en el modelo completo, se estimó el modelo mediante Modelado de Ecuaciones Estructurales (SEM) basado en covarianzas. Los índices de ajuste obtenidos en la estimación del modelo indicaron que las variables convergieron adecuadamente hacia los factores establecidos en el modelo de investigación.

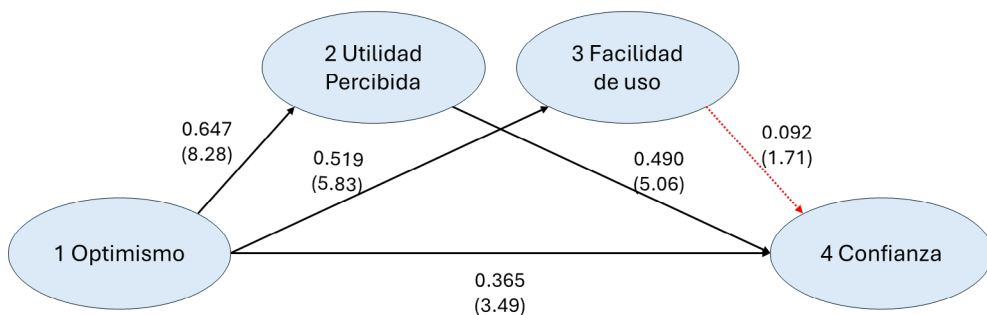
El chi-cuadrado de Satorra–Bentler (χ^2) fue de 182.5, con 49 grados de libertad y un p-valor de 0.000; el χ^2/df resultó ser 3.72, valor dentro del límite aceptable de 5. Además, el error cuadrático medio de aproximación (RMSEA) fue de 0.093, mientras que el índice de ajuste comparativo (CFI) alcanzó un valor de 0.924. Considerando la sensibilidad de la estadística robusta χ^2 y observando los indicadores globales, el ajuste general del modelo fue aceptable (Hair Jr et al., 2010).

La Figura 2 presenta el modelo de investigación con los coeficientes estandarizados y sus valores *t* asociados entre paréntesis. Todos los coeficientes fueron significativos al nivel de 0.05, excepto la relación entre “Facilidad de uso” y “Confianza”. Por lo tanto, tres hipótesis fueron aceptadas, mientras que la cuarta fue rechazada.

Este hallazgo es particularmente relevante y requiere un análisis detallado. Los resultados no demostraron que la “Facilidad de uso” tenga un impacto directo sobre la “Confianza”, lo que sugiere que los estudiantes universitarios confían en ChatGPT principalmente en función de su utilidad percibida, más que por la facilidad con la que pueden utilizar la herramienta.

Figura 2

*Modelo de investigación con coeficientes estandarizados y valores *t* asociados*



Fuente: elaboración propia.

Los resultados de este estudio ofrecen hallazgos significativos sobre cómo los estudiantes universitarios desarrollan confianza en herramientas de inteligencia artificial generativa como ChatGPT. En primer lugar, el análisis revela que el “Optimismo tecnológico” tiene un impacto directo y significativo en la “Confianza”, así como en la “Utilidad percibida” y la “Facilidad de uso”. Estos efectos son igualmente fuertes, lo que sugiere que los estudiantes con una actitud más optimista hacia

las nuevas tecnologías no solo encuentran que ChatGPT es más fácil de usar, sino que también lo perciben como una herramienta más útil en sus tareas académicas. Sin embargo, surge una distinción clave al examinar las rutas que conducen a la confianza: mientras que la “Utilidad percibida” tiene un impacto significativo en la “Confianza”, la “Facilidad de uso” no lo tiene.

La Tabla 5 presenta la descomposición de los efectos de mediación de “Utilidad percibida” y “Facilidad de uso” en la relación entre “Optimismo” y “Confianza”, lo que constituye un elemento central del modelo de investigación. De manera notable, el efecto directo del optimismo sobre la confianza es tan fuerte como el efecto indirecto total (a través de “Utilidad percibida” y “Facilidad de uso”), con ambos mostrando un coeficiente estandarizado idéntico de 0.365. Esta coincidencia resalta la importancia tanto de las rutas directas como de las mediadas en la construcción de la confianza. No obstante, un análisis más profundo de los efectos indirectos revela que la “Utilidad percibida” juega un papel crucial en esta mediación, mientras que la “Facilidad de uso” contribuye poco al desarrollo de la confianza.

Las implicaciones son claras: para fomentar la confianza en ChatGPT, la percepción de utilidad de la herramienta es esencial. Si los estudiantes consideran que la herramienta es fácil de usar o no, resulta menos relevante en este sentido. Esto sugiere que la curva de aprendizaje para el uso de ChatGPT no representa una barrera significativa para el desarrollo de la confianza; una vez que los estudiantes dominan su funcionalidad básica, su confianza depende principalmente de la capacidad de la herramienta para ofrecer resultados valiosos y confiables.

Según Zhao et al. (2010), este hallazgo indica que la mediación a través de la “Utilidad percibida” es complementaria, mientras que la “Facilidad de uso” no actúa como mediador. En otras palabras, la percepción de utilidad de los resultados generados por ChatGPT permite el fortalecimiento de la confianza, independientemente de lo fácil que sea su uso. Esto subraya el papel crucial que desempeñan los beneficios prácticos de la inteligencia artificial generativa en la formación de la confianza de los estudiantes, sugiriendo que los desarrolladores y educadores deberían centrarse en demostrar el valor tangible de estas herramientas en lugar de enfatizar únicamente su facilidad de uso.

Tabla 5
Matriz de correlaciones entre factores latentes

	Efecto Total	Efecto Indirecto Partial	Efecto Indirecto Total	Efecto Directo
Optimismo → Confianza	0.730 (3.89)	(b*c) 0.317 (d*e) 0.048	0.365 (3.91)	(a) 0.365 (3.49)

Fuente: elaboración propia.

Cada celda muestra el valor estandarizado y, entre paréntesis, el p-valor. Las letras a, b, c y d corresponden a la notación en la Figura 1.

DISCUSIÓN

Los hallazgos de este estudio brindan una comprensión más profunda sobre cómo los estudiantes universitarios desarrollan confianza en las herramientas de inteligencia artificial generativa, específicamente en ChatGPT. En consonancia con el modelo adaptado de Marimon et al. (2024), se encontró que el optimismo tecnológico de los estudiantes tiene un impacto significativo tanto en la utilidad percibida como en la confianza en ChatGPT. Sin embargo, el análisis también reveló una diferencia clave en el contexto académico: la “Facilidad de uso” de ChatGPT no tiene un impacto directo significativo en la “Confianza”, a diferencia de lo observado en estudios centrados en el ámbito laboral.

Este hallazgo sugiere que, en el entorno académico, los estudiantes tienden a priorizar la utilidad práctica percibida de la herramienta sobre su facilidad de uso. Una posible explicación es que los estudiantes universitarios, al enfrentar demandas académicas específicas, valoran más los beneficios tangibles de ChatGPT, como su capacidad para generar contenido relevante y preciso, en lugar de la simplicidad de su interfaz. Esto contrasta con el contexto laboral analizado por Marimon et al. (2024), donde la facilidad de uso desempeñó un papel más destacado en la confianza de los empleados, posiblemente debido a la necesidad de una integración rápida y sin fricciones en los flujos de trabajo.

La relación entre “Optimismo” y “Confianza” fue particularmente sólida, lo que indica que aquellos estudiantes con una actitud positiva hacia la adopción de nuevas tecnologías están más inclinados a confiar en ChatGPT. Este resultado se alinea con la literatura sobre adopción tecnológica, que sugiere que el optimismo hacia las innovaciones tecnológicas puede ser un fuerte predictor de la disposición a aceptar nuevas herramientas (Parasuraman y Colby, 2014). Sin embargo, este estudio amplía dicha comprensión al demostrar que, en el contexto de la educación superior, la utilidad percibida actúa como un mediador crucial entre el optimismo y la confianza.

Los hallazgos de este estudio tienen implicaciones prácticas significativas para el diseño de programas de formación tecnológica en universidades. Dado que la utilidad percibida de ChatGPT influye directamente en la confianza de los estudiantes, las universidades deberían centrarse en desarrollar programas que no solo enseñen el uso básico de la herramienta, sino que también resalten sus beneficios prácticos en el ámbito académico.

Por ejemplo, podrían implementarse talleres específicos para demostrar cómo utilizar ChatGPT en la redacción de ensayos, la elaboración de revisiones de literatura y la generación de ideas innovadoras para proyectos de investigación. Estas

actividades permitirían a los estudiantes experimentar de forma directa el valor de la herramienta en diferentes áreas de su formación académica.

Además, la integración de módulos de formación dentro de cursos existentes podría ser una estrategia efectiva para fomentar el uso continuo de ChatGPT. En cursos de metodología de la investigación, por ejemplo, ChatGPT podría emplearse para guiar a los estudiantes en la estructuración y organización de sus trabajos. En cursos de ciencias sociales, la herramienta podría utilizarse para simular debates sobre temas complejos.

Este enfoque práctico permitiría a los estudiantes comprender cómo ChatGPT puede ser una herramienta útil y relevante en su proceso de aprendizaje, aumentando así su predisposición a adoptarlo como parte de su educación universitaria.

Promoviendo un Uso Crítico de la Inteligencia Artificial

La adopción de ChatGPT debe ir acompañada de esfuerzos por parte de las universidades para promover un uso crítico de la inteligencia artificial generativa. Los resultados muestran que los estudiantes con una actitud optimista hacia la tecnología tienden a confiar más en ChatGPT, lo cual es positivo para su adopción. Sin embargo, es fundamental que esta confianza no se traduzca en un uso acrítico de la herramienta. Los educadores deben enfatizar que, si bien ChatGPT puede ser un aliado en el proceso de aprendizaje, no es infalible y puede generar información inexacta o sesgada.

Para fomentar un uso crítico, los programas de formación podrían incluir actividades que incentiven a los estudiantes a comparar la información generada por ChatGPT con fuentes académicas tradicionales. Esto les permitiría evaluar la calidad de las respuestas de ChatGPT y discernir cuándo es más apropiado utilizar la herramienta. Además, los docentes podrían diseñar tareas que requieran a los estudiantes justificar y analizar los resultados obtenidos con ChatGPT, promoviendo así un enfoque reflexivo y crítico en su uso.

Desarrollando Competencias Digitales y Éticas

La integración de la inteligencia artificial en la educación superior debe ir acompañada de un fortalecimiento de las competencias digitales y éticas de los estudiantes. ChatGPT, al igual que otras herramientas de inteligencia artificial, opera a través de algoritmos complejos cuyos procesos no siempre son transparentes para el usuario. Por ello, es fundamental que los estudiantes comprendan el funcionamiento básico de estas tecnologías y los posibles sesgos que pueden influir en sus resultados. Las universidades deberían desarrollar módulos de formación que aborden cuestiones éticas como la privacidad de los datos, la propiedad intelectual y el impacto social de la inteligencia artificial.

Los resultados de este estudio subrayan la importancia de que los estudiantes perciban la utilidad de ChatGPT, ya que aquellos que lo hacen tienen una mayor predisposición a integrarlo en su proceso de aprendizaje. Sin embargo, esta integración debe ir de la mano de una comprensión crítica de las limitaciones de la tecnología, permitiendo que los estudiantes aprovechen sus beneficios de manera consciente y responsable.

Este estudio sugiere que las universidades deberían considerar la utilidad percibida de las herramientas de inteligencia artificial generativa como un factor clave en el diseño de sus planes de estudio. La integración de tecnologías como ChatGPT no debe verse como un fin en sí mismo, sino como un medio para enriquecer la experiencia de aprendizaje y desarrollar habilidades críticas en los estudiantes. Las universidades que lideren la adopción de estas herramientas tienen la oportunidad de crear programas interdisciplinarios que conecten la inteligencia artificial con diversas áreas del conocimiento.

Por ejemplo, los departamentos de ciencias de la computación podrían colaborar con las facultades de ciencias sociales y humanidades para diseñar cursos en los que se enseñe a los estudiantes a aplicar herramientas de IA generativa en el análisis de datos, la interpretación de textos y la creación de proyectos innovadores. Esta colaboración interdisciplinaria no solo fortalecería el uso de la inteligencia artificial en la educación superior, sino que también prepararía a los estudiantes para un entorno profesional cada vez más digitalizado y competitivo.

En términos generales, estos hallazgos destacan la importancia de desarrollar un enfoque centrado en los beneficios tangibles que ChatGPT y otras herramientas de IA pueden ofrecer a los estudiantes universitarios. Al demostrar cómo estas tecnologías pueden apoyar y mejorar el aprendizaje, las universidades pueden promover una adopción más amplia y efectiva de la inteligencia artificial generativa en sus programas. Esto no solo facilitará la integración de la IA en el entorno académico, sino que también empoderará a los estudiantes para que se conviertan en usuarios críticos e informados, preparados para enfrentar los desafíos tecnológicos del futuro.

Principales hallazgos e implicaciones prácticas

La confianza en ChatGPT emerge como un factor determinante para su adopción entre los estudiantes universitarios. El estudio confirma que la disposición de los estudiantes a integrar herramientas de inteligencia artificial en sus actividades académicas está influida significativamente por su percepción sobre la utilidad de dichas herramientas. Este hallazgo refuerza la vigencia de los modelos de aceptación tecnológica, al subrayar la importancia de generar confianza en la inteligencia artificial para fomentar su uso efectivo en el ámbito educativo.

La investigación también pone de manifiesto que la utilidad percibida desempeña un papel más relevante en la adopción de la IA que la facilidad de uso. Si bien una interfaz intuitiva puede facilitar el acercamiento inicial, los estudiantes priorizan

los beneficios prácticos que ChatGPT ofrece para sus procesos de aprendizaje. En consecuencia, las instituciones de educación superior deberían enfocarse en fortalecer las competencias digitales del estudiantado mediante programas de formación estructurados que promuevan una integración significativa y efectiva de la inteligencia artificial en los entornos académicos.

Para garantizar una adopción responsable de la IA, las universidades deben desarrollar políticas claras que respeten los principios éticos y resguarden la integridad académica. Estas políticas deben incluir orientaciones sobre el uso responsable de la IA, al mismo tiempo que fomenten que los docentes incorporen tecnologías basadas en inteligencia artificial en sus prácticas de enseñanza. Al promover un enfoque crítico e informado, los educadores pueden ayudar a los estudiantes a comprender tanto los beneficios como las limitaciones de las herramientas de IA generativa, asegurando así su uso responsable y eficaz en la educación superior.

CONCLUSIONES

Los hallazgos de este estudio tienen el potencial de redefinir cómo las universidades diseñan e implementan políticas de integración de inteligencia artificial generativa, como ChatGPT, en sus programas académicos. Al demostrar que la utilidad percibida de ChatGPT tiene un impacto más significativo en la confianza de los estudiantes que la facilidad de uso, este estudio sugiere que las instituciones de educación superior deberían centrarse en resaltar el valor práctico de estas herramientas para el aprendizaje, en lugar de limitarse a facilitar su uso. Esto implica que las políticas de integración de IA deben priorizar la creación de espacios de formación donde los beneficios concretos de estas tecnologías para el desarrollo de habilidades académicas —como la redacción de ensayos, la investigación y la resolución de problemas complejos— sean evidentes.

Las universidades podrían beneficiarse de la implementación de políticas que promuevan la incorporación de ChatGPT en actividades curriculares específicas, como talleres, módulos de cursos y proyectos interdisciplinarios, con el objetivo de reforzar la percepción de utilidad de la herramienta entre los estudiantes. Además, estas políticas deberían fomentar el desarrollo de competencias críticas y éticas, asegurando que los estudiantes no solo adopten la tecnología, sino que también se conviertan en usuarios reflexivos y responsables.

Los resultados de este estudio proporcionan información valiosa sobre cómo los estudiantes universitarios desarrollan confianza en herramientas de IA generativa como ChatGPT. En primer lugar, el análisis revela que el “Optimismo tecnológico” tiene un impacto directo y significativo en la “Confianza”, así como en la “Utilidad percibida” y la “Facilidad de uso”. Estos efectos son igualmente fuertes, lo que sugiere que los estudiantes que son más optimistas respecto a las nuevas tecnologías no solo encuentran ChatGPT más fácil de usar, sino que también lo perciben como una herramienta más útil para sus tareas académicas.

Sin embargo, surge una distinción clave al examinar los factores que conducen a la confianza: mientras que la “Utilidad percibida” tiene un impacto significativo en la confianza, la “Facilidad de uso” no. Esto sugiere que la curva de aprendizaje de ChatGPT no representa una barrera importante para el desarrollo de la confianza; una vez que los estudiantes dominan su funcionalidad básica, su confianza depende principalmente de la capacidad de la herramienta para proporcionar resultados valiosos y confiables.

Los resultados de este estudio abren nuevas líneas de investigación que podrían ampliar la comprensión del impacto de la IA generativa en el ámbito académico. En primer lugar, sería valioso explorar cómo el uso de ChatGPT afecta la interacción entre estudiantes y docentes, analizando si esta herramienta facilita una comunicación más efectiva y cómo influye en la dinámica del aula. Por ejemplo, podrían diseñarse estudios para evaluar si el uso de ChatGPT mejora la participación de los estudiantes en debates académicos o si ayuda a los docentes a ofrecer una retroalimentación más personalizada.

Otro aspecto relevante para futuras investigaciones es el impacto de ChatGPT en la evaluación del desempeño académico de los estudiantes. Dado que esta herramienta puede apoyar la preparación de trabajos académicos y la comprensión de conceptos complejos, sería relevante investigar si su uso se traduce en mejoras medibles en las calificaciones de los estudiantes o en la calidad de sus trabajos. Esto permitiría a las universidades evaluar con mayor precisión el valor agregado de integrar ChatGPT en los procesos de enseñanza y aprendizaje, así como ajustar sus políticas de adopción tecnológica para maximizar los beneficios académicos.

Dado que este estudio se centra en la percepción de los estudiantes, una limitación inherente es la subjetividad de las respuestas, lo que puede introducir sesgos en la interpretación de los hallazgos. Diferencias individuales, experiencias previas y contextos institucionales podrían haber influido en la evaluación que los participantes hicieron de la utilidad y confiabilidad de ChatGPT. Aunque los datos autoreportados proporcionan información valiosa sobre las actitudes de los usuarios, pueden no reflejar completamente los patrones de comportamiento o las tendencias de adopción a largo plazo. Para abordar estas limitaciones, futuras investigaciones podrían incorporar enfoques cualitativos, como entrevistas en profundidad o grupos focales, para explorar las motivaciones subyacentes detrás de las percepciones de los estudiantes.

Además, estudios longitudinales permitirían analizar cómo evoluciona la confianza en herramientas de IA a lo largo del tiempo, lo que contribuiría a una comprensión más profunda del fenómeno. Esto ayudaría a las universidades a desarrollar estrategias de implementación a largo plazo, adaptándose a las necesidades cambiantes de los estudiantes y a la evolución continua de la IA generativa.

Este estudio proporciona una base sólida para el diseño de políticas educativas más efectivas en la integración de IA, al tiempo que plantea preguntas clave que

merecen una mayor exploración en futuras investigaciones. La integración estratégica de herramientas como ChatGPT no solo puede enriquecer las experiencias de aprendizaje de los estudiantes, sino que también puede preparar a las universidades para un futuro en el que la inteligencia artificial sea un componente central de la experiencia educativa.

REFERENCIAS

- Acemoglu, D., Johnson, S. y Viswanath, K. (2023). Why the Power of Technology Rarely Goes to the People. *MIT Sloan Management Review*, 65(1), 12-14.
- Al-Abdullatif, A. M. y Alsubaie, M. A. (2024). ChatGPT in Learning: Assessing Students' Use Intentions through the Lens of Perceived Value and the Influence of AI Literacy. *Behavioral Sciences*, 14(9). <https://doi.org/10.3390/bs14090845>
- Al-kfairy, M., Mustafa, D., Kshetri, N., Insiew, M. y Alfandi, O. (2024). Ethical Challenges and Solutions of Generative AI: An Interdisciplinary Perspective. *Informatics*, 11(3). <https://doi.org/10.3390/informatics11030058>
- Baron, R. M. y Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51(6), 1173-1182. <https://doi.org/10.1037/0022-3514.51.6.1173>
- Bender, E. M., Gebru, T., McMillan-Major, A. y Shmitchell, S. (2021). On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? *FACCT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*. 610-623. <https://doi.org/10.1145/3442188.3445922>
- Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, 13(3), 319-340. JSTOR. <https://doi.org/10.2307/249008>
- Forero-Corba, W. y Negre Bennasar, F. (2024). Técnicas y aplicaciones del Machine Learning e inteligencia artificial en educación: Una revisión sistemática. *RIED-Revista Iberoamericana de Educación a Distancia*, 27(1), 209-253. <https://doi.org/10.5944/ried.27.1.37491>
- Fornell, C. y Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39-50. <https://doi.org/10.2307/3151312>
- García Peñalvo, F. J., Llorens-Largo, F. y Vidal, J. (2024). La nueva realidad de la educación ante los avances de la inteligencia artificial generativa. *RIED-Revista Iberoamericana de Educación a Distancia*, 27(1), 9-39. <https://doi.org/10.5944/ried.27.1.37716>
- Hair Jr, J. F., Black, W. C., Babin, B. J. y Anderson, R. E. (2010). *Multivariate data analysis*. 7th edition. Prentice Hall.
- Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical Mediation Analysis in the New Millennium. *Communication Monographs*, 76(4), 408-420. <https://doi.org/10.1080/03637750903310360>
- Huang, W., Zhang, H., Zhang, Z., Zhang, B., Sun, X., Huo, Y., Feng, Y., Tian, P., Mo, H. y Wang, C. (2022). A prognostic nomogram based on a new classification of combined micropapillary and solid components for stage IA invasive lung adenocarcinoma. *Journal of Surgical Oncology*, 125(4), 796-808. <https://doi.org/10.1002/jso.26760>

- Ladhari, R. (2012). The lodging quality index: An independent assessment of validity and dimensions. *International Journal of Contemporary Hospitality Management*, 24(4), 628-652. <https://doi.org/10.1108/0959611211217914>
- Ly, B., Doeur, B. y nat, S. (2024). Key factors influencing digital learning adoption among Cambodian university students: An integrated theoretical approach. *Computers in Human Behavior Reports*, 15, 100460-100460. <https://doi.org/10.1016/j.chbr.2024.100460>
- Marimon, F., Mas-Machuca, M. y Akhmedova, A. (2024). Trusting in Generative AI: Catalyst for Employee Performance and Engagement in the Workplace. *International Journal of Human-Computer Interaction*, 1-16. <https://doi.org/10.1080/10447318.2024.2388482>
- Mayer, R. C., Davis, J. H. y Schoorman, F. D. (1995). An Integrative Model Of Organizational Trust. *Academy of Management Review*, 20(3), 709-734. <https://doi.org/10.5465/amr.1995.9508080335>
- Nunnally, J. C. y Bernstein, I. H. (1994). Psychometric theory McGraw-Hill series. *Psychology*, 3.
- Parasuraman, A. (2000). Technology Readiness Index (Tri): A Multiple-Item Scale to Measure Readiness to Embrace New Technologies. *Journal of Service Research*, 2(4), 307-320. <https://doi.org/10.1177/1094670514539730>
- Parasuraman, A. y Colby, C. L. (2014). An Updated and Streamlined Technology Readiness Index: TRI 2.0. *Journal of Service Research*, 18(1), 59-74. <https://doi.org/10.1177/1094670514539730>
- Rogers, E. M. (1995). *Diffusion of innovations*. Free Press.
- Tang, K.-Y., Chang, C.-Y. y Hwang, G.-J. (2023). Trends in artificial intelligence-supported e-learning: A systematic review and co-citation network analysis (1998–2019). *Interactive Learning Environments*, 31(4), 2134-2152. <https://doi.org/10.1080/10494820.2021.1875001>
- Venkatesh, V., Morris, M. G., Davis, G. B. y Davis, F. D. (2003). User Acceptance of Information Technology: Toward a Unified View. *MIS Quarterly*, 27(3), 425-478. <https://doi.org/10.2307/30036540>
- Wolfinbarger, M. y Gilly, M. C. (2003). eTailQ: dimensionalizing, measuring and predicting etail quality. *Journal of Retailing*, 79(3), 183-198. [https://doi.org/10.1016/S0022-4359\(03\)00034-4](https://doi.org/10.1016/S0022-4359(03)00034-4)
- Zawacki-Richter, O., Marín, V. I., Bond, M. y Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education – where are the educators? *International Journal of Educational Technology in Higher Education*, 16(1), 39-39. <https://doi.org/10.1186/s41239-019-0171-0>
- Zhao, X., Lynch, J. G., Jr. y Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and Truths about Mediation Analysis. *Journal of Consumer Research*, 37(2), 197-206. <https://doi.org/10.1086/651257>

Fecha de recepción del artículo: 1 de diciembre de 2024

Fecha de aceptación del artículo: 7 de marzo de 2025

Fecha de aprobación para maquetación: 17 de marzo de 2025

Fecha de publicación en OnlineFirst: 7 de abril de 2025

Fecha de publicación: 1 de julio de 2025

Efectos de la conciencia y confianza en la disposición estudiantil para usar ChatGPT: modelo TAM-ECM integrado

The effects of awareness and trust on students' willingness to use ChatGPT: an integrated TAM-ECM model



- ID Sultan Hammad Alshammari - *University of Ha'il, UoH (Arabia Saudita)*
- ID Abdullah Zaid Almankory - *University of Ha'il, UoH (Arabia Saudita)*
- ID Muna Eid Alrashidi - *University of Ha'il, UoH (Arabia Saudita)*

RESUMEN

Aunque el uso de ChatGPT en el ámbito educativo se ha expandido rápidamente en las universidades, se conoce poco sobre los factores que influyen en la intención de los estudiantes de utilizar esta herramienta para apoyar su aprendizaje. Este estudio aborda esa brecha integrando el Modelo de Aceptación de la Tecnología (TAM), ampliado con las variables de conciencia y confianza, y el Modelo de Confirmación de Expectativas (ECM), con los constructos de confirmación y satisfacción. El modelo propuesto profundiza en la comprensión de la disposición de los estudiantes a utilizar ChatGPT. Se recopiló datos de 322 estudiantes universitarios, analizados mediante Modelado de Ecuaciones Estructurales (SEM) con AMOS. Los resultados revelaron que la conciencia influye significativamente y de forma positiva en la utilidad percibida (PU) y en la facilidad de uso percibida (PEU). Además, la confianza mostró un efecto positivo sobre la PU, pero no impactó de manera relevante la PEU. Asimismo, la PU, la PEU y la confirmación afectaron positivamente la satisfacción de los estudiantes, la cual incide en su intención de utilizar ChatGPT. Por otro lado, tanto la PU como la PEU tuvieron un efecto positivo significativo en la intención de uso (BI) de la herramienta. El estudio ofrece recomendaciones para desarrolladores, responsables de políticas educativas e instituciones académicas, orientadas a comprender y fomentar la adopción de ChatGPT. Además, aporta información valiosa para mejorar el diseño y la seguridad del sistema, favoreciendo una experiencia amigable para el usuario y promoviendo su uso en el ámbito educativo.

Palabras clave: ChatGPT; Modelo de Aceptación de la Tecnología (MAT); Modelado de Ecuaciones Estructurales (MES); satisfacción; conciencia; confianza.

ABSTRACT

Although the use of ChatGPT in the educational field has spread rapidly in universities, there is little information on factors affecting students' intentions toward using ChatGPT to support their learning. This study bridges the gap by integrating the extended Technology Acceptance Model (TAM) (including awareness and trust) with Expectation Confirmation Model (ECM) constructs, namely confirmation and satisfaction. This proposed model provides an in-depth understanding of students' willingness to use ChatGPT. Data was collected from 322 university students and analyzed using a second-generation analysis technique, Structural Equation Modeling (SEM) using AMOS. The results revealed that awareness significantly positively affected students' perceived usefulness (PU) and perceived ease of use (PEU). Furthermore, trust had a significant positive effect on PU, but an insignificant effect on PEU. In addition, PU, PEU, and confirmation significantly positively affected students' satisfaction, affecting their behavioral intention toward using ChatGPT for their learning. Furthermore, PU and PEU significantly positively affected students' behavior intention toward using ChatGPT. This study offers recommendations to developers of ChatGPT, policymakers, and educational institutes by understanding the influential factors on students' willingness to use ChatGPT. This study assists ChatGPT developers and designers by offering insight regarding designing and improving the user's secure and friendly system, which may enhance the use of ChatGPT among students.

Keywords: ChatGPT; TAM; ECM; satisfaction; awareness; trust.

INTRODUCCIÓN

La integración de la inteligencia artificial (IA) y las tecnologías de la información y la comunicación (TIC) ha dado lugar a innovaciones que han transformado radicalmente el panorama educativo (Kelly et al., 2023). El auge acelerado de la IA ha incrementado su presencia en la vida cotidiana, y los chatbots representan un ejemplo destacado de sistemas de IA utilizados para la interacción humano-computadora, conocida por sus siglas en inglés HCI, (Human-Computer Interaction) (Bansal y Khan, 2018). Aunque inicialmente los chatbots fueron diseñados para el entretenimiento y la simulación de conversaciones humanas, sus aplicaciones se han expandido a diversos sectores (Iku-Silan et al., 2023) como los negocios, la recuperación de información, el comercio electrónico y la educación (Ismail et al., 2024). En este contexto, OpenAI lanzó ChatGPT el 30 de noviembre de 2022, atrayendo a más de un millón de suscriptores en un corto período (Baidoo-Anu y Ansah, 2023). ChatGPT es un modelo de lenguaje de gran tamaño, conocido por sus siglas en inglés LLM (Large Language Model), basado en aprendizaje automático, capaz de analizar extensos volúmenes de datos textuales para generar textos coherentes y sofisticados. Este modelo, al igual que otras herramientas basadas en LLM, ha sido ampliamente adoptado por profesionales e investigadores para la redacción de discursos y ensayos, revisión de artículos, resumen de literatura, identificación de brechas en la investigación y generación de código para análisis estadísticos (van Dis et al., 2023). Se han desarrollado diversas versiones de ChatGPT con una cantidad creciente de parámetros: ChatGPT-1 con 117 millones, ChatGPT-2 con 1,5 mil millones, ChatGPT-3 con 175 mil millones y ChatGPT-4 con 100 billones. Este incremento ha permitido mejorar significativamente la capacidad del modelo para generar textos con un nivel de coherencia y fluidez similar al lenguaje humano (OpenAI, 2023). En otras palabras, ChatGPT no solo es capaz de redactar documentos académicos y científicos en distintos estilos y tonos (Lund y Wang, 2023), sino que también puede analizar grandes volúmenes de información, evaluar publicaciones en redes sociales y facilitar la comprensión de múltiples idiomas mediante traducción automática.

Las capacidades de resumen automatizado de ChatGPT permiten a los investigadores mantenerse al tanto de los últimos avances en sus respectivas áreas (Lund y Wang, 2023), mientras que su acceso permanente y respuesta inmediata convierten a esta herramienta en un recurso valioso para la consulta académica (Frieder et al., 2023; Haleem et al., 2022). A medida que ChatGPT gana popularidad y aumenta el interés en sus aplicaciones, los estudiantes han comenzado a utilizarlo en diversas áreas académicas. Esta herramienta puede asistirles en su aprendizaje al ayudarles a comprender y desglosar ideas o conceptos presentados en sus estudios (van Dis et al., 2023). Además, ChatGPT puede considerarse una aplicación de referencia en múltiples disciplinas, ya que permite recopilar información y conocimientos sobre una amplia variedad de temas (Cao et al., 2023). Los estudiantes

lo emplean para comprender conceptos complejos, obtener información adicional sobre determinados temas y superar desafíos académicos (Lund y Wang, 2023). Su disponibilidad ininterrumpida y capacidad de respuesta inmediata brinda a los estudiantes acceso rápido y eficiente a información relevante y recursos educativos (Baidoo-Anu y Ansah, 2023).

Si bien se han realizado diversos estudios sobre el papel de ChatGPT en la educación (por ejemplo, Al-Sharafi et al., 2023; Rudolph et al., 2023; Alshammari y Alshammari, M. H., 2024), existe una escasez de investigaciones centradas en la aceptación y adopción de esta tecnología (Bin-Nashwan et al., 2023; Bernabei et al., 2023). Rahman et al. (2023) llevó a cabo un estudio basado en la revisión de 787 artículos indexados en Scopus, en el que se encontró que solo una fracción de los trabajos abordaba la aceptación de ChatGPT por parte de los estudiantes. Además, la mayoría de estos estudios se han realizado en países desarrollados, como el Reino Unido y Estados Unidos, mientras que hay una notable falta de investigaciones en países en desarrollo, como Arabia Saudita. Asimismo, aunque numerosos estudios han analizado tecnologías emergentes, como la computación en la nube y la realidad virtual (Sestino y D'Angelo, 2023), es fundamental considerar que la adopción de nuevas tecnologías varía según el contexto cultural (Ashraf et al., 2019). A pesar del gran potencial de ChatGPT para mejorar el aprendizaje de los estudiantes, es esencial comprender los factores que influyen en su disposición a utilizar esta herramienta. Al-Sayid y Kirkil (2023) destacan que el Modelo de Aceptación de la Tecnología, conocido por sus siglas en inglés TAM (Technology Acceptance Model), es una referencia clave en la investigación sobre la percepción y el uso de nuevas tecnologías. Según TAM, la utilidad percibida, conocida por sus siglas en inglés PU (Perceived Usefulness), y la facilidad de uso percibida, conocida por sus siglas en inglés PEU (Perceived Ease of Use), son los dos principales determinantes de la adopción de un sistema tecnológico (Davis, 1989). No obstante, este modelo no opera de manera aislada, ya que existen otros factores externos que pueden influir en la percepción y uso de la tecnología, como la conciencia y la confianza en su aplicación (Abdalla, 2024). La conciencia se refiere al grado de familiaridad con una tecnología y el conocimiento de sus beneficios (Mutahar et al., 2018). Antes de utilizar una herramienta como ChatGPT, los estudiantes deben conocer su existencia y comprender sus características y ventajas. La educación en torno a ChatGPT puede desempeñar un papel clave en la manera en que los estudiantes responden a esta tecnología. Por otro lado, la confianza se define como la creencia en la credibilidad y fiabilidad del sistema (Arpaci, 2016), y es un factor crucial en la adopción de nuevas tecnologías (Liu y Tao, 2022). Por otra parte, Bhattacharjee (2001) propone el Modelo de Confirmación de Expectativas, conocido por sus siglas en inglés ECM (Expectation Confirmation Model), como una estructura ampliamente utilizada para evaluar los factores que influyen en el uso continuo de una tecnología. Este modelo se basa en tres constructos principales: PU, confirmación y satisfacción.

A pesar del creciente interés en la implementación de ChatGPT en entornos académicos, todavía existen pocas investigaciones que analicen las relaciones causales entre los factores que influyen en la intención de los estudiantes de adoptar esta tecnología en su proceso de aprendizaje (Yu et al., 2024). Por ello, este estudio busca abordar esta brecha en la literatura, identificando los principales factores que afectan la adopción de ChatGPT en el ámbito educativo. Este trabajo propone un modelo teórico que combina una extensión del TAM, incorporando la conciencia y la confianza como factores externos, con los constructos del ECM, específicamente la confirmación y la satisfacción. Hasta donde se tiene conocimiento, este es el primer estudio que integra el TAM extendido (con conciencia y confianza) con el ECM para evaluar los factores determinantes en la intención de los estudiantes de utilizar ChatGPT en su proceso de aprendizaje.

Revisión de la literatura y desarrollo de hipótesis

ChatGPT es un chatbot basado en un modelo de gran escala (Schulman et al., 2022). Esta herramienta se distingue de otros sistemas de IA por su capacidad para procesar el lenguaje natural de manera similar al cerebro humano. Su aplicación se ha expandido a múltiples campos, incluyendo la investigación académica, la redacción creativa y técnica, la evaluación de pruebas, y la comunicación en el ámbito empresarial y del desarrollo de software (Tung, 2023). Con los avances tecnológicos y la creciente integración global de la IA, ChatGPT se ha implementado ampliamente en universidades, proporcionando enfoques eficientes y rentables para motivar e involucrar a los estudiantes a través de experiencias de aprendizaje personalizadas (Polyportis, 2024; Albayati, 2024). A medida que la IA evoluciona, ChatGPT ha demostrado un gran potencial para transformar el ámbito educativo. La incorporación de diversas herramientas y tecnologías en la enseñanza y el aprendizaje ha optimizado la eficiencia en el aula, y ChatGPT se posiciona como una de las innovaciones más relevantes en este proceso. Investigaciones previas han evaluado su viabilidad en el ámbito educativo, explorando su uso en actividades académicas y profesionales (Brown et al., 2020; Emenike y Emenike, 2023), así como su papel en el aprendizaje de idiomas (Alshammari y Alshammari, M. H., 2024). Su capacidad para proporcionar retroalimentación instantánea sobre tareas, responder preguntas, generar materiales educativos y elaborar planes de estudio lo convierte en una herramienta valiosa para el aprendizaje personalizado (Albayati, 2024).

Desde la perspectiva de los estudiantes, es fundamental analizar los factores que influyen en la adopción y aceptación de ChatGPT (Patel y Lam, 2023). Para ello, el Modelo de Aceptación de la Tecnología (TAM), desarrollado por Davis (1989), ha sido ampliamente utilizado en la evaluación de la adopción de diversas tecnologías (Venkatesh y Davis, 2000; Folkinshteyn y Lennon, 2016; Alshammari y Alshammari, R. A., 2024). Este modelo teórico explica los factores que influyen en la aceptación de una nueva tecnología y ha sido validado en múltiples contextos, incluyendo

aplicaciones móviles, sitios web y redes sociales, demostrando su fiabilidad y aplicabilidad (Alshammari y Alshammari, R. A., 2024). A lo largo del tiempo, los investigadores han ampliado y modificado el TAM para incluir variables adicionales, como la satisfacción del usuario, la calidad del sistema y la capacidad de innovación personal (Esposito et al., 2020; Alshammari y Alshammari, R. A., 2024). Por otro lado, el Modelo de Confirmación de Expectativas (ECM), propuesto por Bhattacharjee (2001), ha sido ampliamente utilizado para evaluar la continuidad en el uso de diversas tecnologías. Su aplicabilidad ha sido validada en contextos como los libros de texto digitales (Joo et al., 2017), aplicaciones móviles (Tam et al., 2020), redes sociales como Facebook (Mouakket, 2015) y dispositivos inteligentes (Park, 2020). Diversos estudios han sugerido la integración de modelos teóricos para comprender mejor los factores que influyen en la intención de uso, conocida por sus siglas en inglés BI (Behavioural Intention) de nuevas tecnologías (Huang y Zhi, 2023; Alshammari y Alshammari, R. A., 2024), así como para superar las limitaciones inherentes a cada modelo individualmente (Wandira et al., 2024). Dado que el objetivo principal de este estudio es analizar en profundidad los factores que afectan la intención de los estudiantes de utilizar ChatGPT en su aprendizaje, se propone un modelo teórico integrado. Este modelo extiende el TAM mediante la incorporación de dos factores externos clave: confianza y conciencia, y lo complementa con los constructos del ECM, a fin de proporcionar un marco más completo para comprender la disposición de los estudiantes a adoptar ChatGPT en su formación académica.

El Modelo de Aceptación de la Tecnología (TAM)

El TAM fue desarrollado por Davis en 1989 y se ha utilizado ampliamente para explicar la interacción de los usuarios con diversas tecnologías. En este modelo, la PU y la PEU son los principales factores que determinan la intención de adoptar una tecnología (Davis, 1989). Desde su formulación, el TAM ha sido ampliado mediante la incorporación de diversos constructos externos, con el objetivo de mejorar su capacidad predictiva y ofrecer una explicación más precisa de los factores que influyen en la intención de los usuarios de adoptar tecnologías de la información (Nikou y Economides, 2017). En el contexto de las aplicaciones de IA, diversos estudios han extendido el TAM al incluir variables adicionales. Mohr y Kühl (2021) ampliaron el modelo para evaluar la aceptación de la IA en la agricultura alemana, incorporando factores como la innovación personal y los derechos de propiedad sobre los datos empresariales. Por su parte, Li (2023) introdujo la motivación para el aprendizaje como una variable externa para analizar su impacto en la aceptación de los sistemas de IA por parte de los estudiantes. Xu et al. (2023) extendieron el TAM para examinar cómo las características técnicas, la experiencia previa y la confianza percibida influyen en la intención de los usuarios de utilizar aplicaciones de pintura basadas en IA. A pesar de la extensa aplicación del TAM y sus versiones extendidas en distintos ámbitos de la IA, el uso del modelo en el estudio de ChatGPT ha sido

limitado. Hasta la fecha, ningún estudio ha integrado el TAM extendido con las variables de conciencia y confianza, junto con el ECM, para evaluar los factores que influyen en la intención de los estudiantes de utilizar ChatGPT en su aprendizaje. Por ello, el presente estudio propone un modelo teórico que combina el TAM extendido con el ECM, con el objetivo de ofrecer un marco integral para comprender los factores determinantes en la adopción de ChatGPT por parte de los estudiantes.

El Modelo de Confirmación de Expectativas (ECM)

El ECM fue desarrollado por Bhattacharjee (2001) para explicar la intención de compra y el uso continuado de una tecnología. Desde su formulación, el ECM ha sido ampliamente validado por investigadores en el contexto del uso sostenido de diversas tecnologías de sistemas de información (SI), como los libros de texto digitales (Joo et al., 2017), Facebook, el gobierno electrónico y los sistemas de gestión del aprendizaje (Ashrafi et al., 2022). De acuerdo con el ECM, la PU y la confirmación son los principales determinantes de la satisfacción del usuario, la cual, a su vez, influye en la intención continuada de uso de una tecnología. A lo largo de los años, diversos estudios han integrado el ECM con otros modelos teóricos. Por ejemplo, se ha combinado con el modelo de éxito de los SI en entornos de aprendizaje virtual (Alshammari y Alshammari, R. A., 2024), con la teoría de la autodeterminación en el ámbito de la banca electrónica (Rahi et al., 2023) y con el TAM en el contexto de sistemas académicos basados en la nube (Wandira et al., 2024). Sin embargo, hasta la fecha, la integración del ECM con el TAM extendido, incorporando las variables de conciencia y confianza como factores externos, no ha sido explorada para evaluar su impacto en la disposición de los estudiantes a utilizar ChatGPT en su aprendizaje. Por lo tanto, el modelo propuesto en este estudio contribuye a la literatura existente al ofrecer un marco teórico integrado que permite comprender los factores que influyen en la aceptación de ChatGPT entre los estudiantes. En la siguiente sección, se presentará el modelo de investigación junto con las hipótesis derivadas de este marco teórico.

Modelo de investigación

Este estudio propone un modelo que integra el TAM extendido con variables adicionales, como la conciencia y la confianza, junto con los factores del ECM, como la satisfacción y la confirmación, con el fin de evaluar los factores que influyen en la intención de los estudiantes de utilizar ChatGPT. A continuación, se describen las variables del modelo y las hipótesis formuladas a partir de la revisión de la literatura y la fundamentación en estos modelos teóricos.

Conciencia

La conciencia es un factor crítico en la adopción de nuevas tecnologías, ya que los usuarios deben conocer su funcionalidad, existencia y beneficios antes de aceptarla. En el contexto de ChatGPT, la conciencia se relaciona con la comprensión de su aplicabilidad, limitaciones y capacidades (Eppler et al., 2024). Desarrollar la conciencia entre los usuarios es un paso crucial para la adopción de cualquier tecnología. Aquellos usuarios que no son conscientes de los beneficios de ChatGPT podrían no estar dispuestos a considerar su aceptación ni a utilizar esta tecnología (Maheshwari, 2023). Si bien la influencia de la conciencia ha sido analizada en diversas áreas, como la banca móvil (Mutahar et al., 2018), los hogares inteligentes (Shuhaiber y Mashal, 2019) y los servicios de IA (Flavián et al., 2022), existe una falta de estudios que examinen su impacto en la PU y la PEU en relación con ChatGPT. Cuando los usuarios comprenden las capacidades de una tecnología, tienden a considerarla útil y fácil de usar. Por lo tanto, se plantean las siguientes hipótesis:

H1: La conciencia influye positivamente en la utilidad percibida (PU).

H2: La conciencia influye positivamente en facilidad de uso percibida (PEU).

Confianza

La confianza percibida se refiere a la creencia de los usuarios con respecto a la credibilidad y confiabilidad de una tecnología (Arpaci, 2016), lo que puede influir en su decisión de adoptarla (Liu y Tao, 2022). En el caso de las aplicaciones de IA, la confianza juega un papel fundamental en la aceptación de estas herramientas (Choung et al., 2023). Diversos estudios han incorporado la confianza en el modelo TAM como un factor externo. Se espera que, cuando los usuarios perciban ChatGPT como una herramienta confiable y precisa, aumente su PU y PEU. Así, se formulan las siguientes hipótesis:

H3: La confianza en ChatGPT influye positivamente en la utilidad percibida (PU).

H4: La confianza en ChatGPT influye positivamente en facilidad de uso percibida (PEU).

Facilidad de uso percibida (PEU)

La PEU se define como el grado en que los usuarios creen que una tecnología es fácil de utilizar y no requiere esfuerzo significativo (Davis, 1989). En el contexto de ChatGPT, esto implica la facilidad con la que los estudiantes y educadores pueden interactuar con la herramienta e integrarla en sus actividades académicas (Baek y Kim, 2023). Estudios empíricos han demostrado que la PEU influye en la satisfacción

del usuario (Kashive et al., 2020). Cuando los usuarios perciben que una tecnología es sencilla de utilizar, es más probable que experimenten satisfacción y continúen utilizándola. Por ello, se plantean las siguientes hipótesis:

- H5: Facilidad de uso percibida (PEU) influye positivamente en la satisfacción del usuario.
- H6: Facilidad de uso percibida (PEU) influye positivamente en la intención de uso (BI).

Utilidad Percibida (PU)

La PU se refiere a la medida en que los usuarios creen que el uso de una tecnología mejorará su desempeño (Davis, 1989). Es una de las variables clave en la adopción de tecnologías de la información (Basak et al., 2015). En el caso de ChatGPT, la PU se asocia con los beneficios académicos percibidos por los estudiantes al utilizar la herramienta (Niu y Mvondo, 2024). Cuando los usuarios consideran que ChatGPT es útil, es más probable que experimenten satisfacción con su uso, lo que, a su vez, influye en su intención de seguir utilizándolo. Por lo tanto, se formulan las siguientes hipótesis:

- H7: La utilidad percibida (PU) influye positivamente en la satisfacción del usuario.
- H8: La utilidad percibida (PU) influye positivamente en la intención de uso (BI).

Confirmación

La confirmación se refiere a la percepción de los usuarios sobre el grado en que su experiencia real con un SI coincide con sus expectativas previas. La literatura ha demostrado que la confirmación influye directamente en la PU de una tecnología (Bhattacharjee, 2001). Además, diversos estudios han identificado la confirmación como un factor clave en la satisfacción de los estudiantes con los sistemas de información (Limayem y Cheung, 2008; Stone y Baker-Eveleth, 2013; Alshammari y Alshammari, R. A., 2024). Cuando la experiencia de uso de un SI cumple con las expectativas del usuario, su nivel de satisfacción tiende a aumentar. En el contexto de ChatGPT, se espera que, cuando los estudiantes confirmen que la herramienta satisface sus expectativas y necesidades académicas, su PU y satisfacción aumenten. Con base en esta relación, se proponen las siguientes hipótesis:

- H9: La confirmación influye positivamente en la utilidad percibida (PU).
- H10: La confirmación influye positivamente en la satisfacción del usuario.

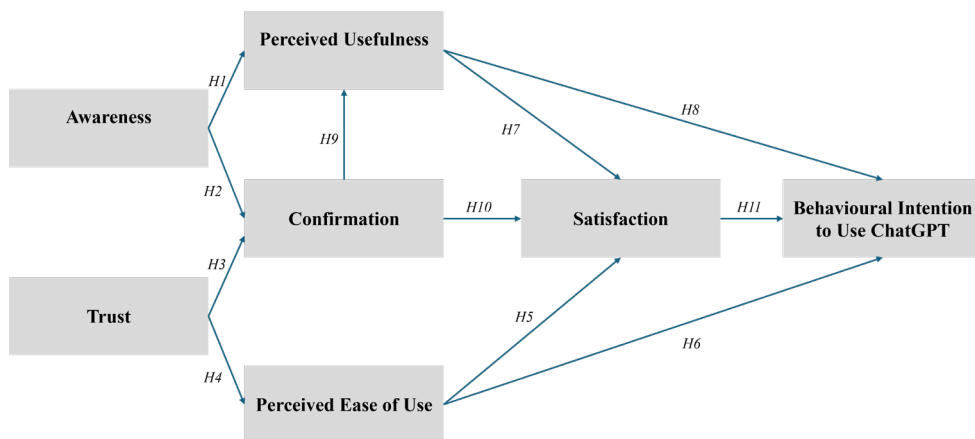
Satisfacción

La satisfacción se define como la evaluación subjetiva que un usuario realiza sobre su experiencia con una tecnología, comparándola con sus expectativas previas (Bhattacharjee, 2001). Satisfacción es un factor esencial en la determinación de la BI continuo de una tecnología (Mouakket, 2015; Hong et al., 2006; Choi et al., 2019). Numerosos estudios han demostrado que los usuarios que experimentan altos niveles de satisfacción con una tecnología tienen una mayor intención de seguir utilizándola. En el ámbito educativo, la satisfacción ha sido identificada como un predictor clave de la intención de los estudiantes de continuar usando tecnologías de aprendizaje (Ifinedo, 2007; Ramadhan et al., 2022). En el contexto de ChatGPT, se espera que los estudiantes que se sientan satisfechos con su uso muestren una mayor intención de seguir utilizándolo para sus actividades académicas. Así, se plantea la siguiente hipótesis:

H11: La satisfacción influye positivamente en la intención de los estudiantes de usar ChatGPT.

A continuación, en la Figura 1, se presenta el modelo teórico propuesto.

Figura 1
Modelo propuesto



METODOLOGÍA

Diseño de la investigación

Este estudio empleó un enfoque cuantitativo, adecuado para alcanzar los objetivos de la investigación, que buscan evaluar la relación entre los constructos del modelo propuesto (Sekaran y Bougie, 2016).

Medición

Para la recolección de datos, se utilizó un cuestionario estructurado dirigido a estudiantes universitarios. El cuestionario consta de dos secciones: 1) Información demográfica, que incluye variables como género, rango de clase y universidad; and 2) Medición de constructos, con ítems diseñados para evaluar las variables del modelo propuesto. Los ítems utilizados para medir los constructos de este estudio fueron adaptados de investigaciones previas reconocidas en la literatura académica. En el caso de la PU, la PEU y la BI, se tomaron como referencia las escalas desarrolladas por Davis (1989). Por otro lado, los ítems que evalúan la conciencia sobre la tecnología fueron adaptados del estudio de Abdalla (2024), mientras que los relacionados con la confianza en su uso se basaron en la investigación de Rahman et al. (2023). En cuanto a los constructos del modelo ECM, los ítems utilizados para medir la confirmación se derivaron de Bhattacharjee (2001), y aquellos que evalúan la satisfacción fueron tomados del estudio de Almulla (2024). Para garantizar la precisión y fidelidad en la interpretación de los cuestionarios, se aplicó un procedimiento de retrotraducción del inglés al árabe, idioma nativo de los participantes del estudio. La medición de todos los ítems se realizó utilizando una escala Likert de cinco puntos.

Procedimiento de recolección de datos

El cuestionario se diseñó y distribuyó a través de Google Forms durante el primer semestre de 2024. Se utilizó un método de muestreo aleatorio para seleccionar a los participantes, asegurando una representación diversa dentro de la población universitaria. Todos los encuestados eran estudiantes de la Universidad de Ha'il, y se obtuvo un total de 322 respuestas completas para el análisis posterior.

Análisis de datos

Para el análisis de los datos, se emplearon dos programas estadísticos: SPSS para analizar la información demográfica de los encuestados, y el Modelado de Ecuaciones Estructurales, conocido por sus siglas en inglés SEM (Structural Equation Modeling) con AMOS para evaluar el modelo de medición y el modelo estructural. El análisis se

realizó en dos fases: 1) Análisis Factorial Confirmatorio (AFC) para evaluar la validez y fiabilidad del modelo de medición; y 2) SEM para analizar las relaciones entre los constructos del modelo propuesto. El SEM es una técnica ampliamente utilizada en estudios psicológicos y de comportamiento (Hair et al., 2010). Kline (2023) también recomienda el uso de SEM para analizar las relaciones entre constructos latentes en modelos teóricos complejos.

RESULTADOS

Un total de 322 estudiantes completaron el cuestionario, y sus respuestas fueron analizadas. La Tabla 1 presenta la distribución demográfica de los encuestados según género y facultad.

El análisis demográfico de la muestra revela que la mayoría de los participantes eran hombres, representando el 61,5 % del total (198 estudiantes), mientras que las mujeres constituían el 38,5 % (124 estudiantes). En cuanto a su distribución por facultades, la mayor proporción de estudiantes estaba inscrita en la Facultad de Ciencias de la Computación e Ingeniería, con un 40,4 % (130 estudiantes). Le seguían la Facultad de Artes, que concentraba el 17,1 % (55 estudiantes), y la Facultad de Administración de Empresas, con un 16,1 % (52 estudiantes). Otras facultades con una presencia significativa en la muestra fueron la Facultad de Ciencias, con un 11,2 % (36 estudiantes), la Facultad de Educación, con un 8,7 % (28 estudiantes), y la Facultad de Medicina, con un 4,7 % (15 estudiantes). En contraste, la menor representación correspondió a la Facultad de Salud Pública e Informática en Salud, donde solo se matriculó el 1,9 % de los encuestados (seis estudiantes).

Tabla 1
Información demográfica

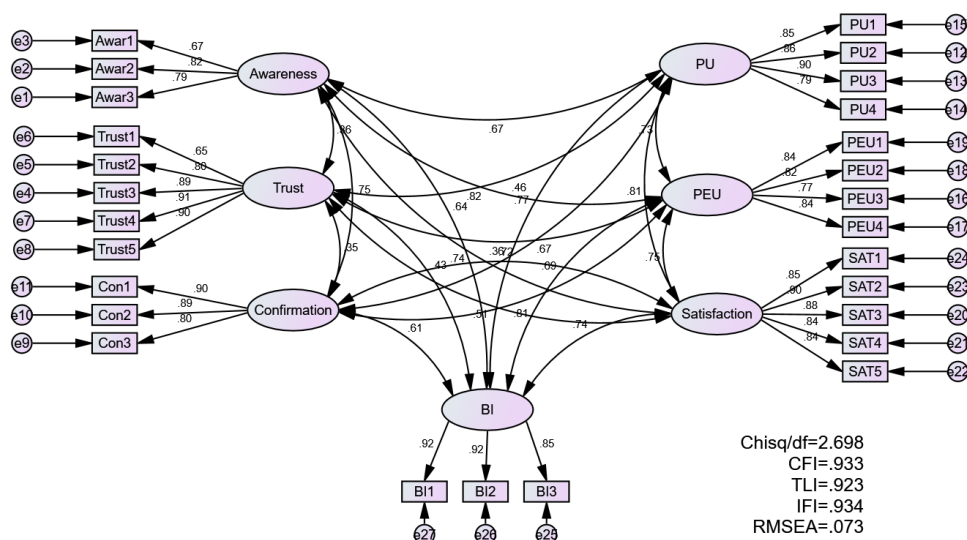
		Frecuencia	Porcentaje
Género	Masculino	198	61.5
	Femenino	124	38.5
Facultad	Ciencias de la Computación e Ingeniería	130	40.4
	Artes	55	17.1
	Administración de Empresas	52	16.1
	Ciencia	36	11.2
	Educación	28	8.7
	Medicina	15	4.7
	Salud Pública e Informática en Salud	6	1.9
	Total	322	100.0

Análisis Factorial Confirmatorio (AFC)

Se realizó un AFC para evaluar la validez del modelo de medición. Los resultados obtenidos se presentan en la Figura 2.

Figura 2

Resultados del Análisis Factorial Confirmatorio (CFA)



Para que la validez de constructo sea aceptable, todos los índices de ajuste deben cumplir con los valores sugeridos por investigadores como Awang (2015), además de que las cargas factoriales deben alcanzar un mínimo de 0,6. Como se muestra en la Tabla 2, todos los índices de ajuste cumplen y superan los valores recomendados en la literatura (Hair et al., 2010; Awang, 2015). Asimismo, las cargas factoriales de todos los ítems para cada constructo oscilaron entre 0,65 y 0,92, superando el umbral mínimo de 0,6 (Yusoff et al., 2021). Estos resultados confirman la validez de constructo del modelo.

Tabla 2

Resumen de los índices de ajuste del modelo

Categoría	Índice	Valor obtenido	Valor de aceptación	Resultados
Ajuste absoluto	RMSEA	0.073	< 0.08	Acceptable (Hair et al., 2010; Awang, 2015)
Ajuste incremental	CFI	0.933	< 0.90	
	TLI	0.923	< 0.90	
	IFI	0.934	< 0.90	
Ajuste parsimonioso	Chi-cuadrado / grados de libertad	2.698	< 3.0	

Por otro lado, la validez convergente se establece cuando el coeficiente de confiabilidad compuesta, conocida por sus siglas en inglés RC (Composite Reliability), es superior a 0,70 (Rahlina et al., 2019) y la varianza media extraída, conocida por sus siglas en inglés AVE (Average Variance Extracted), es superior a 0,50 (Lowry y Gaskin, 2014). Tal como se observa en la Tabla 3, todos los valores de RC y AVE se encuentran por encima de los niveles recomendados, lo que garantiza la validez convergente del modelo. En cuanto a la validez discriminante, los valores de correlación entre los constructos oscilaron entre 0,35 y 0,82, lo que sugiere la ausencia de problemas de multicolinealidad que pudieran comprometer la fiabilidad de los resultados. Esta validez se confirma cuando la raíz cuadrada del AVE (en negrita) es mayor que cualquier otra correlación entre los constructos. Según los datos presentados en la Tabla 3, la validez discriminante se ha alcanzado, cumpliendo con los valores sugeridos por Sarstedt et al. (2021) y Afthanorhan et al. (2021).

Tabla 3

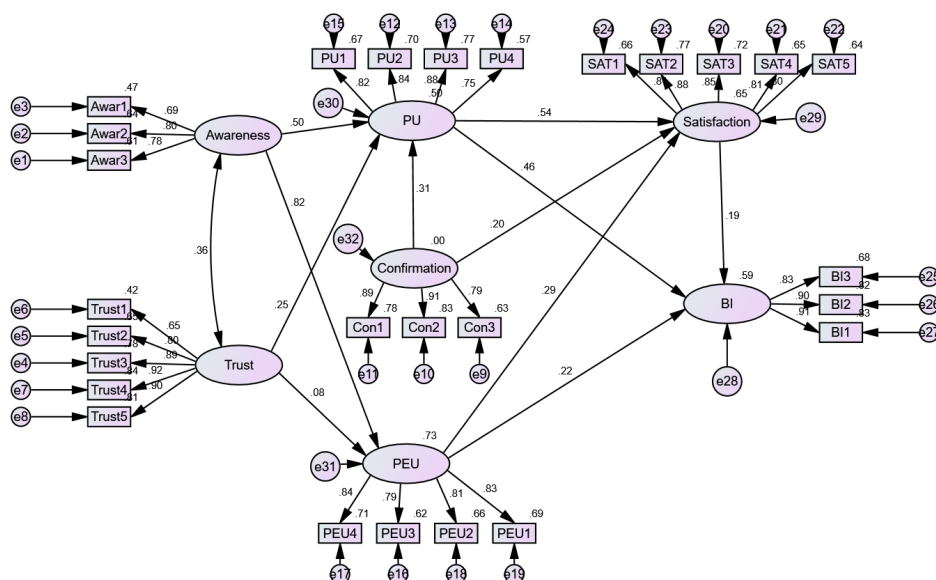
Fiabilidad, validez convergente y discriminante

	CR	AVE	Satis-fac-ción	Con-cien-cia	Con-fianza	Con-firma-ción	PU	PEU	BI
Satisfacción	0.936	0.744	0.863						
Conciencia	0.807	0.585	0.738	0.865					
Confiianza	0.920	0.699	0.507	0.359	0.836				
Confirmación	0.899	0.748	0.723	0.752	0.354	0.865			
PU	0.913	0.725	0.807	0.669	0.455	0.671	0.851		
PEU	0.891	0.672	0.752	0.824	0.362	0.805	0.732	0.839	
BI	0.925	0.805	0.736	0.636	0.430	0.611	0.774	0.692	0.897

Estimación estandarizada

La estimación estandarizada es fundamental para evaluar el coeficiente de determinación (R^2), las cargas factoriales y los coeficientes beta, mientras que la estimación no estandarizada es clave para la prueba de hipótesis. En primer lugar, se ejecutó la estimación estandarizada, cuyos resultados se presentan en la Figura 3.

Figura 3
Estimación estandarizada

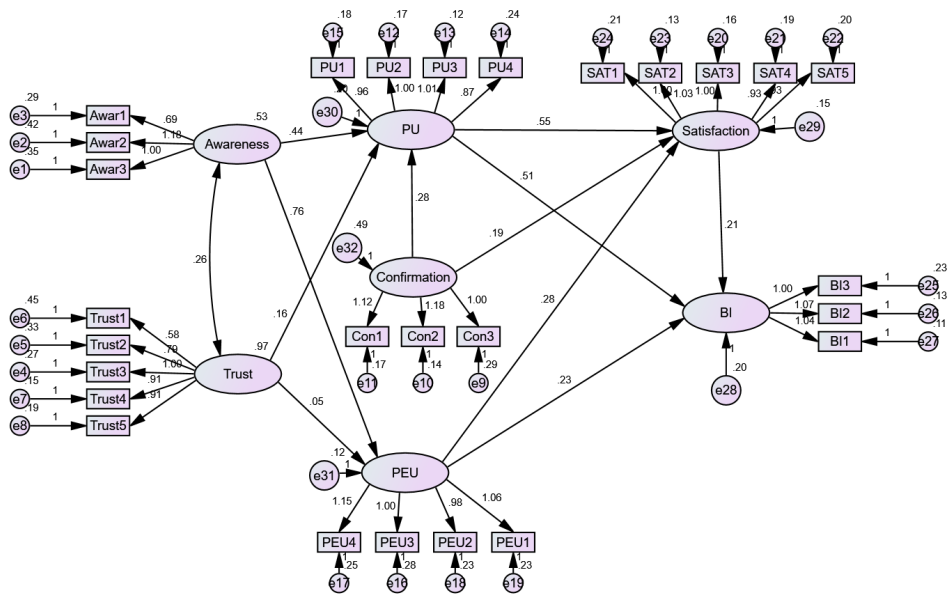


De acuerdo con Cohen (1988), un valor de R^2 superior a 0,26 indica un alto poder explicativo del modelo. Como se observa en la Figura 3, el coeficiente R^2 de la variable dependiente, BI alcanza un valor de 0,59. Esto sugiere que el modelo propuesto explica el 59 % de la variabilidad en la intención de los estudiantes de utilizar ChatGPT, atribuyéndola a factores como la satisfacción, la PU, la confirmación, la PEU, la conciencia y la confianza.

Estimación no estandarizada

La estimación no estandarizada permite analizar los coeficientes de regresión, los valores beta y la razón crítica, elementos esenciales para la prueba de hipótesis. Tras ejecutar este análisis, los resultados se presentan en la Figura 4.

Figura 4
Estimación no estandarizada



Resultados de las pruebas de hipótesis

Los hallazgos confirman que la conciencia tuvo un efecto positivo y significativo tanto en la PU como en la PEU, con valores $\beta = 0,441$, $p < 0,05$ y $\beta = 0,765$, $p < 0,05$, respectivamente. Esto valida las hipótesis H1 y H2. Por otro lado, la confianza tuvo un efecto positivo y significativo en la PU ($\beta = 0,161$, $p < 0,05$), pero su impacto en la PEU no fue significativo ($\beta = 0,054$, $p > 0,05$). En consecuencia, H3 fue aceptada, mientras que H4 fue rechazada. Asimismo, la PEU mostró un efecto positivo sobre la satisfacción ($\beta = 0,297$, $p < 0,05$) y la BI ($\beta = 0,232$, $p < 0,05$), respaldando las hipótesis H5 y H6.

De manera similar, la PU influyó positivamente en la satisfacción ($\beta = 0,549$, $p < 0,05$) y en la BI ($\beta = 0,506$, $p < 0,05$), lo que confirma H7 y H8. Por otro lado, la confirmación tuvo un impacto positivo en la PU ($\beta = 0,279$, $p < 0,05$) y en la satisfacción ($\beta = 0,188$, $p < 0,05$), validando H9 y H10. Finalmente, la satisfacción mostró una influencia significativa en la BI ($\beta = 0,209$, $p < 0,05$), lo que permite aceptar H11. Los coeficientes de regresión correspondientes se presentan en la Tabla 4.

Tabla 4
Ponderaciones de regresión

			Estimado	EE.	CR.	P	Resultados
PU	<---	Conciencia	.441	.053	8.267	***	Significativo
PEU	<---	Conciencia	.765	.066	11.556	***	Significativo
PU	<---	Confianza	.161	.034	4.718	***	Significativo
PEU	<---	Confianza	.054	.032	1.664	.096	Insignificante
PU	<---	Confirmación	.279	.046	6.079	***	Significativo
Satisfacción	<---	PEU	.282	.051	5.496	***	Significativo
Satisfacción	<---	PU	.549	.061	9.007	***	Significativo
Satisfacción	<---	Confirmación	.188	.042	4.457	***	Significativo
BI.	<---	Satisfacción	.209	.083	2.510	.012	Significativo
BI.	<---	PU	.506	.085	5.944	***	Significativo
BI.	<---	PEU	.232	.059	3.947	***	Significativo

DISCUSIÓN

Este estudio integró un modelo ampliado de la Teoría de Aceptación de Tecnología (TAM), incorporando la conciencia y la confianza, junto con el Modelo de Confirmación de Expectativas (ECM), para evaluar los factores que influyen en la intención de los estudiantes de utilizar ChatGPT. Los resultados revelaron que la conciencia tuvo un efecto positivo y significativo tanto en la utilidad percibida (PU) como en la facilidad de uso percibida (PEU), en línea con estudios previos (Mutahar et al., 2018; Almuraqab, 2020; Bamigbola y Adetimirin, 2020; Abdalla, 2024). Esto subraya el papel clave de la conciencia en la percepción de los estudiantes sobre la utilidad y facilidad de uso de ChatGPT. A medida que los estudiantes adquieren mayor conocimiento sobre ChatGPT y sus capacidades, es más probable que reconozcan su utilidad para apoyar su aprendizaje académico. Al estar informados sobre sus funciones, percibirán la herramienta como un recurso valioso para proporcionar información instantánea, explicar conceptos complejos y ofrecer asistencia lingüística.

Asimismo, la familiaridad con ChatGPT facilita su uso, lo que aumenta su adopción. En contraste, un conocimiento limitado sobre la plataforma puede disminuir tanto la PU como la PEU, generando la percepción de que ChatGPT es una herramienta demasiado compleja o difícil de aplicar. En cuanto a la confianza, los hallazgos indicaron que esta tuvo un impacto significativo en la PU (Torrent-Sellens et al., 2021; Liu y Tao, 2022). Cuando los estudiantes confían en la fiabilidad, precisión y uso ético de la información generada por ChatGPT, tienden a considerarlo como una herramienta útil para su aprendizaje. Esta confianza puede motivarlos a utilizar

la plataforma para diversas tareas académicas, como la resolución de problemas, la investigación y la redacción, ya que perciben que ChatGPT puede mejorar su desempeño. No obstante, los resultados mostraron que la confianza no tuvo un efecto significativo en la PEU. Esto sugiere que, aunque la confianza influye en la PU, no necesariamente afecta la PEU. Por ejemplo, un estudiante puede confiar en la información proporcionada por ChatGPT sin que esto implique que lo considere intuitivo o fácil de manejar.

Los hallazgos también revelaron que la confirmación tuvo un efecto positivo significativo en la PU y en la satisfacción, en coherencia con investigaciones previas (Bhattacharjee, 2001; Stone y Baker-Eveleth, 2013; Joo y Choi, 2016). Esto indica que, a medida que los estudiantes adquieren experiencia en el uso de ChatGPT y encuentran que la herramienta cumple con sus expectativas, aumenta su PU y su nivel de satisfacción. En particular, cuando ChatGPT les proporciona funciones y servicios que optimizan su eficiencia y rendimiento académico, facilitando el logro de sus objetivos de aprendizaje con menos esfuerzo, su satisfacción se ve reforzada. Asimismo, se encontró que tanto la PU como la PEU influyeron significativamente en la BI (Nja et al., 2023; Pillai et al., 2023; Chen et al., 2023). Esto sugiere que, cuando los estudiantes perciben que ChatGPT es útil, es más probable que desarrollen una actitud positiva hacia su uso, ya que lo ven como un recurso que optimiza el tiempo en la realización de tareas y mejora la calidad del aprendizaje. Del mismo modo, la PEU influye en la BI de ChatGPT, ya que cuando una plataforma es percibida como accesible y fácil de manejar, su adopción se vuelve más probable. Si los estudiantes consideran que el uso de ChatGPT no requiere un esfuerzo adicional de aprendizaje, es más factible que lo incorporen en su rutina académica.

Además, se confirmó que la PU y la PEU impactaron positivamente la satisfacción de los estudiantes, respaldando estudios previos (Kashive et al., 2020; Al-Fraihat et al., 2020; Yu et al., 2024). Esto indica que, cuando los estudiantes encuentran que ChatGPT les ofrece beneficios tangibles, su motivación para utilizarlo aumenta, lo que se traduce en mayores niveles de satisfacción. De manera similar, una interfaz intuitiva y fácil de manejar contribuye a mejorar la experiencia del usuario, elevando su satisfacción. Por último, se halló que la satisfacción tuvo un efecto positivo en la BI de ChatGPT, lo que confirma hallazgos previos en la literatura (Chen et al., 2020; Yu et al., 2024). Esto sugiere que, cuando los estudiantes se sienten satisfechos con ChatGPT en el contexto de sus actividades académicas, es más probable que desarrollen una intención favorable hacia su uso continuo. Un estudiante satisfecho tiende a percibir la herramienta como valiosa y beneficiosa, lo que incrementa su disposición a integrarla de manera regular en sus tareas de aprendizaje.

Implicaciones

Implicaciones teóricas

Este estudio aporta a la literatura al integrar el TAM con el ECM, incorporando la concienciación y la confianza como factores adicionales para evaluar la intención de los estudiantes de utilizar ChatGPT. Mientras que el TAM enfatiza la PU y la PEU, y el ECM se centra en la confirmación y satisfacción de los usuarios, la inclusión de la concienciación y la confianza proporciona un marco más completo. Este enfoque permite una comprensión más detallada de los factores que influyen en la aceptación de ChatGPT, yendo más allá de las variables originales de ambos modelos. El estudio también llena un vacío en la literatura, ya que la concienciación y la confianza se han discutido en investigaciones sobre adopción tecnológica, pero rara vez han sido analizadas conjuntamente dentro de un solo modelo. Teóricamente, se demuestra que la concienciación influye directamente en la PU y la PEU, mientras que la confianza impacta solo en la PU, lo que resalta los roles diferenciados de estos factores en la adopción de ChatGPT.

Además, aunque TAM y ECM han sido ampliamente utilizados en estudios sobre la adopción de tecnología, su aplicación en el contexto de herramientas de IA como ChatGPT es relativamente reciente. Por lo tanto, este estudio contribuye a la teoría al demostrar cómo la integración de estos modelos puede ayudar a comprender la adopción de tecnologías emergentes basadas en IA. Estos hallazgos pueden inspirar investigaciones futuras que exploren la aplicabilidad del TAM y el ECM en otras herramientas de IA. La combinación de ambos modelos también ha mejorado la capacidad predictiva del estudio, con un valor $R^2 = 0,59$, lo que indica un nivel de predicción sostenible de la BI de ChatGPT. Según Sarstedt et al. (2014), un R^2 de 0,19 es débil, 0,33 es moderado y 0,67 es fuerte. Esto demuestra que el modelo propuesto predice eficazmente la adopción de ChatGPT al considerar tanto el conocimiento previo de los estudiantes (concienciación) como su percepción de la fiabilidad del sistema (confianza), además de los factores clásicos del TAM y el ECM. En este sentido, la integración de ambos modelos representa un avance teórico al ofrecer una mejor explicación de la adopción de ChatGPT que el uso aislado de cualquiera de ellos.

Implicaciones prácticas

Este estudio también tiene implicaciones prácticas para educadores, desarrolladores y responsables de políticas educativas en la promoción del uso de ChatGPT en entornos de aprendizaje. Un aspecto clave es la necesidad de aumentar la concienciación sobre los beneficios de esta herramienta. Integrarla en las actividades de clase permitiría a los estudiantes experimentar directamente su utilidad, facilitando la comprensión de conceptos complejos, optimizando

el tiempo de estudio y proporcionando asistencia académica personalizada. Asimismo, los educadores pueden reforzar este proceso mediante talleres y sesiones informativas que resalten las ventajas de ChatGPT en el aprendizaje. La confianza en la herramienta también desempeña un papel fundamental en su adopción. Para fortalecerla, los docentes e instructores pueden compartir investigaciones sobre el uso de ChatGPT en educación, así como estudios de caso y testimonios de éxito. Esto no solo contribuiría a mejorar la PU de ChatGPT, sino que también fomentaría su uso continuo entre los estudiantes. Además, dado que la PU y la PEU influyen en la satisfacción de los usuarios, es crucial que los educadores recopilen retroalimentación para optimizar su integración en los cursos. Brindar recursos adicionales y apoyo constante facilitaría una mejor experiencia de uso, lo que, a su vez, incrementaría la satisfacción estudiantil.

Desde la perspectiva del desarrollo tecnológico, es fundamental mejorar la interfaz de usuario de ChatGPT para hacerla más intuitiva y accesible, lo que contribuiría a una mayor PEU. Simplificar las funciones de la herramienta podría no solo aumentar la satisfacción de los estudiantes, sino también fomentar una actitud más positiva hacia su adopción. Además, incorporar mecanismos de transparencia, como la explicación de cómo se generan las respuestas y la verificación de la fiabilidad de la información, ayudaría a fortalecer la confianza en el sistema. De manera complementaria, proporcionar a los docentes recursos educativos sobre las capacidades y limitaciones de ChatGPT, a través de tutoriales, guías y casos de uso, facilitaría una integración más efectiva en el entorno académico. Por otro lado, los responsables de políticas educativas pueden desempeñar un papel clave en la promoción del uso ético y responsable de ChatGPT. Implementar iniciativas para fomentar la alfabetización en IA dentro de las instituciones educativas contribuiría a mejorar tanto la confianza como el conocimiento sobre esta tecnología. Asimismo, el establecimiento de directrices claras para su aplicación en el ámbito educativo permitiría abordar inquietudes sobre la precisión del contenido, la privacidad de los datos y el papel de la IA en el aprendizaje. Estas políticas no solo mitigarían posibles preocupaciones, sino que también generarían un marco de confianza que favorecería la adopción de ChatGPT como una herramienta educativa eficaz. En conjunto, estas acciones pueden contribuir a la creación de un entorno propicio para la adopción de ChatGPT en la educación, asegurando que tanto docentes como estudiantes aprovechen su potencial de manera efectiva y ética.

Limitaciones y trabajos futuros

Este estudio presenta algunas limitaciones que deben considerarse al interpretar sus hallazgos. En primer lugar, la muestra se limitó a una única universidad, lo que restringe la generalización de los resultados a otros contextos académicos. La BI de ChatGPT puede variar según la disciplina, el entorno educativo, el sistema de enseñanza o los antecedentes culturales de los estudiantes, lo que sugiere la necesidad

de ampliar la investigación a diferentes instituciones y contextos. Además, el estudio adoptó un enfoque cuantitativo, lo que, si bien permitió obtener datos estadísticos sólidos, no exploró en profundidad las percepciones y experiencias individuales de los estudiantes. Futuros estudios podrían beneficiarse de un enfoque mixto que combine métodos cuantitativos y cualitativos, incorporando entrevistas y grupos de discusión para obtener una comprensión más detallada de los factores que influyen en la adopción de ChatGPT.

Por otro lado, aunque este estudio se centró en la concienciación y la confianza como determinantes clave en la BI de ChatGPT, existen otros factores que podrían desempeñar un papel relevante. Investigaciones futuras podrían examinar variables como el apoyo institucional, la experiencia previa con ChatGPT y la influencia social, con el fin de ampliar el modelo y proporcionar una visión más completa sobre la adopción de esta tecnología en el ámbito educativo.

Conclusión

Este estudio proporciona información valiosa sobre los factores que influyen en la intención de los estudiantes de utilizar ChatGPT para el aprendizaje, integrando el modelo TAM extendido (con la inclusión de la concienciación y la confianza) y el modelo ECM. Los hallazgos destacan la importancia de la concienciación en la percepción de PU y la PEU, mientras que la confianza influye significativamente en la PU, aunque no en la PEU. Además, la PU, la PEU y la confirmación contribuyen a la satisfacción de los estudiantes, lo que a su vez fortalece su intención de adoptar ChatGPT como una herramienta de apoyo académico. Estos resultados subrayan la necesidad de fomentar el conocimiento sobre ChatGPT y generar confianza en su uso para maximizar su adopción y efectividad en el aprendizaje. Al considerar estos factores, educadores, desarrolladores y responsables de políticas educativas pueden diseñar estrategias que faciliten la integración de esta tecnología en la educación, promoviendo así un entorno de aprendizaje más dinámico e innovador.

REFERENCIAS

- Abdalla, R. A. (2024). Examining awareness, social influence, and perceived enjoyment in the TAM framework as determinants of ChatGPT: Personalization as a moderator. *Journal of Open Innovation: Technology, Market, and Complexity*, 10(3), 100327. <https://doi.org/10.1016/j.joitmc.2024.100327>
- Alfthanorhan, A., Ghazali, P. L. y Rashid, N. (2021, May). Discriminant validity: A comparison of CBSEM and consistent PLS using Fornell & Larcker and HTMT approaches. *Journal of Physics: Conference Series*, 1874(1), 012085. <https://doi.org/10.1088/1742-6596/1874/1/012085>
- Albayati, H. (2024). Investigating undergraduate students' perceptions and awareness of using ChatGPT as a regular assistance tool: A user acceptance perspective study. *Computers and Education: Artificial Intelligence*, 6,

- Article 100203. <https://doi.org/10.1016/j.caeai.2024.100203>
- Al-Fraihat, D., Joy, M. y Sinclair, J. (2020). Evaluating e-learning systems success: An empirical study. *Computers in Human Behavior*, 102, 67-86. <https://doi.org/10.1016/j.chb.2019.08.004>
- 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>
- Almuraqab, N. A. S. (2020). Shall universities in the UAE continue distance learning after the COVID-19 pandemic? Revealing students' perspective. *Social Science Research Network*.
- Al-Sayid, F. y Kirkil, G. (2023). Students' web-based activities moderate the effect of human-computer interaction factors on their e-learning acceptance and success during the COVID-19 pandemic. *International Journal of Human-Computer Interaction*, 39(14), 2852-2875. <https://doi.org/10.1080/10447318.2022.2087013>
- Alshammari, S. H. y Alshammari, R. A. (2024). An integration of expectation confirmation model and information systems success model to explore the factors affecting the continuous intention to utilise virtual classrooms. *Scientific Reports*, 14(1), Article 18491. <https://doi.org/10.1038/s41598-024-69401-8>
- Alshammari, S. H. y Alshammari, M. H. (2024). Factors affecting the adoption and use of ChatGPT in higher education. *International Journal of Information and Communication Technology Education (IJICTE)*, 20(1), 1-16. <https://doi.org/10.4018/IJICTE.339557>
- Al-Sharafi, M. A., Al-Emran, M., Iranmanesh, M., Al-Qaysi, N., Iahad, N. A. y Arpaci, I. (2023). Understanding the impact of knowledge management factors on the sustainable use of AI-based chatbots for educational purposes using a hybrid SEM-ANN approach. *Interactive Learning Environments*, 31(10), 7491-7510. <https://doi.org/10.1080/10494820.2022.2075014>
- Arpaci, I. (2016). Understanding and predicting students' intention to use mobile cloud storage services. *Computers in Human Behavior*, 58, 150-157. <https://doi.org/10.1016/j.chb.2015.12.067>
- Ashraf, R. U., Hou, F. y Ahmad, W. (2019). Understanding continuance intention to use social media in China: The roles of personality drivers, hedonic value, and utilitarian value. *International Journal of Human-Computer Interaction*, 35(13), 1216-1228. <https://doi.org/10.1080/10447318.2018.1519145>
- Ashrafi, A., Zareravasan, A., Rabiee Savoji, S. y Amani, M. (2022). Exploring factors influencing students' continuance intention to use the learning management system (LMS): A multi-perspective framework. *Interactive Learning Environments*, 30(8), 1475-1497. <https://doi.org/10.1080/10494820.2020.1734028>
- Awang, P. (2015). *SEM made simple: A gentle approach to learning structural equation modeling*. MPWS Rich Publication.
- Baek, T. H. y Kim, M. (2023). Is ChatGPT scary good? How user motivations affect creepiness and trust in generative artificial intelligence. *Telematics and Informatics*, 83, 102030. <https://doi.org/10.2139/ssrn.4416087>
- Baidoo-Anu, D. y Ansah, L. O. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI*, 7(1), 52-62. <https://doi.org/10.61969/jai.1337500>
- Bamigbola, A. A. y Adetimirin, A. E. (2020). Assessing determinants of perceived ease of use of institutional repositories by lecturers in Nigerian universities.

- International Information & Library Review*, 52(2), 95-107. <https://doi.org/10.1080/10572317.2019.1662261>
- Bansal, H. y Khan, R. (2018). A review paper on human computer interaction. *International Journal of Advanced Research in Computer Science and Software Engineering*, 8(4), 53. <https://doi.org/10.23956/ijarcsse.v8i4.630>
- Basak, E., Gumussoy, C. A. y Calisir, F. (2015). Examining the factors affecting PDA acceptance among physicians: An extended technology acceptance model. *Journal of Healthcare Engineering*, 6(3), 399-418. <https://doi.org/10.1260/2040-2295.6.3.399>
- Bernabei, M., Colabianchi, S., Falegnami, A. y Costantino, F. (2023). Students' use of large language models in engineering education: A case study on technology acceptance, perceptions, efficacy, and detection chances. *Computers and Education: Artificial Intelligence*, 5, 100172. <https://doi.org/10.1016/j.caeai.2023.100172>
- Bhattacharjee, A. (2001). Understanding information systems continuance: An expectation-confirmation model. *MIS Quarterly*, 25, 351-370. <https://doi.org/10.2307/3250921>
- Bin-Nashwan, S. A., Sadallah, M. y Bouteraa, M. (2023). Use of ChatGPT in academia: Academic integrity hangs in the balance. *Technology in Society*, 75, 102370. <https://doi.org/10.1016/j.techsoc.2023.102370>
- Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., ... Amodei, D. (2020). Language models are few-shot learners [Preprint]. arXiv. <https://doi.org/10.48550/arXiv.2005.14165>
- Cao, Y., Li, S., Liu, Y., Yan, Z., Dai, Y., Yu, P. S. y Sun, L. (2023). A comprehensive survey of AI-generated content (AIGC): A history of generative AI from GAN to ChatGPT [Preprint]. arXiv. <https://arxiv.org/abs/2303.04226>
- Chen, C. C., Hsiao, K. L. y Li, W. C. (2020). Exploring the determinants of usage continuance willingness for location-based apps: A case study of bicycle-based exercise apps. *Journal of Retailing and Consumer Services*, 55, 102097. <https://doi.org/10.1016/j.jretconser.2020.102097>
- Chen, S., Qiu, S., Li, H., Zhang, J., Wu, X., Zeng, W. y Huang, F. (2023). An integrated model for predicting pupils' acceptance of artificially intelligent robots as teachers. *Education and Information Technologies*, 28(9), 11631-11654. <https://doi.org/10.1007/s10639-023-11601-2>
- Choi, G., Nam, C. y Kim, S. (2019). The impacts of technology platform openness on application developers' intention to continuously use a platform: From an ecosystem perspective. *Telecommunications Policy*, 43(2), 140-153. <https://doi.org/10.1016/j.telpol.2018.04.003>
- Choung, H., David, P. y Ross, A. (2023). Trust in AI and its role in the acceptance of AI technologies. *International Journal of Human-Computer Interaction*, 39(9), 1727-1739. <https://doi.org/10.1080/10447318.2022.2050543>
- Cohen, J. (1988). *Statistical power for the behavioural sciences*. Lawrence Erlbaum.
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319-340. <https://doi.org/10.2307/249008>
- Emenike, M. E. y Emenike, B. U. (2023). Was this title generated by ChatGPT? Considerations for artificial intelligence text-generation software programs for chemists and chemistry educators.

- Journal of Chemical Education*, 100(4), 1413-1418. <https://doi.org/10.1021/acs.jchemed.3c00063>
- Eppler, M., Ganjavi, C., Ramacciotti, L. S., Piazza, P., Rodler, S., Checcucci, E., Rivas, J. G., Kowalewski, K. F., Belenchón, I. R., Puliaatti, S., Taratkin, M., Veccia, A., Baekelandt, L., Teoh, J. Y.-C., Somani, B. K., Wroclawski, M., Abreu, A., Porpiglia, F., Gill, I. S., ... Cacciamani, G. E. (2024). Awareness and Use of ChatGPT and Large Language Models: A Prospective Cross-sectional Global Survey in Urology. *European Urology*, 85(2), 146-153. <https://doi.org/10.1016/j.eururo.2023.10.014>
- Esposito, C., Tamburis, O., Su, X. y Choi, C. (2020). Robust decentralised trust management for the internet of things by using game theory. *Information Processing & Management*, 57(6), 102308. <https://doi.org/10.1016/j.ipm.2020.102308>
- Flavián, C., Pérez-Rueda, A., Belanche, D. y Casaló, L. V. (2022). Intention to use analytical artificial intelligence (AI) in services-the effect of technology readiness and awareness. *Journal of Service Management*, 33(2), 293-320. <https://doi.org/10.1108/JOSM-10-2020-0378>
- Folkinshteyn, D. y Lennon, M. (2016). Braving Bitcoin: A technology acceptance model (TAM) analysis. *Journal of Information Technology Case and Application Research*, 18(4), 220-249. <https://doi.org/10.1080/15228053.2016.1275242>
- Frieder, S., Pinchetti, L., Griffiths, R. R., Salvatori, T., Lukasiewicz, T., Petersen, P. y Berner, J. (2023). Mathematical capabilities of ChatGPT. En *Advances in Neural Information Processing Systems* (Vol. 36).
- Hair, J. F., Jr., Black, W. C., Babin, B. J., Anderson, R. E. y Tatham, R. L. (2010). *Multivariate data analysis* (7th ed.). Pearson Prentice Hall.
- Haleem, A., Javaid, M. y Singh, R. P. (2022). An era of ChatGPT as a significant futuristic support tool: A study on features, abilities, and challenges. *Benchmark Transactions on Benchmarks, Standards and Evaluations*, 2(4), 100089. <https://doi.org/10.1016/j.tbench.2023.100089>
- Hong, S., Thong, J. Y. y Tam, K. Y. (2006). Understanding continued information technology usage behavior: A comparison of three models in the context of mobile internet. *Decision Support Systems*, 42(3), 1819-1834. <https://doi.org/10.1016/j.dss.2006.03.009>
- Huang, X. y Zhi, H. (2023). Factors influencing students' continuance usage intention with virtual classroom during the COVID-19 pandemic: An empirical study. *Sustainability*, 15(5), 4420. <https://doi.org/10.3390/su15054420>
- Ifinedo, P. (2007). Investigating the antecedents of continuance intention of course management systems use among Estonian undergraduates. *International Journal of Information and Communication Technology Education (IJICTE)*, 3(4), 76-92. <https://doi.org/10.4018/jicte.2007100107>
- Iku-Silan, A., Hwang, G. J. y Chen, C. H. (2023). Decision-guided chatbots and cognitive styles in interdisciplinary learning. *Computers & Education*, 201, 104812. <https://doi.org/10.1016/j.compedu.2023.104812>
- Ismail, A. R. A., Fujita, S., Sakamoto, K., Takahashi, H. y Panessai, I. Y. (2024). Emerging trends and challenges in chatbot technology: A Japanese industry perspective. *International Journal of Recent Technology and Applied Science (IJORTAS)*, 6(1), 46-60. <https://doi.org/10.36079/lamintang.ijortas-0601.708>
- Joo, S. y Choi, N. (2016). Understanding users' continuance intention to use online library resources based on an extended expectation-confirmation model. *The*

- Electronic Library*, 34(4), 554-571. <https://doi.org/10.1108/EL-02-2015-0033>
- Joo, Y. J., Park, S. y Shin, E. K. (2017). Students' expectation, satisfaction, and continuance intention to use digital textbooks. *Computers in Human Behavior*, 69, 83-90. <https://doi.org/10.1016/j.chb.2016.12.025>
- Kashive, N., Powale, L. y Kashive, K. (2020). Understanding user perception toward artificial intelligence (AI) enabled e-learning. *The International Journal of Information and Learning Technology*, 38(1), 1-19. <https://doi.org/10.1108/IJILT-05-2020-0090>
- Kelly, S., Kaye, S. A. y Oviedo-Trespalacios, O. (2023). What factors contribute to the acceptance of artificial intelligence? A systematic review. *Telematics and Informatics*, 77, 101925. <https://doi.org/10.1016/j.tele.2022.101925>
- Kline, R. B. (2023). *Principles and practice of structural equation modeling*. Guilford Publications.
- Li, K. (2023). Determinants of college students' actual use of AI-based systems: An extension of the technology acceptance model. *Sustainability*, 15(6), 5221. <https://doi.org/10.3390/su15065221>
- Limayem, M. y Cheung, C. M. (2008). Understanding information systems continuance: The case of Internet-based learning technologies. *Information & Management*, 45(4), 227-232. <https://doi.org/10.1016/j.im.2008.02.005>
- Liu, K. y Tao, D. (2022). The roles of trust, personalization, loss of privacy, and anthropomorphism in public acceptance of smart healthcare services. *Computers in Human Behavior*, 127, 107026. <https://doi.org/10.1016/j.chb.2021.107026>
- Lowry, P. B. y Gaskin, J. (2014). Partial least squares (PLS) structural equation modeling (SEM) for building and testing behavioral causal theory: When to choose it and how to use it. *IEEE Transactions on Professional Communication*, 57(2), 123-146. <https://doi.org/10.1109/TPC.2014.2312452>
- Lund, B. D. y Wang, T. (2023). Chatting about ChatGPT: How may AI and GPT impact academia and libraries? *Library Hi Tech News*, 40(3), 26-29. <https://doi.org/10.1108/LHTN-01-2023-0009>
- Maheshwari, G. (2023). Factors influencing students' intention to adopt and use ChatGPT in higher education: A study in the Vietnamese context. *Education and Information Technologies*, 1-29. <https://doi.org/10.1007/s10639-023-12333-z>
- Mohr, S. y Kühl, R. (2021). Acceptance of artificial intelligence in German agriculture: An application of the technology acceptance model and the theory of planned behavior. *Precision Agriculture*, 22(6), 1816-1844. <https://doi.org/10.1007/s11119-021-09814-x>
- Mouakket, S. (2015). Factors influencing continuance intention to use social network sites: The Facebook case. *Computers in Human Behavior*, 53, 102-110. <https://doi.org/10.1016/j.chb.2015.06.045>
- Mutahar, A. M., Daud, N. M., Thurasamy, R., Isaac, O. y Abdulsalam, R. (2018). The mediating of perceived usefulness and perceived ease of use: The case of mobile banking in Yemen. *International Journal of Technology Diffusion*, 9(2), 21-40. <https://doi.org/10.4018/IJTD.2018040102>
- Nikou, S. A. y Economides, A. A. (2017). Mobile-based assessment: Investigating the factors that influence behavioral intention to use. *Computers & Education*, 109, 56-73. <https://doi.org/10.1016/j.compedu.2017.02.005>
- Niu, B. y Mvondo, G. F. N. (2024). I am ChatGPT, the ultimate AI Chatbot! Investigating the determinants of users' loyalty and ethical usage concerns of ChatGPT. *Journal of Retailing and*

- Consumer Services, 76, 103562. <https://doi.org/10.1016/j.jretconser.2023.103562>
- Nja, C. O., Idiege, K. J., Uwe, U. E., Meremikwu, A. N., Ekon, E. E., Erim, C. M., Ukah, J. U., Eyo, E. O., Anari, M. I. y Cornelius-Ukpepi, B. U. (2023). Adoption of artificial intelligence in science teaching: From the vantage point of the African science teachers. *Smart Learning Environments*, 10(42). <https://doi.org/10.1186/s40561-023-00261-x>
- OpenAI. (2023). *GPT-4 technical report*. Cornell University. <https://doi.org/10.48550/arXiv.2303.08774>
- Park, E. (2020). User acceptance of smart wearable devices: An expectation-confirmation model approach. *Telematics and Informatics*, 47, 101318. <https://doi.org/10.1016/j.tele.2019.101318>
- Patel, S. B. y Lam, K. (2023). ChatGPT: The future of discharge summaries? *The Lancet Digital Health*, 5(3), e107-e108. [https://doi.org/10.1016/S2589-7500\(23\)00021-3](https://doi.org/10.1016/S2589-7500(23)00021-3)
- Pillai, R., Sivathanu, B., Metri, B. y Kaushik, N. (2023). Students' adoption of AI-based teacher-bots (T-bots) for learning in higher education. *Information Technology & People*. Advance online publication. <https://doi.org/10.1108/ITP-02-2021-0152>
- Polyportis, A. (2024). A longitudinal study on artificial intelligence adoption: Understanding the drivers of ChatGPT usage behavior change in higher education. *Frontiers in Artificial Intelligence*, 6, 1324398. <https://doi.org/10.3389/frai.2023.1324398>
- Rahi, S., Alghizzawi, M. y Ngh, A. H. (2023). Factors influence user's intention to continue use of e-banking during COVID-19 pandemic: The nexus between self-determination and expectation confirmation model. *EuroMed Journal of Business*, 18(3), 380-396. <https://doi.org/10.1108/EMJB-12-2021-0194>
- Rahlina, N. A., Awangb, Z., Afthanorhanc, A. y Aimrand, N. (2019). The art of covariance based analysis in behaviour-based safety performance study using confirmatory factor analysis: Evidence from SMEs. *Measurement*, 7(10).
- 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*, 51, 51-71. <https://doi.org/10.14742/ajet.8956>
- Ramadhan, A., Hidayanto, A. N., Salsabila, G. A., Wulandari, I., Jaury, J. A. y Anjani, N. N. (2022). The effect of usability on the intention to use the e-learning system in a sustainable way: A case study at Universitas Indonesia. *Education and Information Technologies*, 1-34.
- Rudolph, J., Tan, S. y Tan, S. (2023). ChatGPT: Bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning and Teaching*, 6(1), 342-363. <https://doi.org/10.37074/jalt.2023.6.1.9>
- Sarstedt, M., Ringle, C. M. y Hair, J. F. (2014). PLS-SEM: Looking back and moving forward. *Long Range Planning*, 47(3), 132-137. <https://doi.org/10.1016/j.lrp.2014.02.008>
- Sarstedt, M., Ringle, C. M. y Hair, J. F. (2021). Partial least squares structural equation modeling. En *Handbook of Market Research* (pp. 587-632). Springer International Publishing. https://doi.org/10.1007/978-3-319-57413-4_15
- Schulman, J., Zoph, B., Kim, C., Hilton, J., Menick, J., Weng, J., Ceron Uribe, J. F., Fedus, L., Metz, L., Pokorny, M., Gontijo Lopes, R., Zhao, S., Vijayvergiya, A., Sigler, E., Perelman, A., Voss, C., Heaton, M., Parish, J., ... Hesse, C. (2022, November 30). *Optimizing language models for dialogue*. OpenAI.
- Sekaran, U. y Bougie, R. (2016). *Research Methods for Business: A Skill-Building*

- Approach*. 7th Edition, Wiley & Sons, West Sussex.
- Sestino, A. y D'Angelo, A. (2023). My doctor is an avatar! The effect of anthropomorphism and emotional receptivity on individuals' intention to use digital-based healthcare services. *Technological Forecasting and Social Change*, 191, Article 122505. <https://doi.org/10.1016/j.techfore.2023.122505>
- Shuhaiber, A. y Mashal, I. (2019). Understanding users' acceptance of smart homes. *Technology in Society*, 58, 101110. <https://doi.org/10.1016/j.techsoc.2019.01.003>
- Stone, R. W. y Baker-Eveleth, L. (2013). Students' expectation, confirmation, and continuance intention to use electronic textbooks. *Computers in Human Behavior*, 29(3), 984-990. <https://doi.org/10.1016/j.chb.2012.12.007>
- Tam, C., Santos, D. y Oliveira, T. (2020). Exploring the influential factors of continuance intention to use mobile apps: Extending the expectation confirmation model. *Information Systems Frontiers*, 22, 243-257. <https://doi.org/10.1007/s10796-018-9864-5>
- Torrent-Sellens, J., Jiménez-Zarco, A. I. y Saigí-Rubió, F. (2021). Do people trust in robot-assisted surgery? Evidence from Europe. *International Journal of Environmental Research and Public Health*, 18(23), 12519. <https://doi.org/10.3390/ijerph182312519>
- Tung, L. (2023). ChatGPT can write code. Now researchers say it's good at fixing bugs, too. *ZDNet*.
- van Dis, E. A., Bollen, J., Zuidema, W., Van Rooij, R. y Bockting, C. L. (2023). ChatGPT: Five priorities for research. *Nature*, 614(7947), 224-226. <https://doi.org/10.1038/d41586-023-00288-7>
- Venkatesh, V. y Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science*, 46(2), 186-204. <https://doi.org/10.1287/mnsc.46.2.186.11926>
- Wandira, R., Fauzi, A. y Nurahim, F. (2024). Analysis of factors influencing behavioral intention to use cloud-based academic information system using extended technology acceptance model (TAM) and expectation-confirmation model (ECM). *Journal of Information Systems Engineering and Business Intelligence*, 10(2), 179-190. <https://doi.org/10.20473/jisebi.10.2.179-190>
- Xu, J., Zhang, X., Li, H., Yoo, C. y Pan, Y. (2023). Is everyone an artist? A study on user experience of AI-based painting system. *Applied Sciences*, 13(11), 6496. <https://doi.org/10.3390/app13116496>
- Yu, C., Yan, J. y Cai, N. (2024). ChatGPT in higher education: Factors influencing ChatGPT user satisfaction and continued use intention. *Frontiers in Education*, 9, Article 1354929. <https://doi.org/10.3389/educ.2024.1354929>
- Yusoff, M. N. H. B., Zainol, F. A., Hafifi Ridzuan, R., Ismail, M. y Afthanorhan, A. (2021). Psychological traits and intention to use e-commerce among rural micro-entrepreneurs in Malaysia. *Journal of Theoretical and Applied Electronic Commerce Research*, 16(5), 1827-1843. <https://doi.org/10.3390/jtaer16050102>

Fecha de recepción del artículo: 1 de diciembre de 2024

Fecha de aceptación del artículo: 25 de febrero de 2025

Fecha de aprobación para maquetación: 29 de abril de 2025

Fecha de publicación en OnlineFirst: 7 de abril de 2025

Fecha de publicación: 1 de julio de 2025

Marcos normativos para una IA ética y confiable en la educación superior: estado de la cuestión

Normative framework for ethical and trustworthy AI in higher education: state of the art



- id María Obdulia González-Fernández - *Universidad de Guadalajara, UDG (México)*
id María Asunción Romero-López - *Universidad de Granada, UGR (España)*
id Natalia Fátima Sgreccia - *Universidad Nacional de Rosario, UNR (Argentina)*
id María José Latorre Medina - *Universidad de Granada, UGR (España)*

RESUMEN

La investigación sobre el uso de la inteligencia artificial (IA) en instituciones educativas es un área emergente y en evolución. En la actualidad se ve la necesidad de analizar la sinergia entre la ética y el adecuado uso de la IA en el contexto académico. Por ende, se presenta un estudio documental para indagar sobre los marcos normativos y éticos del uso de la IA en las instituciones de educación superior. Se realizó una revisión sistemática de literatura bajo las directrices y procedimientos del modelo PRISMA. La búsqueda de la información se realizó en las bases de datos de Scielo, Scopus, Redalyc, entre otras, que posteriormente fueron exportados a Mendeley y procesados en Rayyan con la intención de analizarlos y categorizarlos. En el proceso de cribación se seleccionaron 28 artículos clasificados en cuatro categorías: desafíos y riesgos éticos, marcos normativos, formación ética y modelos didácticos. Se utilizó el software de MAXQDA para su procesamiento. Además, se realizó un análisis bibliométrico (VOSviewer) que permitió conocer las principales tendencias en torno a la temática. En general, los resultados manifiestan, por un lado, la importancia de generación de políticas y regulaciones específicas que permitan el uso adecuado de la IA en contextos universitarios. Por otro, la necesidad de formación de toda la comunidad educativa involucrada, incorporando su uso responsable. El debate académico sobre cómo afrontar el desafío de equilibrar la innovación tecnológica con la ética, la inclusión y la sostenibilidad continúa abierto.

Palabras clave: inteligencia artificial; educación superior; ética; normativa.

ABSTRACT

The use of artificial intelligence (AI) in higher education represents an emerging and continuously evolving field. Current research highlights the critical synergy required between ethics and appropriate use in the academic application of AI. Therefore, a documentary study is presented to investigate the regulatory and ethical frameworks of the use of AI in higher education institutions. A systematic literature review was conducted following the guidelines and procedures of the PRISMA model. The information search was carried out in the databases of Scielo, Scopus, Redalyc, among others, which were subsequently exported to *Mendeley* and *Rayyan* platforms for analysis and categorization. In the screening process, 28 articles were selected, classified into four categories: ethical challenges and risks, regulatory frameworks, ethical training and didactical models, utilizing the *MAXQDA* software. Additionally, a bibliometric analysis (*VOSviewer*) was performed to identify major research trends in this domain. In general, the results show, on the one hand, the importance of generating specific policies and regulations that allow the appropriate use of AI in university contexts. On the other hand, the need for training of the entire educational community involves incorporating its responsible use. The academic debate on how to balance technological innovation with ethics, inclusion, and sustainability remains an open and pressing challenge.

Keywords: artificial intelligence; higher education; ethics; legislations.

INTRODUCCIÓN

La literatura académica reciente pone de manifiesto el papel clave que juegan las tecnologías en el ámbito de la enseñanza superior (García-Fuentes et al., 2024; Villegas y Delgado, 2024), y el uso que hacen de ellas tanto docentes (Sánchez, 2023) como estudiantes universitarios (Baldrich y Domínguez, 2024). En concreto, centra la atención en las tecnologías emergentes e informa del impacto de la IA en el campo educativo; desde la perspectiva de los Objetivos de Desarrollo Sostenible (ODS) de la Agenda 2030 de la UNESCO, se describen las oportunidades de uso por parte del estudiantado y profesorado. Concretamente, “el ODS4 destaca que las tecnologías de IA se utilizan para garantizar un acceso equitativo e inclusivo a la educación” (Flores y García, 2023, p. 38).

No obstante, cabe plantearse si los procesos educativos en el plano académico han sido capaces de aprovechar todo el potencial que brindan las tecnologías. Si no ha sido así, ¿qué lo ha imposibilitado?, ¿la infraestructura, la falta de recursos tecnológicos, la formación del profesorado, la aptitud del alumnado...? El actual contexto universitario, donde las herramientas de IA han cobrado fuerza (García Peñalvo et al., 2024; López-Regalado et al., 2024), no puede permanecer impasible al respecto, sino que ha de posicionarse y adoptar medidas. El binomio educación universitaria-tecnología IA convoca a ello.

El presente trabajo se enmarca en esta línea de estudio. En una investigación de mayor envergadura desarrollada por investigadores iberoamericanos, que pretende analizar el uso de la IA en la formación práctica universitaria, de distintas titulaciones del campo de las Ciencias Sociales, Ingenierías y Administrativas, desde un enfoque crítico, ético y responsable. La investigación se titula *Prácticum IA: Experiencias del Prácticum integrando Inteligencia Artificial* (bienio 2023-2025) y se está desarrollando en el marco de una comunidad de conocimiento profesional, RedTICPraxis. Uno de los objetivos iniciales está relacionado con el análisis documental sobre marcos normativos y éticos del uso de la IA en la educación superior (ES).

Ética, Inteligencia Artificial y Educación Superior

La IA en la ES representa un asunto novedoso y de actualización permanente. Las investigaciones subrayan la sinergia entre ética y responsabilidad, así como privacidad y seguridad, al emplear IA en contextos académicos (Mora Naranjo et al., 2023). En ese marco emerge la necesidad de generar políticas y regulaciones específicas en los niveles institucionales correspondientes (Rodríguez-Argueta, 2020), así como actualizaciones que contemplen competencias digitales y conciencia ética en los planes y programas de estudio (Vélez-Rivera et al., 2024).

Las respuestas y/o reacciones que esta tecnología ha provocado en el profesorado, en sus percepciones, valoraciones y posicionamientos, han sido abordadas en

diferentes estudios. En concreto, los trabajos realizados por Humble y Mozeliuss (2022), McGrath et al. (2023) y Svaricek (2024) muestran un docente universitario que alberga cierto escepticismo sobre los beneficios de la IA para la enseñanza, y con preocupaciones arraigadas en incertidumbres sobre su impacto en el contexto educativo. El desconocimiento de la IA, sobre todo de sus aplicaciones prácticas y potencialidades como recurso didáctico (Sánchez, 2023), está contribuyendo a que el profesorado se sienta inseguro y adopte una postura cautelosa, que ha frenado su inclusión en ámbitos específicos de docencia, investigación y/o gestión universitaria.

En la otra esfera, están los estudiantes universitarios. También se muestran reacios a utilizar herramientas IA por temor a la estigmatización social y a ser acusados de deshonestidad académica. El entorno académico tradicionalmente ha ponderado los esfuerzos y logros individuales, y a algunos estudiantes les preocupa que emplear la IA pueda percibirse como desinterés, comodidad, falta de compromiso, capacidad de esfuerzo o la autenticidad de su trabajo. En medio de este escenario, han empezado a surgir experiencias de integración de la IA en diseños curriculares de la ES (Gairín y Alguacil, 2024). Las directrices sugeridas por González Alcaide (2024) brindan pistas para su integración en la práctica docente. La clave parece estar en el conocimiento y la formación en/para un uso efectivo, ético y responsable de estas herramientas (Mora Naranjo et al., 2023), que inspire confianza e invite a utilizarlas abiertamente en el ámbito universitario, dando prioridad a las consideraciones éticas a la hora de adoptar tecnologías de IA (Vélez-Rivera et al., 2024) para garantizar la equidad, transparencia y privacidad de los datos, con la premisa de contribuir a una sociedad más justa e inclusiva.

Desde la segunda coordenada, existencia de marcos normativos sobre el uso de la IA, la investigación alerta que, en el contexto académico actual, es posible que algunas instituciones de ES carezcan aún de directrices claras sobre el uso de la IA en el trabajo académico, dirigidas tanto al alumnado como al profesorado. El estudio de Seldon et al. (2020) resulta revelador sobre el impacto que tiene en la comunidad académica el uso desregulado de la IA. Si bien la literatura es todavía escasa, se advierte la necesidad e importancia de generar políticas, regulaciones y normativas éticas específicas para el uso responsable de la IA en los planes y programas educativos (Nguyen et al., 2023; Rodríguez-Argueta, 2020; Vélez et al., 2024), con el propósito de aprovechar las ventajas que esta herramienta puede aportar a las trayectorias académicas y educativas.

De cara a dar respuesta a este reto, se puede tomar como referente el trabajo de Chan (2023), en el que se recoge un marco político de educación ecológica para la IA bajo tres dimensiones: pedagógica, de gobernabilidad y operativa. El autor recomienda investigar la inclusión de la IA en la enseñanza universitaria a través de este marco ético para mejorar la comprensión de sus implicaciones. En una línea similar, en el estudio de Flores y García (2023, p. 45) se insta al desarrollo de marcos de políticas públicas consensuadas para regular y concienciar el uso ético de la IA en la educación. Según ambos autores:

Esto debe pasar por el diseño y puesta en marcha de un Plan de Alfabetización digital y de IA, considerando que el principal obstáculo para la adopción generalizada de herramientas IA es la falta de evidencia sólida sobre su eficacia y su impacto en los logros académicos de los estudiantes. De ahí que resulte necesario mejorar la gobernanza, la accesibilidad y la confiabilidad de la IA.

Desde la tercera coordenada, capacitación en materia de ética digital, se advierten algunas consideraciones a tener en cuenta. Es cierto que la investigación subraya la importancia crítica de la ética y la responsabilidad en la aplicación de la IA (Mora Naranjo et al., 2023), pero ¿Están los estudiantes y el profesorado capacitados al respecto?, ¿Qué uso hacen de la IA?, ¿Es un uso crítico, ético y responsable? En definitiva, ¿son competentes en materia de ética digital? Los estudios consultados instan a fortalecer los planes de formación docentes en materia de ética digital (Marín y Tur, 2024), e incluso a valorar el impacto que tales planes van teniendo en el alumnado (Mora Naranjo et al., 2023); realmente, ¿están promoviendo un uso responsable de las herramientas de IA? En este punto, el trabajo de Vélez-Rivera et al. (2024) apela a la necesidad de trabajar con la comunidad académica, profesorado y alumnado, en la utilización de estas herramientas, en correspondencia con entornos académicos seguros, éticos y responsables.

Ante este desafío, la investigación subraya la relevancia de la capacitación ética digital del profesorado universitario. Más aún, González Alcaide (2024) presenta un modelo de alfabetización digital docente, mientras que Flores y García (2023, p. 45) hablan de un “Plan de Alfabetización IA que permita formar a los docentes no solo en capacidades técnicas, sino en debates éticos-filosóficos”. Por otro lado, la investigación avala la necesidad de preparar al estudiantado para afrontar los retos éticos asociados a la IA. Por ejemplo, Lopezosa et al. (2023) plantean un modelo formativo nuclear en IA adaptado al ámbito educativo. Incluye competencias, objetivos y temas de estudio esenciales sobre IA, y un esquema que puede formar parte de un plan de estudios más amplio, para grado/posgrado o seminarios especializados. En una línea similar, González Alcaide (2024) desarrolla una unidad didáctica dirigida al alumnado para promover un uso adecuado de la IA, así como un rendimiento óptimo de estas herramientas. La propuesta de Aler et al. (2024) también va en este sentido. Sucintamente, focaliza su atención en las estrategias educativas, competencias éticas y recursos necesarios para una implementación de la IA, digna de confianza en las distintas disciplinas de la ES. Su objetivo es que la próxima generación de jóvenes pueda contribuir a la utilización de una IA ética, segura y de vanguardia. ¿Es posible establecer prácticas educativas justas, éticas y sostenibles en la era de la IA en el contexto académico? ¿Bajo qué normativa? La tríada IA-ética-ES, vista desde cada una de estas coordenadas, nos sitúa ante un panorama alentador pero necesitado aún de estudio y análisis.

Objetivo del estudio

El presente trabajo tiene como objetivo analizar los estudios recientes sobre los marcos normativos y éticos en torno al uso de la IA en las instituciones de ES.

METODOLOGÍA

Acorde al objetivo formulado, se ha llevado a cabo una revisión de la literatura a través de un método sistemático para identificar, evaluar e interpretar los artículos científicos (Mendoza, 2021). En las revisiones sistemáticas se trata de responder a una interrogante bien delimitada a partir del análisis y discusión de la información más relevante de las publicaciones revisadas (Torres-Fonseca y López-Hernández, 2014). Para esta revisión se emplearon las directrices y procedimientos de la declaración PRISMA (*Preferred Reporting Items for Systematic reviews and Meta-Analyses*) actualizada al 2020 (Page et al., 2021). Como primer paso, conforme al objetivo del estudio, se definieron las siguientes preguntas de investigación:

- PI1. ¿Cuáles son los desafíos y riesgos éticos que se han detectado en las universidades?
- PI2. ¿Cuáles son las medidas normativas y éticas que han aplicado o sugerido las universidades en torno a la IA?
- PI3. ¿Cómo se interpreta el uso ético de la IA en los procesos educativos?
- PI4. ¿Cuáles son los enfoques o modelos didácticos que se han implementado en las universidades para el uso confiable de la IA?

Estrategia de búsqueda

La búsqueda de literatura se realizó en las bases de datos de *Scielo*, *Redalyc*, *Scopus*, *Dialnet*, *Scopus* y *Web of Science*. Asimismo, se emplearon herramientas de IA, como Consensus, para la identificación de referencias. La ecuación de búsqueda para obtener resultados significativos fue: (“challenges” AND “promises of AI”); (“ethics AND “IA” AND “higher education”); (“ethical considerations” AND “AI”); (“practicum” AND “AI”).

Criterios de inclusión y exclusión de documentos

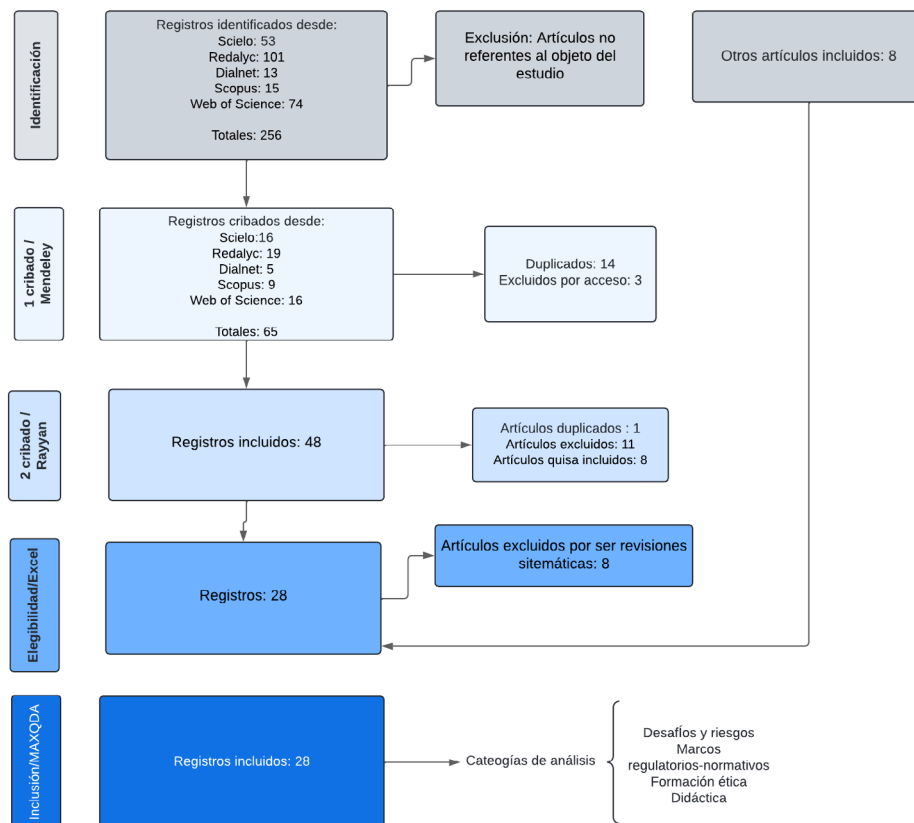
Los criterios principales de inclusión de los documentos fueron: artículos publicados del 2020 a fines de mayo de 2024; idioma español, inglés o portugués; producciones que recojan una investigación empírica en el nivel educativo de ES. Los criterios de exclusión fueron: artículos sin acceso al documento en extenso y/o que no narran experiencias del campo ético aplicado a la educación.

Descripción del proceso

En la búsqueda inicial de los artículos se identificaron 256 registros, de los cuales se seleccionaron 65. De estos, 14 estaban duplicados y otros 3 no eran *open access*, por lo que fueron eliminados. Los documentos seleccionados fueron procesados en *Mendeley* y *Rayyan* con la intención de analizarlos y categorizarlos. En la Figura 1 se representa el modelo metodológico a partir de la declaración PRISMA.

En la plataforma de *Rayyan* se exportaron 48 registros para una segunda etapa de cribado y clasificado de los artículos. Del total, uno fue duplicado, once se excluyeron y ocho pasaron a una segunda revisión; el resto se mantuvo. A su vez, se incluyeron ocho artículos de otras fuentes de información, mientras que otros ocho se descartaron por no ser del área educativa. Las etiquetas que se utilizaron para la clasificación de los artículos en esta plataforma de *Rayyan* fueron: uso ético, modelos didácticos, desafíos y riesgos éticos, integridad académica y marcos regulatorios.

Figura 1
Diagrama de flujo del modelo PRISMA



Al final se contó con 28 artículos, que se procesaron y categorizaron en el programa Excel para finalmente analizar su contenido a partir de las categorías: desafíos y riesgos éticos, marcos normativos-regulatorios, formación ética y modelos didácticos; estas mismas categorías se contemplaron en el software de análisis cualitativo MAXQDA.

RESULTADOS

En primer lugar, se brindan algunas características que ayudan a circunscribir el conjunto de artículos analizados. A continuación, estos se interpretan en clave de las categorías de interés, para dar cuenta de los aspectos éticos en relación con el uso de la IA en la ES.

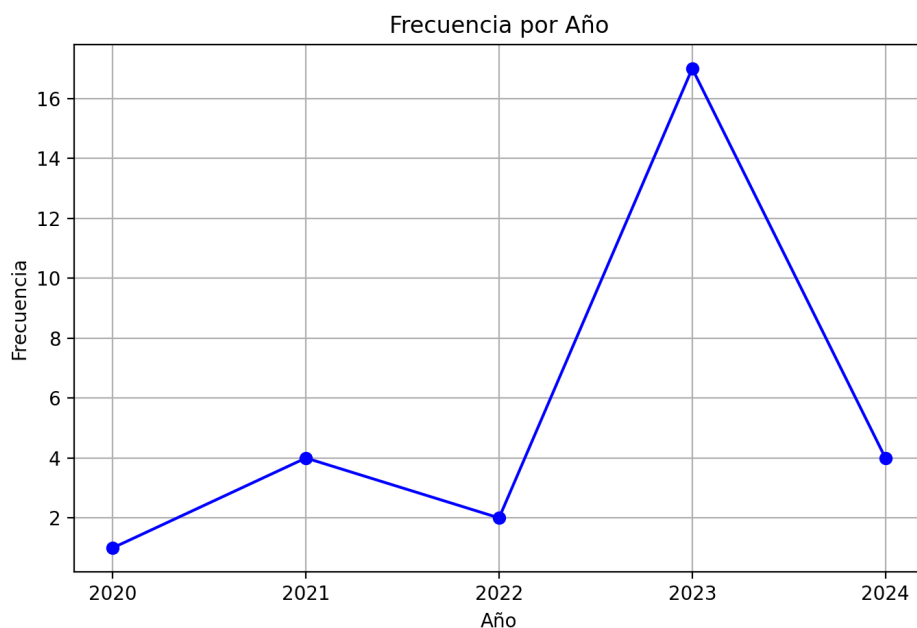
Caracterización de los artículos analizados

La síntesis de los principales resultados se presenta en la Tabla 1 (véase *Anexo*), organizada en orden alfabético. En ella, se recoge una síntesis de la información contextual de cada estudio, incluyendo referencia, metodología, lugar de estudio, palabras clave y etiquetas.

La Figura 2 ilustra la evolución de la producción científica sobre esta temática entre 2020 y mayo 2024. Destaca especialmente el año 2023, en el que se registró un notable aumento en la publicación de investigaciones, con un total de 17 trabajos, coincidiendo con el impacto generado tras el lanzamiento de ChatGPT el 30 de noviembre de 2022. En 2021 y 2024 se registraron cuatro publicaciones cada año, mientras que en 2022 se documentaron dos trabajos y en 2020, un único estudio.

Figura 2

Distribución temporal de los artículos



Distribución geográfica por lugar del estudio

La producción científica se concentra principalmente en España e Irlanda, que lideran como los países con mayor frecuencia de estudios sobre el fenómeno, con un total de cuatro publicaciones cada uno. Les siguen los Países Bajos y Suecia, con dos publicaciones. Los demás países registrados cuentan con una sola publicación, entre ellos: Bangladesh, Canadá, Chile, Croacia, China, Costa Rica, Corea del Sur, Ecuador, Italia, Nigeria, Perú, Reino Unido y Suiza. Además, se identificaron 12 estudios cuyo contexto geográfico no se especifica.

Diseño metodológico

El análisis del enfoque metodológico ofrece una visión clara de cómo se está abordando actualmente la investigación en esta temática. Según los datos presentados en el anexo 1, predominan los estudios de análisis bibliográfico en comparación con aquellos de enfoque cuantitativo, cualitativo o mixto. En concreto, 17 artículos (58,6 %) corresponden a revisiones de la literatura, ya sean sistemáticas o no. El 42 % restante se distribuye entre estudios mixtos (descripción de una experiencia basada en un modelo), representando cada uno un 3,4 %.

Abordaje de las categorías

A partir de un adentramiento gradual en el material, desde la técnica de análisis de contenido (Ruiz Bueno, 2021), ha sido posible identificar cuatro cuestiones centrales donde se focaliza la atención en términos éticos y normativos: desafíos y riesgos éticos, marcos normativos-regulatorios, formación ética y modelos didácticos.

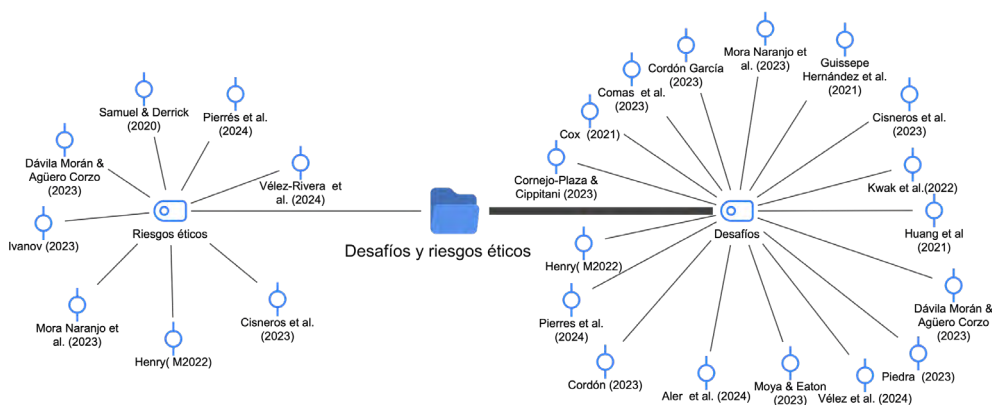
PI1. *¿Cuáles son los desafíos y riesgos éticos que se han detectado en las universidades?* (las referencias sobre este punto se muestran en la Figura 3).

El uso de la IA en los procesos de enseñanza, aprendizaje y evaluación presenta desafíos que vienen siendo reconocidos en estudios recientes a escala mundial, tanto desde la perspectiva de docentes y estudiantes. En particular, es considerable su impacto en la ES, generando oportunidades y riesgos (Chan y Hu, 2023). Por ejemplo, se presentan casos acerca de cómo estudiantes universitarios se van adaptando al uso de la IA en su aprendizaje (Cisneros et al., 2023). Puntualmente advierten que, aunque la IA facilita el acceso a información y recursos, al mismo tiempo puede generar sobrecarga de información con escasez de habilidades críticas para evaluar contenidos. Incluso, ante tal prevalencia de herramientas de IA en constante crecimiento, en ocasiones los estudiantes incurrir en plagio académico (Comas Forgas et al., 2023). La facilidad para generar contenido conlleva a esa tendencia y

abre cuestionamientos sobre la integridad académica, así como la originalidad en las producciones estudiantiles.

Figura 3

Referencias asociadas a los desafíos y riesgos éticos



Sin dejar de reconocer las oportunidades que ofrece la IA, como la personalización de aprendizajes (Córdón García, 2023), se identifican potenciales riesgos para las personas y sociedades. En efecto, se llega a hablar de la “deshumanización” de la educación, de la mano de una dependencia de las tecnologías. De ahí la importancia de incorporar la ética en una formación que contemple el uso de la IA de modo vital (Guiseppe Hernández et al., 2021; Huang et al., 2021). Encontramos también reportes que muestran que en ocasiones la ansiedad y la falta de autoeficacia en el uso de la IA por parte de los estudiantes (Kwak et al., 2022) pueden limitar su adopción y aplicación efectiva en entornos específicos de actuación. En esos términos se concibe una conciencia ética sobre la IA, que influye en las intenciones de comportamiento de quienes la usan.

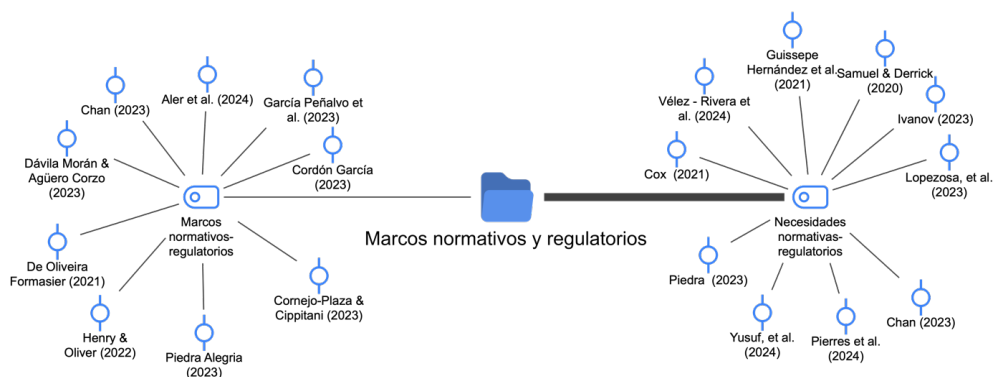
En síntesis, es posible subrayar que, aunque la IA ofrece oportunidades significativas para mejorar la enseñanza, el aprendizaje y la evaluación, resulta fundamental abordar los desafíos éticos, la integridad académica y la preparación adecuada de los estudiantes para mitigar los riesgos asociados con su uso. Esto también involucra a los docentes e implica un enfoque proactivo en la formación y la regulación de la IA en contextos educativos. De hecho, informes de investigaciones sobre el tema (Ivanov, 2023; Vélez-Rivera et al., 2024) advierten la relevancia creciente que al binomio IA-ética le vienen dando las universidades, tanto a nivel institucional general como en carreras/áreas puntuales.

PI2. ¿Cuáles son las medidas normativas y éticas que han aplicado o sugerido las universidades en torno a la IA?

Un nutrido número de documentos revisa la *necesidad de marcos normativos y/o regulaciones* sobre el uso de la IA en la ES (Figura 4). Ha sido posible reconocer numerosas iniciativas en torno a medidas normativas y éticas que han sido propuestas o aplicadas por universidades alrededor del mundo en los últimos años. Desde el Parlamento Europeo se acuerda en que la implementación de un Código de Ética Académica proporciona un marco normativo que establece los principios y valores con la premisa de guiar la conducta en relación con el uso de IA, tanto por parte de estudiantes como de profesores (Cornejo-Plaza y Cippitani, 2023). En este sentido, las universidades han definido actos como el plagio y la falsificación de datos, entre otras prácticas deshonestas, como violaciones a este código, con consecuencias establecidas para tales infracciones.

Figura 4

Referencias asociadas a las medidas regulatoria y normativas



Se han propuesto políticas que guían el uso ético y responsable de la IA en entornos académicos (Aler et al., 2024). Estas políticas abordan la necesidad de asegurar que las tecnologías de IA se utilicen de manera que respeten la privacidad, equidad y transparencia. Por ello se espera que las universidades establezcan normas claras sobre el uso de este tipo de herramientas, al mismo tiempo que puedan dotar a docentes y estudiantes de pautas para su integración en el proceso educativo.

De ahí la relevancia de avanzar en la formación sobre IA, con la implementación de talleres y cursos que aborden la importancia de la integridad académica, el plagio, la citación y las buenas prácticas de investigación (Cordón García, 2023). Conscientes de ello, varias instituciones españolas han desarrollado módulos obligatorios en el currículo, con estudiantes de primer año, de cara a capacitarlos en materia de ética digital.

Las propuestas regulatorias en el nivel universitario no son restrictivas al ámbito de docencia, puesto que la Unión Europea también ha establecido los Comités de

Ética en Investigación (Piedra Alegría, 2023). La propuesta es revisar y aprobar los proyectos de investigación antes de ejecutarse para garantizar la ética en el tratamiento de datos y la protección de los participantes. Yusuf et al. (2024) sugieren la implementación de evaluaciones de impacto ético en proyectos que involucren IA, revisando las implicaciones sociales y éticas, lo que otorga relevancia a marcos normativos en este punto.

Además de las propuestas anteriores, se pretende que existan regulaciones para generar ambientes inclusivos basados en IA, que fomenten la participación de los grupos marginados (Cox, 2021) mediante acciones como normativas para la admisión y becas para garantizar la igualdad y accesibilidad. Otro de los aspectos analizados es la responsabilidad social universitaria y la IA, por lo que los principios de FATE (*Fairness, Accountability, Transparency and Ethics*) son una muestra de las iniciativas vinculadas con las actividades académicas y de investigación para el bienestar de la comunidad (Henry y Oliver, 2022). Todo ello mediante proyectos de extensión y colaboración estudiantil y comunitarios, que alientan a las universidades a ser transparentes en el uso de los algoritmos de IA y la toma de decisiones resultante de estos.

La protección y privacidad de datos es otro aspecto por garantizar en las instituciones cuando incorporan IA, por lo que las regulaciones van encaminadas para que las universidades cumplan con dichas medidas (García Peñalvo et al., 2024) a través de una adecuada gestión y resguardo de la información. A este respecto, cobra relevancia la evaluación de las políticas en torno al uso de la IA (Moya y Eaton, 2023), por lo que se invita a las universidades a revisar y actualizar periódicamente sus normativas para respaldar su eficacia ante los nuevos desafíos.

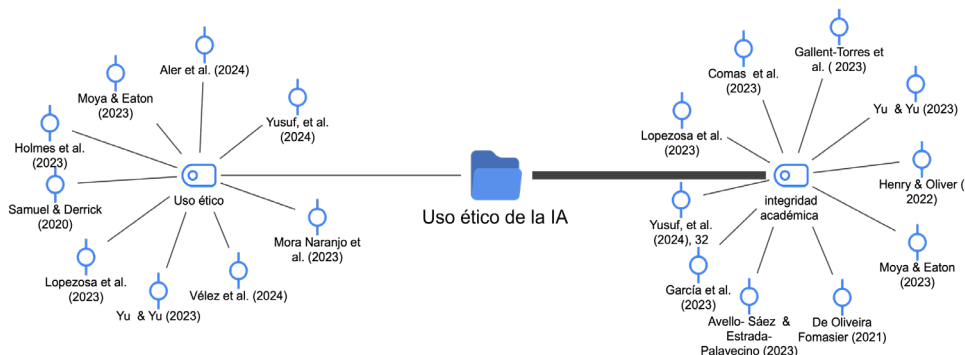
Junto a lo expuesto, se alienta a las universidades a investigar las repercusiones sociales, económicas, ambientales y culturales de las tecnologías de IA para comprender su impacto en la sociedad (Chan, 2023). Esta sinergia entre investigación-docencia-vinculación-comunidad puede ayudar a las instituciones a anticiparse a los problemas y ajustar sus políticas de acuerdo con los hallazgos.

PI3. *¿Cómo se interpreta el uso ético de la IA en los procesos educativos?* (Figura 5)

Los hallazgos obtenidos han permitido conocer a fondo las implicaciones “éticas” del uso de la IA en ES; ello ha desencadenado la necesidad de contar con un marco regulatorio. Concretamente, en esta pregunta se examina el panorama actual de las políticas universitarias con respecto al uso de IA. De particular interés son las universidades asiáticas, que presentan un contexto peculiar para observar la respuesta institucional a IA (Dai et al., 2024). Se espera que los hallazgos ofrezcan información útil sobre las direcciones futuras del desarrollo de políticas al respecto en tales universidades y en todo el mundo.

Figura 5

Referencias asociadas al uso ético de la IA en los procesos educativos



La integridad académica conforma un pilar esencial en la educación, en general, y en la formación de profesionales, en particular, sobre todo en un mundo donde la IA juega un papel cada vez más marcado. Múltiples herramientas, entre ellas el conocido ChatGPT, están presentes en la formación de competencias de terapeutas ocupacionales (Avello-Sáez y Estrada-Palavecino, 2023). En este escenario, amerita prestar atención a la integridad académica para evitar dependencias indebidas y promover aprendizajes auténticos que se requerirán en el desempeño laboral; entre ellos, la capacidad de interpretar las decisiones de la IA en consonancia con criterios éticos (Yu y Yu, 2023). Esto ha de trabajarse desde el inicio de la formación universitaria, con estrategias que refuercen la ética académica; como se advirtió en las carreras de Trabajo Social y Educación Social (Comas Forgas et al., 2023), el plagio académico puede prevalecer entre estudiantes.

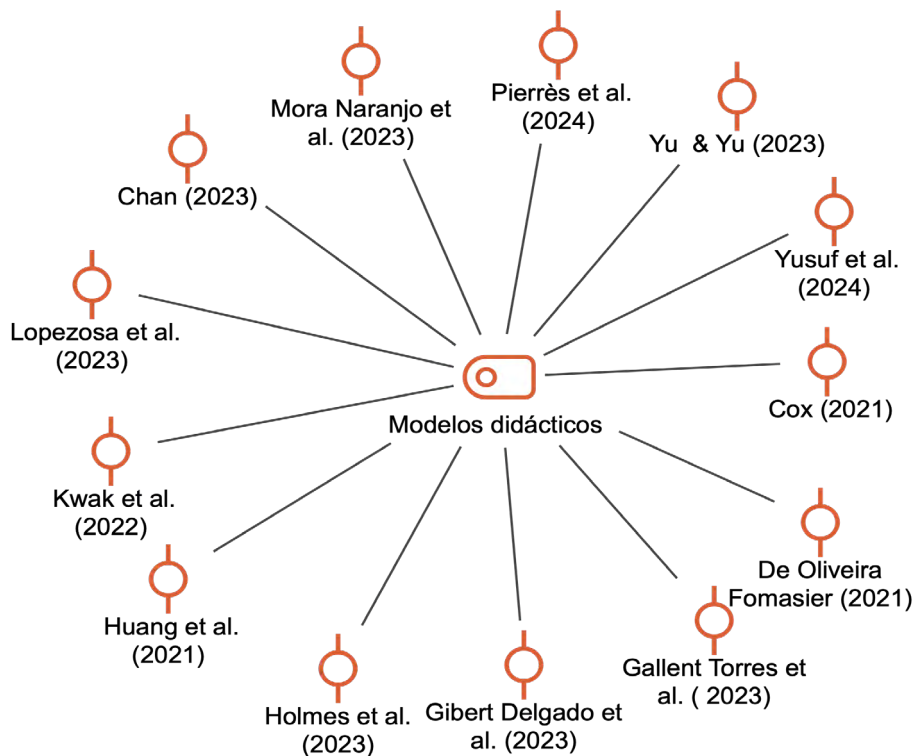
Sin duda, existe una responsabilidad ineludible en la implementación de la IA en los distintos niveles del sistema educativo, en particular el superior (Mora Naranjo et al., 2023), donde una base ética de la mano de la integridad académica resulta esencial para su genuina aceptación y sostenibilidad.

PI4. *¿Cuáles son los enfoques o modelos didácticos que se han implementado en las universidades para el uso confiable de la IA?* (Figura 6)

La creciente integración de la IA en la ES ha provocado debates en torno a los modelos didácticos y el uso ético de estas tecnologías. Los modelos educativos que se alientan refieren una integración efectiva de la IA en prácticas educativas (Chan, 2023) a través de estrategias específicas para su uso en los diversos contextos educativos. Ello en el marco de estándares de integridad académica a delimitar con políticas integrales y lineamientos lo más claros posible. Se reconocen beneficios para los aprendizajes, al mismo tiempo que preocupaciones, que giran en torno a la conjunción autonomía/dependencia, así como originalidad/plagio.

Figura 6

Referencias asociadas a los modelos didácticos implementados en las universidades



Una de las herramientas que ha cobrado mayor popularidad en diversidad de ámbitos, entre ellos académicos de nivel superior y de desarrollo profesional de los graduados, es el ChatGPT. Dada la importancia de educar a los estudiantes y profesionales en el empleo adecuado de estas herramientas, se ofrecen guías para favorecer el uso de esta IA conversacional de manera ética para complementar, y no reemplazar, las competencias humanas. Esto resulta acorde con iniciativas para fomentar una cultura de integridad académica y responsabilidad ética (Gallent Torres et al., 2023), particularmente en el uso de tecnologías avanzadas.

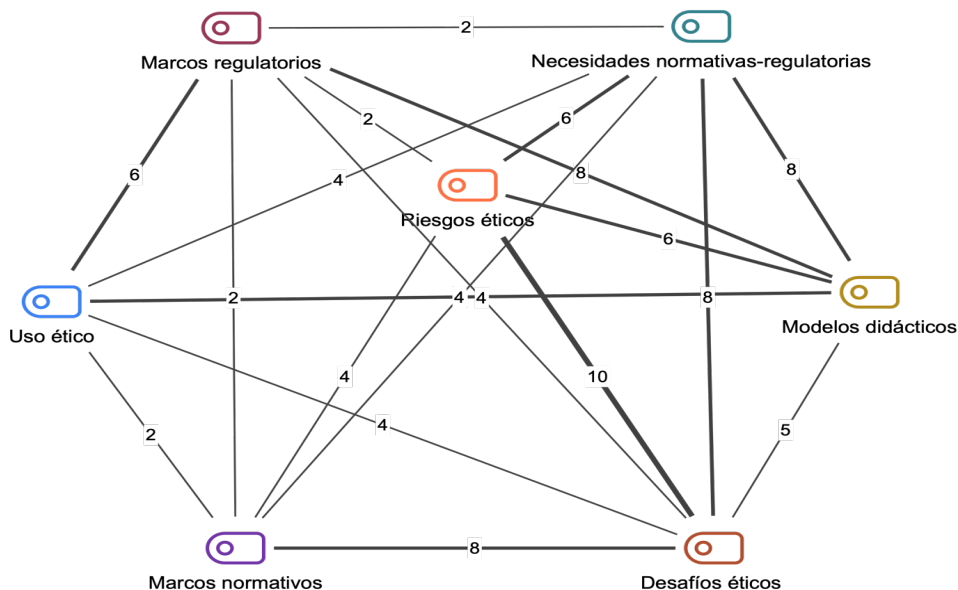
También se ha acuñado el concepto de Educación 4.0 (Gibert Delgado et al., 2023) en referencia a un enfoque que incorpora IA de manera innovadora con el propósito de promover el desarrollo de competencias digitales y pensamiento crítico en los futuros profesionales, afín a los desafíos de un mundo cada vez más digitalizado. Incluso la ética ha de constituir una parte nodal de los programas educativos que incorporen IA, de modo que los profesionales estén preparados para manejar los

dilemas éticos que puedan surgir en sus ámbitos laborales (Huang et al., 2021). Se espera, en ese contexto, que los sistemas mantengan altos estándares éticos (Alert et al., 2024), mediante evaluaciones transparentes y justas, conscientes de los sesgos potenciales que pueden surgir al incluir IA.

De cara a resumir las relaciones entre las categorías analizadas en este estudio, se muestra un mapa de co-ocurrencia de códigos (Figura 7), en los que se representa el uso ético de la IA y su relación intrínseca con los requerimientos normativos sólidos para la utilización responsable, confiable y transparente de la tecnología. Se evidencia la necesidad de una normativa adecuada para mitigar los riesgos, mediante modelos didácticos que promuevan un equilibrio entre la innovación tecnológica y los principios éticos con los cuales regir su uso en la ES.

Figura 7

Co-ocurrencia de códigos en los documentos



Análisis bibliométrico

Tras realizar el análisis descriptivo del impacto de los artículos seleccionados y con el objetivo de identificar las posibles tendencias en la producción científica de los últimos casi cinco años (2020 - mayo 2024), se llevó a cabo un estudio de las relaciones entre sus palabras clave. Para ello, se utilizó la técnica de análisis de redes

(Knoke y Yang, 2008), empleando el programa *VOSviewer*. De los 28 estudios que conforman la muestra, se extrajeron 90 palabras clave, cada una con un índice de co-ocurrencia de al menos dos veces. En el mapa de relaciones etiquetado (Figura 8), los nueve clústeres o nodos de palabras clave se relacionan, en función del grado de similitud, mostrando cómo se interrelacionan entre sí.

El clúster rojo se centra en los enfoques de investigación que abordan la relación entre ética y tecnología, especialmente en los campos de la investigación con seres humanos, como la educación y la medicina. Aumenta el interés por usar estas tecnologías en los entornos educativos, mostrando el papel de las plataformas digitales tanto en la comunicación de hallazgos como en la integración de la comunidad educativa.

El clúster verde hace foco en el uso de la tecnología en ES para ayudar a promover inclusión en este nivel, especialmente de personas con discapacidad. Enfatiza el desarrollo de tecnologías de asistencia para mejorar una educación inclusiva, enfrentando desafíos éticos en su diseño y acceso para evitar desigualdades.

El clúster azul puntualiza los retos que presenta la transformación digital en los sectores de la investigación y la educación, subrayando el papel de la IA como impulsor de cambio. Destaca la necesidad de asegurar la autonomía y la integridad académica, adoptando un enfoque que prioriza al ser humano.

El clúster amarillo aborda la transformación que la IA y el análisis de grandes datos están provocando en el sector educativo, puntualmente en la optimización de los procesos de enseñanza y aprendizaje. Estas tecnologías permiten la personalización del contenido, ajustándose a las necesidades específicas de cada alumno, con un énfasis particular en el uso de robots sociales para potenciar la experiencia educativa.

El clúster morado comprende los retos y las oportunidades que los modelos de IA ofrecen en el sector educativo, con atención a la integridad académica y a la evaluación del aprendizaje. La implementación de esta tecnología ha incrementado las inquietudes relacionadas con el plagio, lo que requiere la creación de métodos de evaluación innovadores que desalienten esta conducta y fomenten un uso responsable.

El clúster turquesa abarca el ámbito legal, enfatiza el papel de la IA en la creación de herramientas de digitalización y automatización que, en el ámbito educativo, ofrecen múltiples oportunidades para mejorar la accesibilidad y adaptar la enseñanza a las necesidades individuales. No obstante, también resalta la importancia de asumir una responsabilidad en la implementación de estas tecnologías, para reducir los sesgos que podrían condicionar desfavorablemente los resultados de aprendizaje.

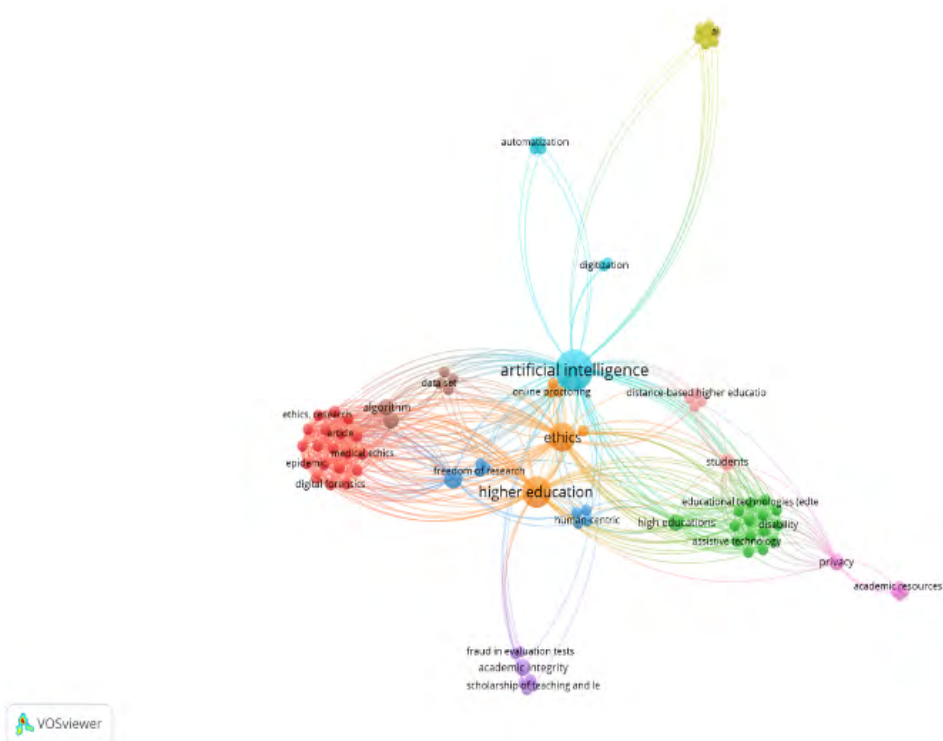
El clúster naranja analiza la relación entre la utilización de algoritmos, la supervisión en línea y los sistemas informáticos en los laboratorios de la ES, focaliza la dificultad de desarrollar y emplear estas tecnologías de manera ética.

El clúster marrón se refiere a la toma de decisiones fundamentadas en datos y algoritmos, considerando además posibles efectos adversos, tales como las implicaciones éticas y los sesgos que pueden surgir.

Finalmente, el clúster rosa examina el impacto de la IA en el acceso y uso de los recursos académicos, al mismo tiempo que aborda los desafíos de adaptación y privacidad que surgen con su implementación.

Figura 8

Mapa de relaciones etiquetado



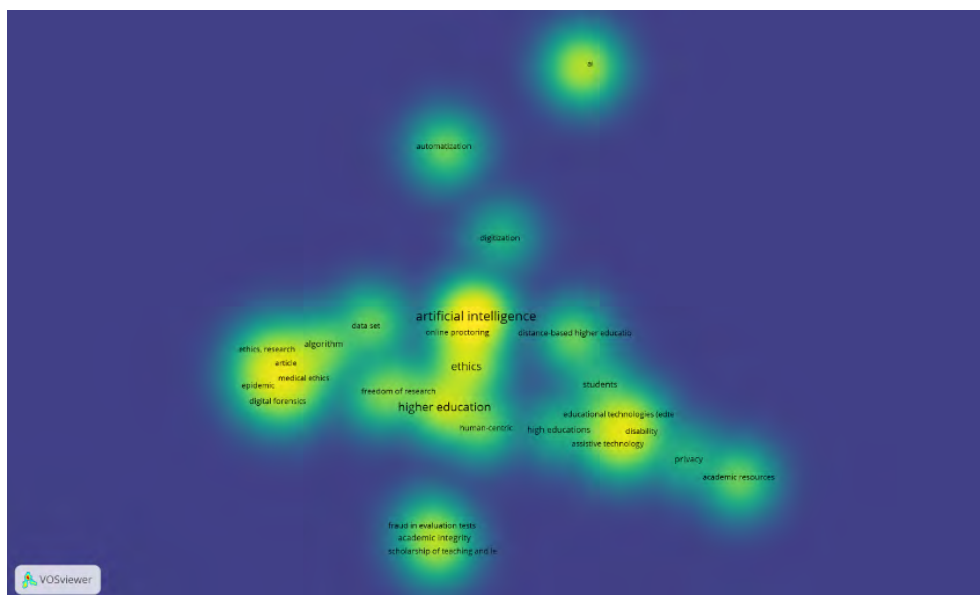
Atendiendo al tamaño de las etiquetas de las palabras clave y al color de las diferentes zonas de aproximación, se observa el nivel de relevancia de los estudios a través del mapa de visualización de densidad en la Figura 9.

De esta manera, en la zona central del mapa (color amarillo) se sitúan, por su importancia y co-ocurrencia, las palabras clave pertenecientes a la línea de investigación relacionada con la IA, la ES y la ética, entre las que se encuentran: AI, high education, ethics, online proctoring, article, ethics research, disability, entre otras.

Por otro lado, la zona periférica del mapa (color próximo al verde) muestra la menor densidad de co-ocurrencia de los términos clave. Se localizan expresiones más

específicas en torno a los que se han direccionado las investigaciones analizadas. Se distinguen palabras como: academic resources, automatization, digitization, privacy, students, distance-based higher education, data set, algorithm.

Figura 9
Mapa de densidades



DISCUSIÓN

El propósito de este estudio ha sido sintetizar aquellas investigaciones que se hacen eco del uso de la IA en el ámbito de la ES, ante los desafíos éticos emergentes y/o riesgos potenciales. Los estudios generados al respecto arrojan luz sobre cómo enfrentar el desafío de equilibrar la innovación tecnológica con la ética, la inclusión y la sostenibilidad. Cuatro parecen ser las claves: (1) la superación de *desafíos y/o riesgos éticos* a través del (2) diseño y desarrollo de *normativas* que aseguren un (3) *uso cívico y responsable de la IA* en los contextos académicos, y promuevan, a su vez, la (4) implementación de prácticas educativas y/o *modelos didácticos éticos*, que inciten al pensamiento crítico tanto en su uso como en su enseñanza.

No obstante, los resultados obtenidos en este trabajo advierten una necesidad, con algunos avances, en torno a las medidas normativas y éticas existentes sobre la IA, propuestas por las propias universidades, entre otros órganos competentes.

Si bien las universidades han estado desarrollando normativas para guiar su integración y abordar los desafíos más relevantes (Michel-Villarreal et al., 2023) –lo cual ha supuesto nuevas oportunidades y posibilidades de repensar la formación de los futuros profesionales (Dai et al., 2024)–, aún queda camino por recorrer. En este punto, las instituciones de ES quedan convocadas a desarrollar sus respectivas políticas de orientación, administración y regulación de la IA (UNESCO, 2023). La existencia de estos marcos normativos fomenta un entorno académico transparente y de apoyo, que educa y guía a los estudiantes para hacer un uso adecuado y confiable de la IA, y promueve su integración en los procesos educativos.

De cara al diseño de estos marcos normativos, Nguyen et al. (2023) recogen en su trabajo siete principios que deberían contemplarse en el desarrollo de cualquier normativa sobre IA en educación (IAED, como refieren los autores): 1) gobernanza y administración; 2) transparencia y responsabilidad; 3) sostenibilidad y proporcionalidad; 4) privacidad; 5) seguridad y protección; 6) inclusión; 7) IAED centrada en el ser humano.

Recientemente, en España, el Instituto Nacional de Tecnologías Educativas y de Formación del Profesorado del Ministerio de Educación, Formación Profesional y Deportes ha publicado una guía sobre el uso de la IA en el ámbito educativo. En concreto, se recoge un decálogo para atender a los principales dilemas éticos hallados (Instituto Nacional de Tecnologías Educativas y de Formación del Profesorado [INTEF], 2024, p. 20): “Integración respetuosa y sostenible; Transparencia y conciencia; Equidad y no discriminación; Robustez y seguridad; Privacidad y protección de datos; Supervisión humana; Compatibilidad humana; Promoción del bienestar social; Aprendizaje colaborativo; Reflexión y anticipación”.

CONCLUSIONES

A raíz de los datos obtenidos, se puede identificar una serie de conclusiones, que dan respuesta a las preguntas de investigación que han guiado el trabajo.

Atendiendo a la **PI1**, los diferentes estudios reflexionan ante la transición de adopción de la IA en las instituciones de ES y cómo estas enfrentan desafíos como la integridad académica; se advierte que el uso desregulado de la herramienta conlleva consecuencias negativas como sobrecarga o dependencia, aumento de la ansiedad, falta de autoeficacia y promoción de infracciones (plagio, falsificación). Ante este escenario, se sugiere el diseño e implementación de Códigos de Ética Académica, que definen normas y penalizan acciones como el plagio y la alteración de datos (Cornejo-Plaza y Cippitani, 2023), así como el establecimiento de políticas de empleo responsable de la IA, garantizando su uso en entornos que respeten la privacidad, equidad y transparencia (Aler et al., 2024).

Respecto a la **PI2**, se detectó una tendencia en cuanto a las medidas normativas y éticas que han sugerido las universidades, que se alinean a políticas internacionales como las de la Unión Europea y la UNESCO (Cornejo-Plaza y Cippitani, 2023), para

adaptar sus posibles normativas. Sin embargo, la aplicación de los reglamentos implica promover una responsabilidad compartida entre docentes y estudiantes para impulsar el pensamiento crítico en la toma de decisiones. Al respecto, diferentes instituciones han implementado cursos y talleres para reforzar la integridad académica, el uso ético de la IA, competencias digitales, buenas prácticas de investigación (Cordón García, 2023) y la creación de Comités de Ética (Piedra Alegría, 2023). Al tiempo que se han implementado evaluaciones de impacto ético en proyectos de IA para prever sus implicaciones sociales y culturales (Moya y Eaton, 2023; Yusuf et al., 2024) y para asegurar la diversidad e inclusión (Cox, 2021). Ejemplos muy específicos de recomendaciones éticas son las que ha sugerido el TEC de Monterrey en México (<https://bit.ly/4317udp>), donde se han establecido nueve principios para el uso de la IA por docentes y estudiantes. Sucintamente, estas normativas buscan sostener estándares académicos y fomentar una cultura de ética y responsabilidad en las universidades.

Considerando la **PI3**, los datos analizados indican un rol protagónico de las universidades para propiciar, gestionar, evaluar y actualizar marcos relevantes y eficaces, capaces de anticiparse a los problemas y proponer estándares. Ello engloba, asimismo, al ámbito de “buenas prácticas” en investigación y, en correspondencia con ello, la integración de la capacitación en competencias digitales, la creación de comités de ética y la realización de la evaluación del impacto ético en proyectos de IA. En efecto, el foco de atención está puesto en mitigar los riesgos de deshonestidad académica y promover un uso responsable de estas herramientas. Un ejemplo de ello se encuentra en el último boletín publicado por el Consejo Interuniversitario Nacional de Argentina (<https://www.cin.edu.ar/>), que asume a la IA como un recurso educativo válido en la ES.

Finalmente, en torno a la **PI4**, se concluye que la incorporación de la IA en la ES no solo se centra en la regulación, sino también en la promoción del uso adecuado y la reflexión crítica de su impacto en los diferentes sectores de la sociedad. Aparecen así modelos didácticos para integrar la IA en la práctica docente de modo que logren inspirar confianza.

La integración de la IA en la ES requiere no solo un enfoque regulatorio, sino también una reflexión crítica sobre su impacto en la formación académico-profesional, la creación de conocimiento y el desarrollo de competencias digitales. Las universidades están llamadas a prepararse para formar profesionales que utilicen estas tecnologías de manera responsable y colaborativa, integrándose en sus estrategias didácticas sin comprometer los principios éticos fundamentales. Este desafío implica un esfuerzo conjunto de la comunidad académica para asegurar que la adopción de la IA se haga de manera reflexiva, ética y alineada con los objetivos de calidad educativa.

A través de esta revisión se ha pretendido comprender el alcance y la naturaleza de las políticas existentes, así como identificar prácticas prometedoras, brechas y desafíos (Pham et al., 2014). No obstante, el estudio no está exento de limitaciones.

Cabe mencionar, fundamentalmente, la relativa a la muestra; la mayoría de la literatura revisada y analizada procede de fuentes europeas, que pueden reflejar marcos culturales, regulatorios y educativos específicos, exclusivos de esa zona. Y esta centralización geográfica –o sesgo geográfico– puede afectar significativamente a la generalización de los hallazgos a un contexto global más amplio. Atendiendo a esta limitación, y a sus posibles implicaciones, sería posible mejorar la aplicabilidad de las propuestas, señaladas más arriba, en contextos internacionales realizando más estudios que incluyeran diversas perspectivas geográficas y culturales.

En conclusión, si bien este estudio ofrece información rigurosa sobre el uso ético de la IA en el contexto académico, sus limitaciones con respecto al enfoque geográfico de la muestra exigen cautela a la hora de generalizar los hallazgos. Abordar estas limitaciones, por ejemplo, mediante una investigación inclusiva, puede ayudar a garantizar que las propuestas sean relevantes y efectivas en diferentes contextos internacionales. Sea como fuere, el reto principal será fomentar un entorno universitario que priorice las consideraciones éticas y, a la vez, incorpore los últimos avances tecnológicos.

REFERENCIAS

- Aler, A., Mora, M. y Nieves, J. C. (2024). How to teach responsible AI in Higher Education: challenges and opportunities. *Ethics and Information Technology*, 26(3). <https://doi.org/10.1007/s10676-023-09733-7>
- Avello-Sáez, D. y Estrada-Palavecino, L. (2023). ChatGPT e seu impacto na formação de competências para terapeutas ocupacionais: uma reflexão sobre a integridade acadêmica. *Cadernos Brasileiros de Terapia Ocupacional*, 31, e3534. <https://doi.org/10.1590/2526-8910.ctoen271035343>
- Baldrich, K. y Domínguez, J. C. (2024). El uso de ChatGPT en la escritura académica: un estudio de caso en educación. *Pixel-Bit. Revista de Medios y Educación*, 71, 141-157. <https://doi.org/10.12795/pixelbit.103527>
- Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. *International Journal of Educational Technology in Higher Education*, 20(1), 38. <https://doi.org/10.1186/s41239-023-00408-3>
- Chan, C. K. Y. y Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education*, 20(43). <https://doi.org/10.1186/s41239-023-00411-8>
- Cisneros, J. D. D., Limo, F. A. F., Tinoco, L. M. B., Aybar, H. N. C., Alarcon, V. G. C., Romero, M. Á. M. y Flores, R. A. R. (2023). Adjustment of Peruvian university students to artificial intelligence. *Arts Educa*, 36, 237-248. <https://artseduca.com/article-detail/?id=145>
- Comas Forgas, R., Cerdá-Navarro, A., Touza Garma, C. y Moreno Herrera, L. (2023). Prevalencia y factores asociados al plagio académico en estudiantes de nuevo ingreso de Trabajo Social y Educación Social: un análisis empírico. *RELIEVE. Revista Electrónica de Investigación y Evaluación Educativa*, 29(2), 1-20. <https://doi.org/10.30827/relieve.v29i2.29055>
- Cordón García, O. (2023). Inteligencia Artificial en Educación Superior:

- Oportunidades y Riesgos. *RiiTE. Revista interuniversitaria de investigación en Tecnología Educativa*, (15), 16-27. <https://doi.org/10.6018/riite.591581>
- Cornejo-Plaza, I. y Cippitani, R. (2023). Consideraciones éticas y jurídicas de la Inteligencia Artificial en Educación Superior: desafíos y perspectivas. *Revista de Educación y Derecho*, (28). <https://doi.org/10.1344/REYD2023.28.43935>
- Cox, A. M. (2021). Exploring the impact of Artificial Intelligence and robots on higher education through literature-based design fictions. *International Journal of Educational Technology in Higher Education*, 18(3). <https://doi.org/10.1186/s41239-020-00237-8>
- Dai, Y., Lai, S., Lim, C. P. y Liu, A. (2024). University policies on generative AI in Asia: promising practices, gaps, and future directions. *Journal of Asian Public Policy*, 1-22. <https://doi.org/10.1080/17516234.2024.2379070>
- Dávila Morán, R. C. y Agüero Corzo, E. C. (2023). Desafíos éticos de la inteligencia artificial: implicaciones para la sociedad y la economía. *Conrado*, 19(94), 137-144.
- Flores, J. M. y García, F. J. (2023). Reflexiones sobre la ética, potencialidades y retos de la Inteligencia Artificial en el marco de la Educación de Calidad (ODS4). *Comunicar*, 31(74), 37-47. <https://doi.org/10.3916/C74-2023-03>
- Fornasier, M. de O. (2021). Artificial intelligence and democratic rule of law. *Revista de Estudos Constitucionais, Hermenêutica e Teoria do Direito (RECHTD)*, 13(3), 351-369. <https://doi.org/10.4013/rechtd.2021.133.06>
- Gairín, J. y Alguacil, L. (Coords.). (2024). *La gestión de la inteligencia artificial en los contextos universitarios iberoamericanos*. EDO-Serveis Universitat Autònoma de Barcelona. <https://ddd.uab.cat/record/300143>
- Gallent Torres, C., Zapata González, A. y Ortego Hernando, J. L. (2023). El impacto de la inteligencia artificial generativa en educación superior: una mirada desde la ética y la integridad académica. *RELIEVE - Revista Electrónica de Investigación y Evaluación Educativa*, 29(2). <https://doi.org/10.30827/relieve.v29i2.29134>
- García-Fuentes, O., Raposo-Rivas, M., Martínez-Figueira, M. E. y Sarmiento-Campos, J. A. (2024). Análisis de vídeo-anotaciones sobre el uso de recursos tecnológicos durante el Prácticum. *Pixel-Bit. Revista de Medios y Educación*, 70, 143-161. <https://doi.org/10.12795/pixelbit.102048>
- García Peñalvo, F. J., Llorens-Largo, F. y Vidal, J. (2024). The new reality of education in the face of advances in generative artificial intelligence. *RIED. Revista Iberoamericana de Educación a Distancia*, 27(1), 9-39. <https://doi.org/10.5944/ried.27.1.37716>
- Gibert Delgado, R. P., Gorina Sánchez, A., Reyes-Palau, N. C., Tapia-Sosa, E. V. y Siza Moposita, S. F. (2023). Educación 4.0: Enfoque innovador apoyado en la inteligencia artificial para la educación superior. *Universidad y Sociedad*, 15(6), 60-74. <https://rus.ucf.edu.cu/index.php/rus/article/view/4122/4032>
- González Alcaide, G. (2024). *1 d.C. (después de ChatGPT). Inteligencia artificial generativa en la educación superior*. Publicaciones de la Universitat de València. <https://puv.uv.es/1-d-c-despues-de-chatgpt.html>
- Guisseppe Hernández, Y. K., Hernández Chávez, W. J. y Moucharrafiéh Moucharafiéh, S. (2021). Orientaciones éticas en la educación superior para el uso de la inteligencia artificial en ambientes disruptivos. *REDHECS. Revista Electrónica de Humanidades, Educación y Comunicación Social*, 29(19), 69-92. <https://ojs2.urbe.edu/index.php/redhecs/article/view/4423>
- Henry, J. V. y Oliver, M. (2022). Who Will Watch the Watchmen? The Ethico-

- political Arrangements of Algorithmic Proctoring for Academic Integrity. *Postdigital Science and Education*, 4(2), 330-353. <https://doi.org/10.1007/s42438-021-00273-1>
- Holmes, W., Iniesto, F., Anastopoulou, S. y Boticario, J. G. (2023). Stakeholder perspectives on the ethics of AI in distance-based higher education. *International Review of Research in Open and Distributed Learning*, 24(2), 96-117. <https://doi.org/10.19173/irrodl.v24i2.6089>
- Huang, C., Samek, T. y Shiri, A. (2021). AI and ethics: Ethical and educational perspectives for LIS. *Journal of Education for Library and Information Science*, 62(4), 351-365. <https://doi.org/10.3138/jelis-62-4-2020-0106>
- Humble, N. y Mozeliuss, P. (2022). The threat, hype, and promise of artificial intelligence in education. *Discover Artificial Intelligence*, 2(1), 22. <https://doi.org/10.1007/s44163-022-00039-z>
- Instituto Nacional de Tecnologías Educativas y de Formación del Profesorado. (2024). *Guía sobre el uso de la Inteligencia Artificial en el ámbito educativo*. Ministerio de Educación, Formación Profesional y Deportes. https://code.intef.es/wp-content/uploads/2024/07/Gu%C3%ADa-sobre-el-uso-de-la-IA-en-el-ámbito-educativo-INTEF_2024.pdf
- Ivanov, S. (2023). The dark side of artificial intelligence in higher education. *The Service Industries Journal*, 43(15-16), 1055-1082. <https://doi.org/10.1080/02642069.2023.2258799>
- Knoke, D. y Yang, S. (2008). *Social network analysis*. Sage. <https://doi.org/10.4135/9781412985864>
- Kwak, Y., Ahn, J. W. y Seo, Y. H. (2022). Influence of AI ethics awareness, attitude, anxiety, and self-efficacy on nursing students' behavioral intentions. *BMC nursing*, 21(1), 267. <https://doi.org/10.1186/s12912-022-01048-0>
- López-Regalado, O., Núñez-Rojas, N., López-Gil, O. R. y Sánchez-Rodríguez, J. (2024). Análisis del uso de la inteligencia artificial en la educación universitaria: una revisión sistemática. *Pixel-Bit. Revista de Medios y Educación*, 70, 97-122. <https://doi.org/10.12795/pixelbit.106336>
- Lopezosa, C., Codina, L., Pont-Sorribes, C. y Váñez, M. (2023). Use of generative artificial intelligence in the training of journalists: challenges, uses and training proposals. *Profesional de la información*, 32(4), e320408. <https://doi.org/10.3145/epi.2023.jul.08>
- Marín, V. I. y Tur, G. (2024). Ethical Issues in the Use of Technologies in Education Settings: A Scoping Review. *Education in the Knowledge Society*, 25, e31301. <https://doi.org/10.14201/eks.31301>
- McGrath, C., Pargman, T. C., Juth, N. y Palmgren, P. J. (2023). University teachers' perceptions of responsibility and artificial intelligence in higher education. An experimental philosophical study. *Computers and Education: Artificial Intelligence*, 4, 100139. <https://doi.org/10.1016/j.caeai.2023.100139>
- Mendoza, D. S. O. E. (2021). Las revisiones sistemáticas como proceso de selección artificial del conocimiento científico en el área de salud. *Revista de Ciencias Sociales*, 30(59). <https://doi.org/10.20983/noesis.2021.1.9>
- Michel-Villarreal, R., Vilalta-Perdomo, E., Salinas-Navarro, D. E., Thierry-Aguilera, R. y Gerardou, F. S. (2023). Challenges and Opportunities of Generative AI for Higher Education as Explained by ChatGPT. *Education Sciences*, 13(9), 856. <https://doi.org/10.3390/educsci13090856>
- Mora Naranjo, B. M., Aroca Izurieta, C. E., Tiban Leica, L. R., Sánchez Morrillo, C. F. y Jiménez Salazar, A. (2023). Ética y Responsabilidad en la implementación de la Inteligencia Artificial en la educación. *Ciencia Latina. Revista Científica Multidisciplinar*, 7(6), 2054-

2076. https://doi.org/10.37811/cl_rem.v7i6.8833
- Moya, B. A. y Eaton, S. E. (2023). Examinando Recomendaciones para el Uso de la Inteligencia Artificial Generativa con Integridad desde una Lente de Enseñanza y Aprendizaje. *RELIEVE. Revista Electrónica de Investigación y Evaluación Educativa*, 29(2). <https://doi.org/10.30827/relieve.v29i2.29295>
- Nguyen, A., Ngo, H. N., Hong, Y., Dang, B. y Nguyen, B. P. T. (2023). Ethical principles for artificial intelligence in education. *Education and Information Technologies*, 28(4), 4221-4241. <https://doi.org/10.1007/s10639-022-11316-w>
- 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., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372(71). <https://doi.org/10.1136/bmj.n71>
- Pham, M. T., Rajić, A., Greig, J. D., Sargeant, J. M., Papadopoulos, A. y McEwen, S. A. (2014). A scoping review of scoping reviews: Advancing the approach and enhancing the consistency. *Research Synthesis Methods*, 5(4), 371-385. <https://doi.org/10.1002/jrsm.1123>
- Piedra Alegría, J. (2023). Anotaciones iniciales para una reflexión ética sobre la regulación de la Inteligencia Artificial en la Unión Europea. *Revista de Derecho*, (28), e3264. <https://doi.org/10.22235/rd28.3264>
- Pierrès, O., Christen, M., Schmitt-Koopmann, F. M. y Darvishy, A. (2024). Could the use of AI in Higher Education hinder students with disabilities? A scoping review. *IEEE Access*, 12, 27810-27828. <https://doi.org/10.1109/ACCESS.2024.3365368>
- Rodríguez-Argueta, C. M. (2020). Tendencias de la oferta en educación superior en El Salvador - Relevancia de las carreras en Ciencia, Tecnología, Ingenierías y Matemáticas (por sus siglas en inglés STEM) ante la nueva economía digital. *Entorno*, (70), 22-32. <https://doi.org/10.5377/entorno.voi69.9559>
- Ruiz Bueno, A. (2021). *El contenido y su análisis: enfoque y proceso*. Universitat de Barcelona. <http://hdl.handle.net/2445/179232>
- Samuel, G. y Derrick, G. (2020). Defining ethical standards for the application of digital tools to population health research. *Bulletin of the World Health Organization*, 98(4), 239. <https://doi.org/10.2471/BLT.19.237370>
- Sánchez, M. M. (2023). La inteligencia artificial como recurso docente: usos y posibilidades para el profesorado. *Educare*, 60(1), 33-47. <https://doi.org/10.5565/rev/educar.1810>
- Seldon, A., Lakhani, P. y Luckin, R. (2020). *The Ethical Framework for AI in Education*. The Institute for Ethical AI in Education. <https://bit.ly/3Q9yobB>
- Svaricek, R. (2024). Sceptical teacher and silent students: a case study of teacher changes during a teacher professional development programme. *Journal of Education for Teaching*, 50(2), 310-328. <https://doi.org/10.1080/02607476.2023.2296459>
- Torres-Fonseca, A. y López-Hernández, D. (2014). Criterios para publicar artículos de revisión sistemática. *Revista de Especialidades Médico-quirúrgicas*, 19(3), 393-399.
- UNESCO. (2023). *UNESCO survey: Less than 10% of schools and universities have formal guidance on AI*. UNESCO. <https://www.unesco.org/en/articles/unesco-survey-less-10-schools-and-universities-have-formal-guidance-ai>
- Vélez-Rivera, R., Muñoz-Álvarez, D., Leal-Orellana, P. y Ruiz-Garrido, A.

(2024). Uso de Inteligencia Artificial en educación superior y sus implicancias éticas. Mapeo sistemático de literatura. *Hacheteteapé. Revista científica de Educación y Comunicación*, (28). <https://doi.org/10.25267/Hachetetepe.2024.i28.1105>

Villegas, V. y Delgado, M. (2024). Inteligencia artificial: revolución educativa innovadora en la Educación Superior. *Pixel-Bit. Revista de Medios y Educación*, 71, 159-177. <https://doi.org/10.12795/pixelbit.107760>

Yu, L. y Yu, Z. (2023). Qualitative and quantitative analyses of artificial intelligence ethics in education using VOSviewer and CitNetExplorer. *Frontiers in Psychology*, 14, 1061778. <https://doi.org/10.3389/fpsyg.2023.1061778>

Yusuf, A., Pervin, N. y Román-González, M. (2024). Generative AI and the future of higher education: a threat to academic integrity or reformation? Evidence from multicultural perspectives. *International Journal of Educational Technology in Higher Education*, 21(21). <https://doi.org/10.1186/s41239-024-00453-6>

Fecha de recepción del artículo: 1 de diciembre de 2024
Fecha de aceptación del artículo: 7 de marzo de 2025
Fecha de aprobación para maquetación: 31 de marzo de 2025
Fecha de publicación en OnlineFirst: 22 de abril de 2025
Fecha de publicación: 1 de julio de 2025

ANEXO

Tabla 1
Relación de referencias analizadas

Referencia	Metodología	Lugar del estudio	Palabras clave	Etiquetas
Aler et al. (2024)	Revisión de literatura	Croacia. Irlanda. Países Bajos. España. Suecia. Estados Unidos	Trustworthy. AI Higher Education. AI ethics. Educational strategies	Uso ético, marcos regulatorios-normativos, riesgos éticos
Avello-Sáez y Estrada-Palavecino (2023)	Revisión de literatura	No se especifica	IA. Formación profesional. Terapia ocupacional. Ética basada en principios. Competencia profesional. Estudiantes	Integridad académica

González-Fernández, M. O., Romero-López, M. A., Sgreccia, N. F., & Latorre Medina, M. J. (2025). Normative framework for ethical and trustworthy AI in higher education: state of the art. [Marcos normativos para una IA ética y confiable en la educación superior: estado de la cuestión]. *RIED-Revista Iberoamericana de Educación a Distancia*, 28(2), pp. 181-208. <https://doi.org/10.5944/ried.28.2.43511>

Referencia	Metodología	Lugar del estudio	Palabras clave	Etiquetas
Chan (2023)	Mixto	China	AI policy framework. AI. ChatGPT. Ethics. Assessment	Uso ético, marcos regulatorios-normativos, riesgos éticos
Cisneros et al. (2023)	Mixto	Perú	AI. Adaptation. Privacy. Academic Resources. Perú	Desafíos éticos, riesgos
Comas et al. (2023)	Cuantitativo	España	Plagio académico. Integridad académica. Fraude en pruebas de evaluación. Educación superior	Integridad académica, desafíos éticos
Cordón García (2023)	Revisión de literatura	No se especifica	IA. Enseñanza superior. Nuevas tecnologías. Innovación pedagógica. Tecnología de la educación	Marcos regulatorios-normativos, desafíos
Cornejo-Plaza y Cippitani (2023)	Teórico reflexivo	No se especifica	Educación. IA. Integridad científica. Libertad de investigación. ChatGPT	Desafíos éticos, marcos normativos
Cox (2021)	Narrativas	No se especifica	AI. Robots. Social robots. Learning analytics. Big data. AIEd. Design fiction	Desafíos éticos, marcos regulatorios-normativos
Dávila Morán y Agüero Corzo (2023)	Revisión de literatura	No se especifica	IA. Ética. Responsabilidad. Sesgo. Automatización	Desafíos éticos, riesgos
Fornasier (2021)	Revisión de literatura	No se especifica	Legal education. Digitization. AI	Modelos didácticos, integridad académica
Gallent Torres et al. (2023)	Revisión de literatura	No se especifica	Ética. Integridad académica. IAG. Educación superior	Modelos didácticos
García Peñalvo et al. (2024)	Revisión de literatura	No se especifica	IA. IAG. ChatGPT. Educación	Integridad académica, marcos regulatorios-normativos

Referencia	Metodología	Lugar del estudio	Palabras clave	Etiquetas
Gibert Delgado et al. (2023)	Revisión de literatura	No se especifica	Educación 4.0. IA. Enseñanza personalizada. Investigación científica. Educación superior	Modelos didácticos
Guisseppe Hernández et al. (2021)	Revisión de literatura	No se especifica	IA. Ética. ES. Ambiente Disruptivo	Desafíos éticos, marcos regulatorios-normativos
Henry y Oliver (2022)	Revisión de literatura	No se especifica	Care. Ethics. Online proctoring. Higher education. Algorithms. AI	Marcos regulatorios-normativos, desafíos
Holmes et al. (2023)	Cualitativo	España	AI. Ethics. Distance-based higher education. Students. Teachers. Institutions. Theoretical framework	Uso ético, modelos didácticos
Huang et al. (2021)	Revisión de literatura	No se especifica	AI. Critical LIS. Ethics. Higher education	Desafíos éticos, modelos didácticos
Ivanov (2023)	Revisión de literatura	No se especifica	AI. Higher education. Negative impacts. Ethics. Decision-making approaches	Marcos regulatorios-normativos, riesgos
Kwak et al. (2022)	Cuantitativo	Corea	AI ethics awareness. Attitude toward AI. Anxiety. Self-efficacy. Behavioral intention	Desafíos éticos, modelos didácticos
Lopezosa et al. (2023)	Cualitativo	España	IA. Periodismo. Periodistas. Comunicación. Entrevistas. Formación. Innovación. Planes docentes. Universidad. Estudios universitarios. Competencias. Capacitación. Ética. ChatGPT	Usos éticos, integridad académica, modelos didácticos, marcos regulatorios-normativos
Mora Naranjo et al. (2023)	Mixto	Ecuador	IA. Ética. Responsabilidad. Educación superior	Uso ético, riesgos, modelos didácticos

Referencia	Metodología	Lugar del estudio	Palabras clave	Etiquetas
Moya y Eaton (2023)	Revisión de literatura	Canadá	IA. IAG. Modelos de lenguaje avanzados. Integridad académica. Scholarship of teaching and learning. Enfoque de sistemas	Integridad académica, desafíos éticos
Piedra Alegría (2023)	Revisión de literatura	Costa Rica	Ética. Regulación. Autorregulación. IA. Derechos humanos	Marcos regulatorios-normativos, desafíos éticos
Pierrès et al. (2024).	Revisión de literatura	Suiza	AI. Disabilities. Higher education. Educational technologies. Ethics. Risk assessment	Desafíos éticos, riesgos, modelos didácticos, marcos regulatorios-normativos
Samuel y Derrick (□2020)	Experiencia de un modelo	Reino Unido de Gran Bretaña. Irlanda del Norte	Investigación sanitaria. ES	Uso ético, marcos regulatorios-normativos, riesgos
Vélez-Rivera et al. (2024)	Revisión de literatura	Chile	Educación superior. Tecnologías emergentes. IA. Ética. Integridad académica	Uso ético, marcos regulatorios-normativos, riesgos y desafíos éticos
Yu y Yu (2023)	Revisión de literatura	Italia	AI. Ethics. Bibliometric analysis. VOSviewer. CitNetExplorer	Uso ético, modelos didácticos, integridad académica
Yusuf et al. (2024)	Mixto	Nigeria. España	GenAI. Higher education. Potential. Concerns. Ethical regulations. Cultural dimensions	Marcos regulatorios-normativos, integridad académica, uso ético, modelos didácticos

Desafíos de la Inteligencia Artificial generativa en educación superior: fomentando su uso crítico en el estudiantado

Challenges of generative Artificial Intelligence in higher education: promoting its critical use among students



- Teresa Romeu Fontanillas - *Universitat Oberta de Catalunya, UOC (España)*
- Marc Romero Carbonell - *Universitat Oberta de Catalunya, UOC (España)*
- Montse Guitert Catasús - *Universitat Oberta de Catalunya, UOC (España)*
- Pablo Baztán Quemada - *Universitat Oberta de Catalunya, UOC (España)*

RESUMEN

La irrupción generalizada de la Inteligencia Artificial generativa (IAG) plantea nuevos desafíos para su integración en la Educación Superior. El conocimiento de la IAG, de sus posibilidades y de sus riesgos, debe ir acompañado de la reflexión sobre los aspectos normativos y éticos que esta nueva tecnología suscita. Este artículo presenta los resultados de incluir, en el currículum de la asignatura de una universidad en línea, un recurso de aprendizaje sobre la IAG y una reflexión compartida sobre las implicaciones éticas de su uso. Se analizan los datos cuantitativos y cualitativos obtenidos de más de 900 estudiantes universitarios mediante un cuestionario en línea con preguntas abiertas y cerradas, distinguiendo dos grupos: uno que no tuvo acceso al recurso de aprendizaje y posterior debate y otro que sí. Se analiza comparativamente cómo influye la formación seguida en el conocimiento autopercibido y la influencia en este conocimiento de otras variables, como el nivel formativo de los participantes o la tipología de estudios cursados. Del análisis cualitativo han emergido siete categorías que aglutinan los principios éticos que el estudiantado debería tener en cuenta a la hora de utilizar la IAG. Los resultados demuestran cómo una formación específica sobre IAG mejora el conocimiento de esta, ayudando a discernir sobre sus potencialidades y aspectos críticos de forma informada y consciente, al tiempo que ayudan a la mejora de los elementos de la asignatura relativos a la IAG.

Palabras clave: inteligencia artificial; estudios universitarios; ética; pensamiento crítico.

ABSTRACT

The widespread emergence of generative artificial intelligence (GenAI) presents significant challenges for its integration into Higher Education. Understanding AIG, its possibilities, and its risks must go hand in hand with a critical reflection on the regulatory and ethical issues it raises. It is essential for students to develop informed perspectives that allow them to use GenAI responsibly. This article presents the results of incorporating a specific learning resource on GenAI, along with a shared ethical discussion, into the curriculum of an online university course. The study analyzes both quantitative and qualitative data collected from over 900 university students through an online questionnaire featuring open and closed questions. Two groups were compared: one that did not have access to the learning resource and debate, and another that did. The research examines how this educational experience influences students' self-perceived knowledge of GenAI and explores the impact of other variables, such as the participants' level of education and the type of studies they pursued. From the qualitative analysis, seven categories emerged, grouping the ethical principles that students should consider when engaging with GenAI. The findings demonstrate that specific training on GenAI improves students' understanding, helping them critically assess its potential and challenges. Additionally, the results contribute to enhancing course components related to GenAI, fostering a more responsible and reflective approach to its use.

Keywords: artificial intelligence; university studies; ethics; critical sense.

INTRODUCCIÓN

La irrupción y rápida adopción de la Inteligencia Artificial generativa (IAG) ha revolucionado la forma de generar contenido textual, visual y audiovisual. La IAG, definida como “una tecnología que (i) utiliza modelos de aprendizaje profundo para (ii) generar contenido similar al humano (por ejemplo, imágenes, texto) en respuesta a (iii) indicaciones complejas y variadas (como idiomas, instrucciones, preguntas)” (Lim et al., 2023, p. 2), conlleva grandes promesas y desafíos. Las herramientas de IAG mejoran continuamente sus prestaciones respecto a versiones anteriores, permitiendo completar una amplia variedad de tareas sin necesidad de conocimientos especializados (Jarrahi et al., 2023), al tiempo que se retroalimentan con la información existente en la red, incluso con la nueva información generada.

La IAG permite que cualquier persona con acceso a internet pueda interactuar con ella, ofreciendo formas nuevas de relacionarse, consumir y compartir información, con implicaciones en prácticamente todos los ámbitos, por lo que no es sencillo abarcar el alcance de esta nueva realidad.

Otra importante amenaza que se presenta es el compromiso para la seguridad y privacidad que supone el flujo de millones de datos y conocimiento de los usuarios (Ruschmeier, 2024; Shi, 2023); así como el posible sesgo o falta de rigor en la información, ya que, tal y como señala Grassini (2023), las respuestas de modelos como Chat GPT pueden no ser siempre exactas, sobre todo, en el caso de información especializada, o como la generación de contenido que, aunque simula ser auténtico, carece de una base real, aspecto que aumenta la incertidumbre sobre la veracidad de la información (Raman et al., 2024; Wach et al., 2023).

Además, la IAG amplifica las diferencias asociadas a la brecha digital (Farahani y Ghasemi, 2024; Wilmers, 2024), suponiendo un nuevo frente de desigualdad, así como una falta de democratización del conocimiento (Pragya, 2024). Se añaden las implicaciones negativas de su uso para el medio ambiente, ya que el entrenamiento y la utilización de grandes modelos de IAG demandan cantidades significativas de energía, incrementando las emisiones de gases de efecto invernadero (Ponce del Castillo, 2024; Rane et al., 2024).

Es igualmente crucial reflexionar sobre cómo la IAG transforma y transformará los ámbitos y campos profesionales. Según Zarifhonarvar (2024), esta tecnología tendrá un gran impacto en sectores como, por ejemplo, la industria, la banca, la tecnología o las ciencias sociales. En consecuencia, afectará, y de hecho está afectando, al mercado laboral, provocando la eliminación de algunas profesiones y generando otras nuevas. De este modo, “la IAG tiene el potencial de cambiar la anatomía del trabajo, aumentando las capacidades de los trabajadores automatizando algunas de sus tareas” (Chui et al., 2023, p. 5), pero puede provocar un aumento significativo de las tasas de desempleo (Frey y Osborne, 2023).

Este complejo escenario implica que la dimensión ética de la utilización de la IAG adquiera una relevancia fundamental, particularmente en los esfuerzos

por establecer un marco regulatorio adecuado, como la nueva normativa europea conocida como Reglamento (UE) 2024/0138 sobre la Inteligencia Artificial de la UE que, adoptado por el Parlamento Europeo, establece normas coordinadas en esta materia.

En este contexto, la educación emerge como uno de los campos de mayor sensibilidad frente a la aplicación de la IAG, planteando grandes interrogantes sobre su transformación. Entre las cuestiones más relevantes destaca el papel del profesorado en este nuevo escenario (Chan y Tsi, 2024) y su incidencia en la adquisición de las competencias por parte del estudiantado (Lim et al., 2023). La UNESCO ha sido una de las principales organizaciones que ha planteado la necesidad de establecer un marco para abordar estas inquietudes, como, por ejemplo, en el Consenso de Beijing sobre IA y educación estableciendo recomendaciones sobre lo que implica un enfoque humanizado para el uso de la IA en el contexto educativo (UNESCO, 2019). Asimismo, la Recomendación sobre la ética de la inteligencia artificial (UNESCO, 2021a) proporciona un marco normativo de los aspectos críticos relacionados con la IAG, incluyendo la educación y la investigación. En esa misma línea, publicó la guía IA y educación (UNESCO, 2021b) que ofrece recomendaciones concretas para formular políticas dirigidas al uso adecuado de la IA en el ámbito educativo. Más recientemente, presentó una Guía para la IAG en educación e investigación (UNESCO 2023), con el foco en las implicaciones de su uso en dichos ámbitos.

Centrándonos en el ámbito de la Educación Superior (ES), y tal como señalan García-Peñalvo et al. (2024), esta tecnología también genera un impacto significativo. Instituciones, investigadores, profesorado y alumnado universitario se enfrentan a un cambio de paradigma (Huang et al., 2021), donde la velocidad de los cambios deja escaso margen para su asimilación (Grassini, 2023; Hwang y Chen, 2023). Esta situación ha llevado a la proliferación de estudios como el de Al Shloul et al. (2024), que analiza cómo las universidades se están posicionando ante herramientas como ChatGPT, o analizando la incidencia de la IAG en los entornos en línea de las universidades (Wang y Lei, 2024).

Tal y como destaca la Conferencia de Rectores de las Universidades Españolas (CRUE), “las herramientas de IAG dan la posibilidad de que los estudiantes puedan recibir más y mejor formación. Estas herramientas pueden actuar ya en etapas previas, contribuyendo a facilitar el acceso al mundo universitario” (Cruz Argudo et al., 2024, p. 7). No obstante, la IAG plantea desafíos importantes debido a su capacidad para generar textos académicos de gran complejidad que puede conllevar que el alumnado realice un trabajo sin necesidad de desarrollar competencias como la de la indagación, el pensamiento crítico o la capacidad de discernir entre la información real y la falsa (Rahman y Watanobe, 2023).

Todo ello interpela a los distintos actores del sistema educativo en la búsqueda de nuevas formas de trabajar y establecer nuevos parámetros legales y éticos.

Así, el aumento significativo de las investigaciones abarca temáticas variadas como la percepción del alumnado universitario sobre las implicaciones de su uso

(Johnston et al., 2024), la mejora del rendimiento académico y la motivación (Mohamed et al., 2024), o la aceptación de esta tecnología por el estudiantado y el personal académico (Kanont et al., 2024; Shahzad et al., 2024). También sobre sus posibles efectos negativos, como la proliferación del plagio académico (Cotton et al., 2024) o su incidencia en los procesos psicológicos y la regulación de emociones como la ansiedad producida por su uso (Kaya et al., 2024).

A partir del amplio marco de investigación existente, este trabajo pretende, como objetivo principal, indagar en el conocimiento sobre la IAG y las herramientas basadas en ella de los y las estudiantes de una universidad en línea, atendiendo también a los criterios éticos y las principales preocupaciones que su uso les suscita.

Contexto

Desde la adaptación al EEES en 2009, la Universitat Oberta de Catalunya (UOC) imparte la asignatura transversal “Competencias TIC” (CTIC) en la que los estudiantes adquieren de forma gradual e integrada la competencia digital, fundamental en el caso de la UOC por ser una universidad íntegramente en línea, y en la que los estudiantes desarrollan un proyecto digital en equipo según el modelo ABPCL (Romero et al., 2024).

En el segundo semestre del curso 23/24 se incorporó un recurso de aprendizaje sobre la IAG, que refuerza un debate virtual evaluable sobre las implicaciones de su uso en el ámbito académico, que es el objeto de este estudio. Los datos recogidos permitieron establecer una comparación entre los dos semestres, dado que durante el primero de ellos no se trató la incidencia de la IAG al no disponer del recurso ni realizarse el debate.

El diseño del recurso (Sebastià, 2024) fue coordinado por el profesorado de la asignatura ante un uso cada vez más generalizado de la IAG para la realización de actividades académicas. La asignatura CTIC es el marco ideal para introducir este conocimiento, dado que se imparte en los distintos grados de la UOC, contemplando la dimensión ética de la competencia digital, promoviendo una utilización responsable en el desarrollo de actividades de aprendizaje y atendiendo siempre a los retos que esta implica.

Con esta finalidad, el recurso se estructura en los siguientes apartados:

1. ¿Qué es la inteligencia artificial generativa y qué nos aporta?
2. Herramientas de la IAG
3. ¿Cómo interactuar con la IA?
4. Criterios éticos para el uso de la IA
5. Usos más extendidos de la IA en el aprendizaje

El apartado “Criterios éticos para el uso de la IA”, adquiere especial importancia en este estudio. Se estructura en dos partes: una en la que se detallan los desafíos que

comporta la IAG, como la desinformación o la seguridad y privacidad de los datos, y otra, bajo el título “Consideraciones que hay que tener en cuenta como estudiantado de la UOC”, que ofrece pautas para su utilización en el ámbito académico como aplicar la normativa establecida, citar dónde se ha utilizado la IAG, contrastar críticamente la información aportada, no introducir datos personales, confidenciales o protegidos y planificar su uso para evitar la dependencia de la herramienta (Sebastià, 2024).

Como se ha mencionado anteriormente, y como actividad que implica una lectura analítica y profunda del recurso expuesto, los y las estudiantes participaron en un debate propuesto a partir del siguiente caso ficticio de trabajo en equipo:

“Un grupo de estudiantes está desarrollando un proyecto digital en equipo. Han distribuido las tareas de forma equitativa de manera que cada uno es el responsable del desarrollo de una parte del proyecto y, a su vez, cada miembro del equipo valida la elaboración de otro. Cristina revisa la tarea de Leo, detectando errores importantes tanto a nivel de coherencia del texto como de autenticidad de los datos y argumentos expuestos. Cuando le comenta este hecho, Leo contesta que ha resuelto su tarea empleando una herramienta de IAG que le ha redactado el contenido de toda la actividad. Dado que el enunciado del trabajo advierte sobre el plagio, y que cree que se está vulnerando la normativa académica, Cristina opina que Leo debe rehacer su parte”.

Partiendo del caso, se plantean una serie de cuestiones relativas a la actuación de Leo, que los estudiantes debaten en el foro del aula, plantean posibles soluciones a la situación generada fomentando una reflexión compartida sobre cómo utilizar este tipo de herramientas sin vulnerar la normativa académica y aportando valor a los aprendizajes.

Por lo tanto, en la asignatura CTIC se trata el uso de la IAG desde una perspectiva reflexiva y, a la vez, buscando su aplicación en la práctica de actividades académicas.

METODOLOGÍA

El objetivo central de esta investigación es indagar en el conocimiento que el estudiantado de ES tiene de la IAG. Con este fin, la investigación se estructura siguiendo tres preguntas de investigación:

- ¿Qué conocimiento autopercebido de la IAG tiene el estudiantado?
- ¿Qué variables influyen en este conocimiento?
- ¿Cuáles son los elementos éticos fundamentales para el alumnado en el uso de la IAG en la educación superior?

Se sigue una metodología mixta integrando datos cualitativos y cuantitativos tanto en la recogida de información como en su posterior análisis (Creswell, 2014; Zhou et al., 2024).

Para la obtención de los datos se diseñó y aplicó un cuestionario en línea en cada uno de los semestres del curso 2023-24. El alumnado del primer semestre no disponía del recurso mencionado ni realizó una actividad evaluable de debate, mientras que el segundo semestre sí tuvo acceso al recurso y llevó a cabo la actividad. La participación de estos dos grupos diferenciados permite abordar el estudio comparativo que se presenta en las sucesivas secciones.

El cuestionario incluía preguntas cerradas de múltiple opción, una pregunta cerrada con respuesta en escala de Likert del 1 al 5, y preguntas abiertas sobre distintos aspectos relacionados con la IAG. Se incluyeron preguntas demográficas (formación máxima obtenida, estudio que se cursa, edad y género). Fue validado por nueve expertos especializados en tecnología y educación del grupo de investigación Edul@b.

La población de estudio está formada por el alumnado de la asignatura CTIC, transversal en los programas de Grado de la UOC, con 3.574 estudiantes matriculados en el primer semestre y 2.532 en el segundo semestre del curso 2023-24. La recogida de datos se realizó durante la finalización de ambos semestres de forma confidencial y anónima, mediante un formulario de Google. La muestra obtenida es el resultado de la participación voluntaria de los estudiantes, siendo representativa de cada uno de los estudios donde se imparte la asignatura. Incluye al 21 % de la población del primer semestre y al 11 % de la población del segundo, con un nivel de confianza del 95 % y un margen de error del 3,32 % y del 5,93 % respectivamente. La Tabla 1 presenta dichos datos por estudio:

Tabla 1
Datos poblacionales y muestrales del conjunto de participantes distribuidos por estudio

		Total	E1	E2	E3	E4	E5
			Derecho y Ciencia Polí- tica	Economía y Empresa	Artes y Hu- manidades	Psicología y Ciencias de la Educac.	Informática y Telecomu- nicaciones
1er semestre	Población	3242	882	891	298	738	433
	Muestra	686	157	216	64	146	103
	% pobl.	21	18	24	21	20	24
2º semestre	Población	2181	539	534	227	532	349
	Muestra	243	46	74	31	38	54
	% pobl.	11	9	14	14	7	15

Se analizaron cuantitativamente los datos con el programa estadístico de software libre JASP.

Atendiendo a la perspectiva cualitativa de la investigación, los estudiantes pudieron manifestar aquellos aspectos éticos que consideran más importantes en el uso de la IAG contestando a la pregunta abierta: “Menciona tres elementos fundamentales para utilizar la IAG de forma ética y responsable”, que permite ahondar en la tercera pregunta de investigación.

Utilizando el software libre QCAmap, las respuestas se fueron codificando en vivo en ambos grupos (utilización del recurso y no utilización) hasta que no emergieron nuevas categorías. La codificación axial de las trece categorías así obtenidas se estableció entre cuatro investigadores siguiendo el criterio de similitud conceptual entre ellas, triangulando posteriormente el resultado en contraste con la literatura, dando lugar a las siete categorías finales y su definición (Tabla 3 sección Resultados) con las que se completó el análisis de la totalidad de las respuestas (Bonilla-García y López-Suárez, 2016). La Figura 1 muestra un ejemplo de codificación utilizando QCAmap: la categoría inicial Plagio se integra durante la categorización axial en la categoría CAT 2 (Honestidad y propiedad intelectual).

Figura 1

Ejemplo del proceso de codificación

Eso si, hay que tener en cuenta también los plagios que puedan llegar a surgir por ciertas IAs, como lo es la DaVinci, puesto que a lo largo de unos pocos meses mejora muchísimo la calidad de las imágenes que llegaba a generar de



El mismo software se utilizó para el análisis de los debates generados a partir del caso, mediante una búsqueda intencionada de las categorías emergidas del análisis de los cuestionarios, citándose alguna de las intervenciones para enriquecer la discusión.

RESULTADOS Y DISCUSIÓN

Los resultados se organizan siguiendo las preguntas de investigación. En cuanto a la primera, ¿Qué conocimiento autopercibido de la IAG tiene el estudiantado?, se analiza el grado de conocimiento percibido discriminando entre los dos semestres, dado que el uso del recurso es una variable que podría influir en los resultados.

El cuestionario formula la pregunta “¿Conoces la Inteligencia Artificial Generativa?” con cinco niveles de respuesta en escala de Likert, que definen la variable Conocimiento:

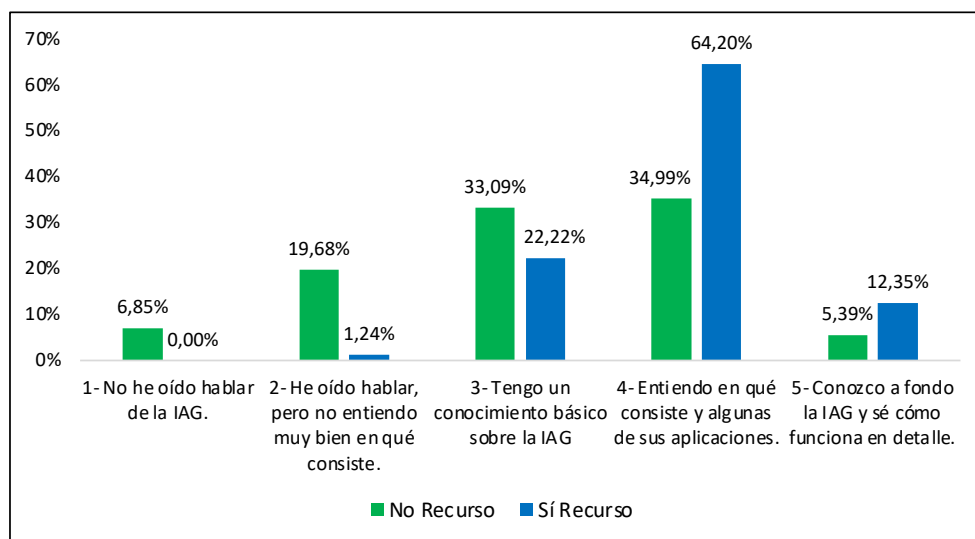
1. No he oído hablar de la IAG.

2. He oído hablar de ello, pero no entiendo muy bien en qué consiste.
3. Tengo un conocimiento básico sobre la IAG.
4. Entiendo en qué consiste y algunas de sus aplicaciones.
5. Conozco a fondo la IAG y sé cómo funciona en detalle.

Se observan diferencias significativas ($t = -10.922$, $p < 0.001$) entre quienes no han utilizado el recurso (No recurso, $M = 3,12$; $SD = 1,01$) y quienes sí lo han hecho (Sí recurso, $M = 3,57$; $SD = 0,85$). Estos resultados sitúan al primer grupo en un estado medio de conocimiento básico (nivel 3) sobre la IAG y al segundo grupo más cerca del nivel 4: “Entiendo en qué consiste y algunas de sus aplicaciones”. Para estudiar la relación entre estas dos variables con más detalle, la Figura 2 muestra los resultados por niveles la variable Conocimiento respecto a la variable Recurso.

Figura 2

Resultados de la variable Conocimiento para los grupos No recurso y Sí recurso por niveles



Cabe destacar que un 26,53 % del grupo No recurso declara no haber oído hablar de la IAG o haber oído hablar de ella, pero no entender en qué consiste (agrupando los niveles 1 y 2 de la escala), mientras el porcentaje baja hasta un 1,24 % en el grupo Sí recurso. Así, una cuarta parte de estudiantes adquieren como mínimo un conocimiento básico sobre la IAG con la utilización del recurso. Por otro lado, quienes accedieron al recurso y al debate los valoran muy positivamente, con un 3,96 y un 4,01 sobre 5 respectivamente.

Destacan por la diferencia relativa los resultados del nivel 4, pasando de un 34,99 % a un 64,20 % con el uso del recurso, mientras que los resultados conjuntos de los niveles 4 y 5 indican cómo el porcentaje de estudiantes que declaran un mayor conocimiento pasa del 40,38 % al 76,55 %. Estos resultados refuerzan cómo la formación específica sobre IAG se muestra eficaz a la hora de aumentar el conocimiento sobre la misma (Ruiz Mendoza et al., 2024).

Para profundizar en estos resultados, y dando respuesta a la segunda pregunta de investigación, ¿Qué variables influyen en este conocimiento?, la Tabla 2 presenta los resultados medios obtenidos para la variable Conocimiento en relación con las variables Formación, Edad y Estudio.

Tabla 2

Resultados medios para la variable Conocimiento en relación con las variables Formación, Edad, y Estudio

	Formación					Edad					Estudio				
	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
M	3,32	3,21	3,55	3,41	3,92	3,09	3,29	3,31	3,32	3,47	3,24	3,3	3,52	3,05	3,64
SD	0,94	1,01	0,92	0,95	0,97	1,04	0,96	0,99	0,99	0,9	1,01	0,92	0,93	0,99	0,91

Nota: Formación: 1- Secundaria, 2- Grado, 3- Máster, 4- Posgrado, 5- Doctorado. Edad (años): 1- menos de 20, 2- de 20 a 30, 3- de 30 a 40, 4- de 40 a 50, 5- más de 50. Estudio: 1- Derecho, 2- Empresa, 3- Humanidades, 4- Psicología y Educación, 5- Informática y Telecomunicaciones.

En la variable Formación, en los grupos de menor nivel formativo (Secundaria y Grado) las medias son algo menores a las de los grupos de mayor formación (Máster, Postgrado y Doctorado), siendo $M = 3,27$ para el primer grupo y $M = 3,55$ para el segundo. El test ANOVA refleja diferencias significativas respecto a la formación, pero quedando delimitadas post hoc (p tukey = 0,004) a la diferencia entre Secundaria ($M = 3,21$) y Grado ($M = 3,55$). Se desprende que el nivel de formación académica previa no es una variable que explique las diferencias encontradas en el conocimiento de la IAG.

Respecto a la variable Edad, la mayor diferencia se da entre los menores de 20 años ($M = 3,09$) y los mayores de 50 ($M = 3,47$), pero no siendo estas diferencias significativas ($t = -10,922$, $p < 0,001$).

En relación con las diferencias encontradas en la variable Estudio (Tabla 2), los análisis realizados mediante pruebas ANOVA y post hoc Tukey revelan diferencias significativas entre las siguientes combinaciones de disciplinas académicas: Derecho e Informática ($p = 0,005$), Economía y Empresa e Informática ($p = 0,002$), Humanidades y Psicología ($p < 0,001$) y Psicología e Informática ($p < 0,001$). Estos

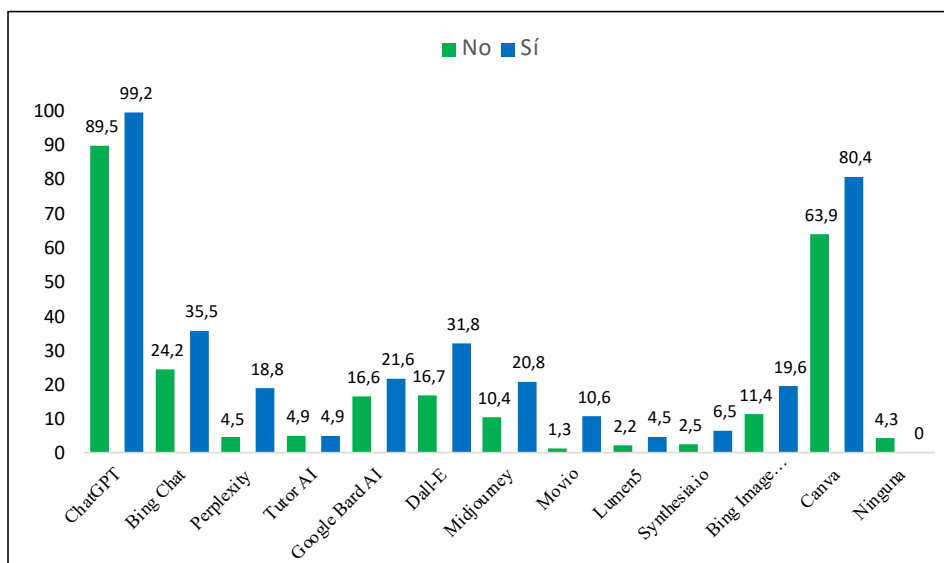
resultados evidencian variaciones sustanciales en la variable conocimiento según la disciplina académica cursada.

Ordenando los estudios de mayor a menor valoración media del conocimiento sobre la IAG se encuentra la siguiente disposición: 5- Informática y Telecomunicaciones, 3- Humanidades, 1- Derecho, 2- Empresa, 4- Psicología y Educación. Destaca la disciplina relacionada con las ingenierías del ámbito informático, aspecto lógico, teniendo en cuenta sus conocimientos técnicos respecto a la IAG. El segundo lugar de los estudios de Humanidades se explicaría por el perfil de sus estudiantes, que cursan su segunda carrera y preocupados por los aspectos éticos y sociales que interpela la IAG. Llama la atención la última posición de los estudios de Psicología y Educación, que puede atribuirse a cierta reticencia del uso de la IAG en el campo académico de la Psicología, en consonancia con lo señalado por Newell (2023) al plantear la oralidad como forma de paliar su uso en el contexto académico.

Estos resultados coinciden con investigaciones como la de Wang y Li (2024) que determinaron diferencias significativas en el impacto de la intención de uso de la IAG entre distintas disciplinas académicas. En una línea parecida, Stöhr et al. (2024) encontraron también diferencias significativas entre la familiaridad de los estudiantes con herramientas de IAG y su ámbito de estudio.

Se analiza ahora la relación de la variable Recurso con el conocimiento de las diferentes herramientas de IAG. La Figura 3 muestra el porcentaje de respuestas para cada una de las herramientas en ambos grupos (No Recurso y Sí Recurso).

Figura 3
Conocimiento de las herramientas basadas en IAG



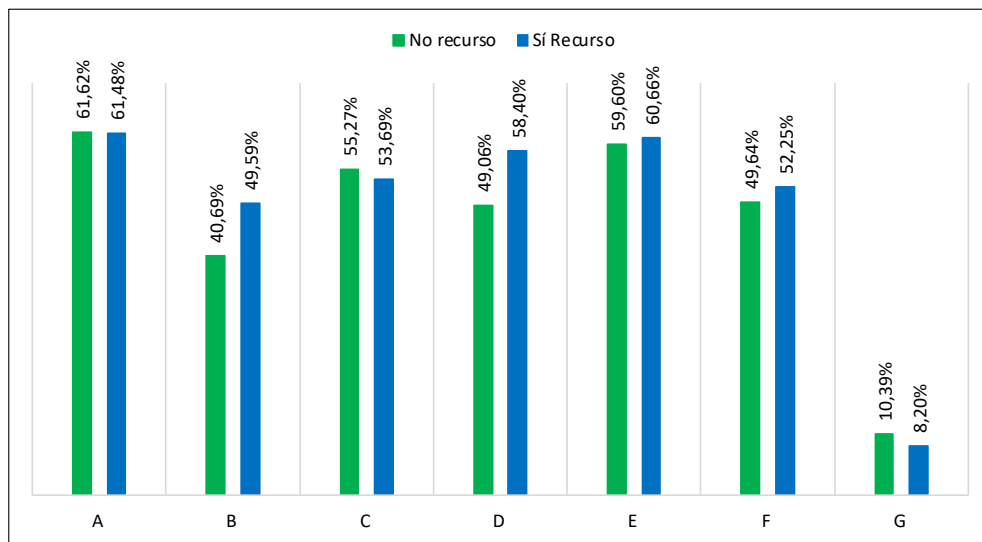
ChatGPT destaca como la herramienta más conocida, con casi un 90 % de quienes no habían utilizado el recurso, pasando a prácticamente un 100 % en el caso de los que sí lo han utilizado, en consonancia con la literatura científica, que constata cómo ChatGPT es la herramienta más utilizada para la realización de actividades académicas (Acosta-Enriquez et al., 2024; Al Shloul et al., 2024; Strzelecki, 2023).

Estos resultados confirman la utilidad del recurso para aumentar el conocimiento de herramientas concretas. Llama la atención el aumento respecto al resto de algunas de ellas, como la herramienta de diseño gráfico Canva (de 63,92 % a 80,41 %). Puede atribuirse a su uso previo a la implementación de funcionalidades de IA, ya que era una herramienta conocida en la asignatura, como confirman estudios como el de Ruiz-Rojas et al. (2024), que destaca su impacto en actividades colaborativas. Otras herramientas con más menciones son Perplexity (de 4,47 % a 18,78 %), un motor de búsqueda que utiliza el lenguaje natural, o Dall-E (de 16,74 % a 31,84 %) herramienta para crear imágenes a partir de una descripción textual.

Una vez analizadas las diferencias en el conocimiento de la IAG entre el estudiantado, se pasa a valorar su visión sobre distintos aspectos éticos relacionados con la utilización de la IAG, dando respuesta a la tercera pregunta de investigación ¿Cuáles son los elementos éticos fundamentales para el alumnado en el uso de la IAG en el ámbito académico? En este caso, se pidió a los estudiantes que indicaran cuáles de los siguientes aspectos críticos les resultaban más preocupantes, pudiendo escoger una o más respuestas:

- A. El desconocimiento que existe en torno a la IAG y de las consecuencias de su uso.
- B. Proporcionar mis datos a la herramienta IAG.
- C. La IAG puede crear contenido que sea engañoso u ofensivo.
- D. La información que me proporciona puede estar sesgada.
- E. Desinformación y manipulación (deepfakes).
- F. Vulneración de los derechos de autor y propiedad intelectual.
- G. No tengo claro que sea legal hacerlo.

La Figura 4 muestra los resultados obtenidos para cada uno de los grupos de la variable Recurso:

Figura 4*Preocupación por los aspectos críticos en el uso de la IAG*

Los aspectos que más preocupan a ambos grupos, y sobre los que apenas presentan diferencias, son el A: El desconocimiento que existe en torno a la IAG y de las consecuencias de su uso, el E: Desinformación y manipulación (deepfakes) y el C: La IAG puede crear contenido que sea engañoso u ofensivo. Este hecho se explica a partir de la alarma social que existe con base en estos aspectos (Łabuz y Nehring, 2024; Shoaib et al., 2023) abordada también en el ámbito universitario (Roe et al., 2024).

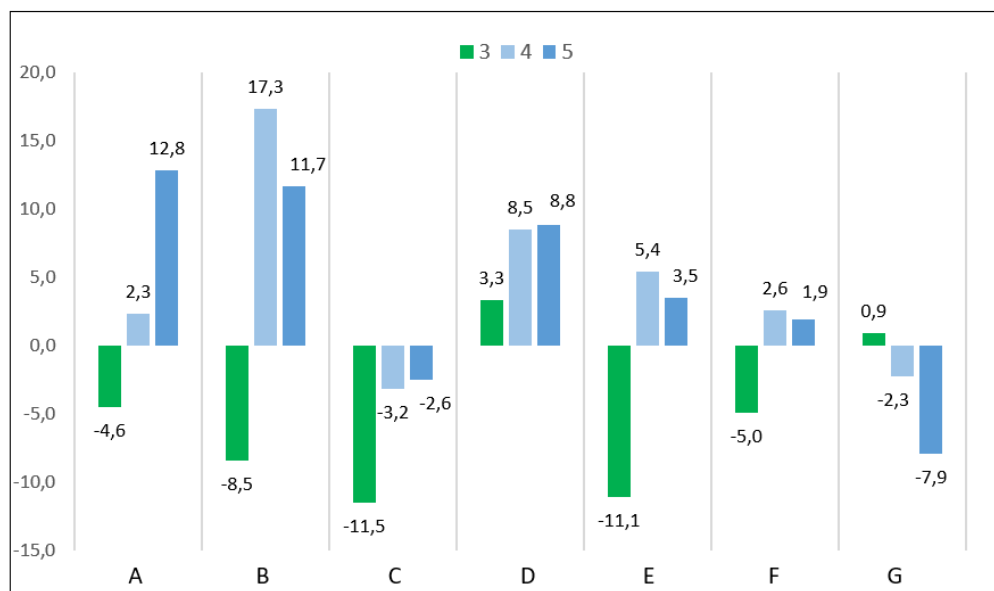
Por otro lado, los dos aspectos críticos cuya valoración presenta un mayor cambio son el B: Proporcionar mis datos a la herramienta IAG, y el D: La información que me proporciona puede estar sesgada. Se observa en ambos un aumento con la inclusión del recurso (8,90 y 9,34 puntos respectivamente). La valoración del aspecto G: No tengo claro que sea legal hacerlo es significativamente más baja en ambos grupos, por lo que, aunque en el recurso se plantean desafíos legales en cuanto a su uso, no ha tenido incidencia en este aspecto.

Profundizando en el análisis de la influencia del uso del recurso y la participación en el debate, la Figura 5 presenta la relación entre las variables Conocimiento y Aspectos Críticos. Al no encontrarse apenas estudiantes en los niveles 1 y 2 una vez han utilizado el recurso (ver Figura 2), el análisis se centra en las diferencias observadas entre los dos grupos (media grupo Sí - media grupo No) para los niveles 3, 4 y 5. Se observa por ejemplo que, para el grupo con un nivel de conocimiento 3, la utilización de recurso provoca una disminución de su preocupación por el aspecto

crítico A en 4,6 puntos respecto a quienes no habían utilizado el recurso, mientras que aumenta en los grupos de nivel de conocimiento 4 y 5 en 2,3 y 12,8 puntos respectivamente. Este patrón se repite en los aspectos críticos B, E y F, mientras que en el aspecto crítico G se invierte. Este último dato puede quedar explicado por la formulación del aspecto G en negativo. En el aspecto D, la preocupación aumenta para todos los niveles, así que sólo los participantes del nivel 3 cambian el patrón anterior. Sobre el aspecto C (La IAG puede crear contenido que sea engañoso u ofensivo), se observa una disminución en la preocupación para todos los niveles de conocimiento, atribuible a que en este caso se les interpela como usuarios de la IA al generar su propio contenido.

Figura 5

Variación en la preocupación por los aspectos críticos de la IAG



Estos resultados confirman que la formación en IAG modifica significativamente la forma en la que los estudiantes universitarios se posicionan ante diferentes aspectos críticos que su uso conlleva. Para entender las causas subyacentes en el cambio del nivel de preocupación para cada uno de los aspectos críticos, sería necesario llevar a cabo un estudio en mayor profundidad.

Durante el estudio cualitativo emergieron las categorías presentadas en la Tabla 3, donde también se presenta su definición y un ejemplo de la literatura.

Así, por ejemplo, en la categoría Análisis crítico y contraste de la información (CAT 1) los participantes hablan de “Verificar la información contrastándola con otras fuentes” y en la literatura, como en Gallent-Torres et al. (2023), se constata la necesidad de que, para utilizar la IAG, se debe contar con unas competencias que permitan evaluar la calidad de los resultados.

Tabla 3

Categorías emergentes sobre el uso de la IAG de forma ética y responsable

ID	Categoría	Definición
CAT 1	Análisis crítico y contraste de la información	Contrastar distintas fuentes y analizar críticamente la información obtenida. (Gallent-Torres et al., 2023)
CAT 2	Honestidad y propiedad intelectual	Citar el origen de la información obtenida, tanto de la aportada en texto como de las ideas utilizadas, siendo explícito con el uso hecho de la IAG. (Hadi et al., 2024)
CAT 3	Protección de datos y seguridad	No exponer datos personales, propios o ajenos, y tomar precauciones garantizando la seguridad al usar la IAG. (Kajiwarra y Kawabata, 2024).
CAT 4	Responsabilidad y propósitos adecuados	Utilizar la IAG con responsabilidad y propósitos adecuados, evitando malos usos. (Jobin et al., 2019).
CAT 5	Uso de la IAG como complemento	La IAG debe ser una ayuda para realizar algunas tareas sin sustituir el proceso de aprendizaje ni la interacción humana. (Chan, 2023)
CAT 6	Conocimiento y conciencia de sus implicaciones	Formación para estudiantes y profesores en el uso e integración efectiva de IAG, incluyendo conceptos, habilidades y aspectos éticos en su creación y uso, fomentando la toma de consciencia en un entorno transparente (Lee et al., 2024)
CAT 7	Regulación	Utilizarla con base en unas normas establecidas, bien adoptando una regulación superior, bien estableciéndose en el contexto académico correspondiente. (Parlamento Europeo, 2024)

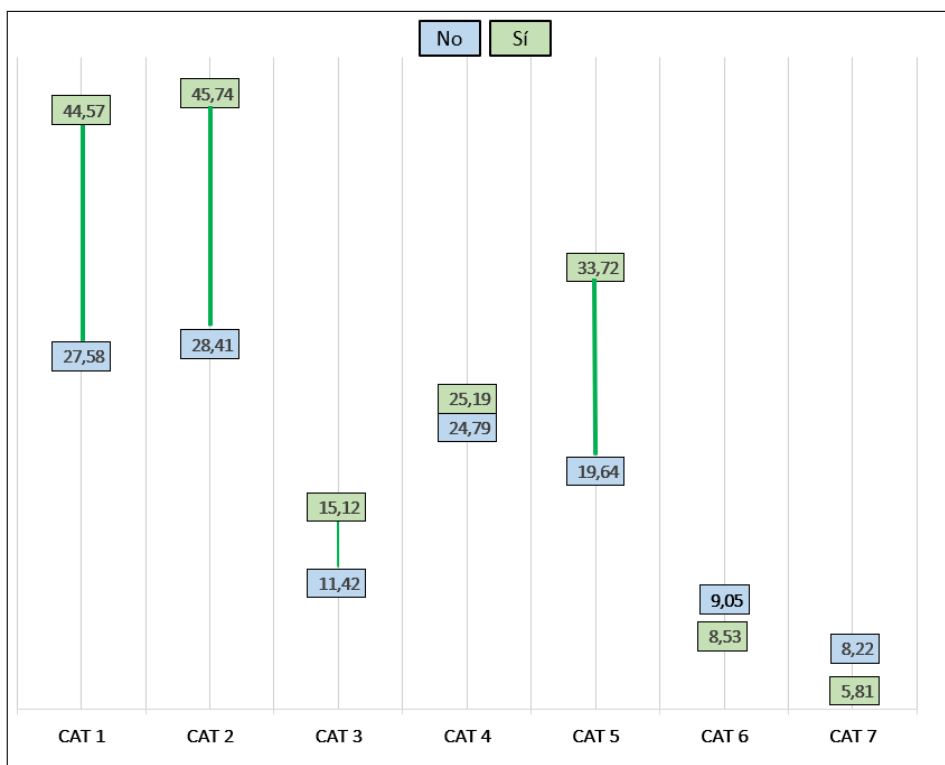
Se agruparon las respuestas en blanco junto a las referidas a una falta de conocimiento y las que no guardaban relación con la pregunta. En este caso se observa una disminución en casi ocho puntos porcentuales entre ambos grupos (26,5 % para el grupo No y 18,1 % para el grupo Sí) por lo que la inclusión del recurso y la realización del debate en la asignatura CTIC ha permitido mejorar el conocimiento del estudiantado sobre la IAG.

El resultado de analizar con las categorías definidas las respuestas a la pregunta “Menciona tres elementos fundamentales para utilizar la IAG de forma ética y

responsable” se muestra en la Figura 6, indicando en porcentaje el número de veces que se menciona cada categoría respecto al número total de respuestas obtenidas discriminando si han utilizado el Recurso sobre IAG y han realizado la actividad de debate, o no lo han hecho.

Figura 6

Porcentaje en el que los participantes mencionan las categorías emergentes sobre el uso de la IAG de forma ética y responsable



En el estudiantado que sí ha utilizado el recurso se observa un aumento considerable en el porcentaje de tres de las categorías: CAT 1 (Análisis crítico y contraste de la información), CAT 2 (Honestidad y propiedad intelectual) y CAT 5 (Uso de la IAG como complemento), coincidiendo con la revisión de la literatura realizada por Gallent-Torres et al. (2023) donde se analizan publicaciones relacionadas con experiencias pedagógicas para potenciar el uso ético de la IAG en la ES (Kong et al., 2023). Para el resto de categorías las diferencias se sitúan en un rango muy pequeño de (0,4 - 3,7) puntos porcentuales.

En ambos grupos las categorías más mencionadas son CAT 2 (Honestidad y propiedad intelectual) y CAT 1 (Análisis crítico y contraste de la información), así como las tres menos mencionadas CAT 3 (Protección de datos y seguridad), CAT 6 (Conocimiento y consciencia de sus implicaciones) y CAT 7 (Regulación). La tercera categoría más mencionada en el grupo Sí es CAT 5 (Uso de la IAG como complemento), siendo CAT 4 (Responsabilidad y propósitos adecuados) en el grupo No.

Se discuten ahora, por orden de mayor a menor valoración media, las categorías representadas en la Figura 6 complementadas con los testimonios del estudiantado a partir de sus intervenciones en la actividad de debate del caso práctico, e indicadas como est. 1, etc.

La categoría más mencionada en ambos grupos es CAT 2: Honestidad y propiedad intelectual (M Sí = 45,74 %, M No = 28,41 %, dif. = 17,33 puntos). El plagio, consistente en dar como propia una creación ajena sin citar la fuente, es una de las mayores preocupaciones del profesorado concernientes al uso de la IAG en ES (Lee et al., 2024). Los estudios sobre la relación del alumnado con el plagio son insuficientes (Gallent-Torres et al., 2023) pero sí reflejan lo generalizado que está y, en menor medida, la preocupación por sus consecuencias (Sullivan et al., 2023).

Como estrategias para minimizar el plagio en la universidad, se mencionan la formación (Cebrián-Robles et al., 2018), o la integración del uso ético de la IAG en los propios planes de estudio (Kajiware y Kawabata, 2024; Lim et al., 2023), estrategia que también se menciona en los debates que el alumnado de esta universidad mantuvo: “fomentar un enfoque equilibrado que promueva su utilización ética y responsable, junto con la formación adecuada para estudiantes y profesores sobre cómo utilizarla de manera efectiva” (est. 25).

La segunda categoría más mencionada es CAT 1: Análisis crítico y contraste de la información (M Sí = 44,57 %, M No = 27,58 %, dif. = 16,99 puntos). El aumento de casi 17 puntos entre quienes sí utilizaron el recurso y quienes no, revela cómo los estudiantes eran menos conscientes del papel proactivo que requiere el uso de la IAG. En los debates este es uno de los elementos que se menciona con más frecuencia, también en relación con el pensamiento crítico: “se debe fomentar el aprendizaje activo y la creatividad entre los estudiantes... la realización de actividades que requieran pensamiento crítico” (est. 54). Siendo el pensamiento crítico una de las competencias clave para el siglo XXI (Martínez-Bravo et al., 2021), la problemática suscitada por la irrupción de la IAG refuerza su importancia.

La tercera categoría más mencionada en el grupo Sí recurso, con una diferencia relevante entre ambos grupos es CAT 5: Uso de la IAG como complemento. (M Sí = 33,72 %, M No = 19,64 %, dif. = 14,08 puntos). Esta categoría está muy ligada a las dos categorías analizadas anteriormente, como muestran estas dos citas extraídas de los debates entre el alumnado: “es esencial que los estudiantes entiendan que la IA debe ser utilizada como un complemento, no como un reemplazo de su propio pensamiento crítico y creativo” (est. 23), o “soy la primera en utilizar herramientas

de IA para ampliar información, pero sólo como información complementaria” (est. 8). Un enfoque positivo de esta categoría es aprovechar la capacidad de la IAG para agilizar muchas de las tareas: generación de lluvia de ideas sobre un tema, traducciones, herramientas de autoaprendizaje, etc. (Andión y Cárdenas, 2023).

Se comentan a continuación brevemente los resultados del resto de categorías, en las que las diferencias entre ambos grupos son pequeñas:

La valoración relativamente alta de la categoría Responsabilidad y propósitos adecuados (CAT 4) habla de una conciencia social en ambos grupos sobre la finalidad con la que se utiliza IAG y la propia responsabilidad, que no ha sido influida por la utilización del recurso o el debate, y que sí se menciona en el plan docente de la asignatura: “ha realizado un mal uso de la IA, el plan docente es claro con respecto a esta cuestión y él no lo ha tenido en consideración” (est. 2).

Sorprende sin embargo la baja mención de la Protección de datos y seguridad (CAT 3), dada la sensibilidad social sobre este aspecto.

Las categorías menos destacadas, Conocimiento y consciencia de sus implicaciones (CAT 6) y Regulación (CAT 7), pueden entenderse como elementos externos al usuario. Aunque en el caso presentado y en los propios debates se menciona el cumplimiento de la normativa existente: “es necesario avanzar sobre su uso mediante normativas y regulaciones” (est. 11), la categoría “Regulación” se refiere a la demanda de más control institucional y no a una implicación ética de los usuarios.

CONCLUSIONES

Los resultados presentados en este trabajo reflejan que el tratamiento conferido a la IAG en la asignatura contribuye significativamente a aumentar el conocimiento del estudiantado. Se validan tanto el recurso como la actividad reflexiva de debate a partir de un caso práctico, ratificándolos como estrategias útiles.

En cuanto a las variables que influyen en el conocimiento del estudiantado sobre la IAG, puede señalarse que la edad y el nivel formativo previo no tienen impacto estadísticamente significativo. Sin embargo, la disciplina de estudio sí que influye de forma relevante, destacando las ingenierías del ámbito informático y las humanidades frente a áreas como psicología y educación. Estos resultados subrayan la necesidad de adaptar el enfoque sobre la IAG a disciplinas específicas, especialmente en aquellas menos familiarizadas con esta tecnología, para hacer consciente a todo el estudiantado del potencial y los riesgos de la IAG en el ámbito académico.

La formación implementada ha incrementado de manera generalizada el grado de conciencia y conocimiento del estudiantado sobre los aspectos críticos que el uso de la IAG conlleva en el ámbito académico. Este aspecto se observa especialmente en cuestiones como la honestidad intelectual y el uso responsable de estas herramientas. No obstante, los estudiantes no manifestaron preocupación significativa por los desafíos legales asociados al uso de la IAG.

A raíz de los resultados presentados, y desde el punto de vista de su aplicación práctica, se hace necesario un rediseño de las actividades de la asignatura. En primer lugar, se identifica la necesidad de evolucionar el caso práctico objeto de debate, incorporando explícitamente algún aspecto crítico en el ámbito legal. En segundo lugar, es necesario diseñar una actividad de análisis colaborativa de los términos y condiciones de las aplicaciones de IAG, con el fin de reflexionar sobre las implicaciones legales de su uso (como, por ejemplo, la gestión de sus datos). Y, en tercer lugar, con la finalidad de que el estudiantado utilice la IAG de forma estructurada, se planteará la realización de parte de las búsquedas de información con una de estas herramientas, de tal manera que apliquen y reflexionen sobre cómo afecta al proceso de producción académica. Cabe decir que las tres propuestas se implementarán en el próximo semestre.

Las principales limitaciones del estudio son el posible sesgo de la muestra, al ser una participación voluntaria, y el desbalance de la relación muestra/población entre el grupo que sí accedió al recurso y el que no lo hizo, aunque ambas limitaciones se atenúan con el amplio número de respuestas. El estudio se ha realizado en un contexto concreto, por lo que su generalización queda supeditada a realizar estudios similares en otros ámbitos, como podrían ser la educación secundaria o universidades presenciales.

En cuanto a las líneas de investigación futuras, parece pertinente explorar con mayor profundidad las diferencias observadas por disciplina de estudios, además de analizar las causas en las diferencias observadas en el grado de percepción sobre los aspectos críticos y éticos del uso académico de la IAG.

También sería relevante investigar cómo la irrupción de estas herramientas influye en la adquisición de las competencias digitales por parte del estudiantado, añadiendo instrumentos de evaluación objetivos de su aprendizaje como instrumento de investigación.

Otra línea de investigación incluiría el estudio de las percepciones del profesorado de la asignatura para detectar sus preocupaciones y mejorar la calidad de su tarea docente.

En definitiva, y considerando el poco tiempo transcurrido desde la irrupción generalizada de la IAG, los resultados permiten avanzar en un nuevo marco para ayudar a evolucionar la enseñanza y aprendizaje del uso de la IAG y sus implicaciones a través de una asignatura de una universidad en línea, mediante un modelo metodológico validado.

REFERENCIAS

Acosta-Enriquez, 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), 1-18. <https://doi.org/10.1186/s40359-024-01764-z>

- Al Shloul, T., Mazhar, T., Abbas, Q., Iqbal, M., Ghadi, Y. Y., Shahzad, T., Mallek, F. y Hamam, H. (2024). Role of activity-based learning and ChatGPT on students' performance in education. *Computers and Education: Artificial Intelligence*, 6, 100219. <https://doi.org/10.1016/j.caeai.2024.100219>
- Andión, M. y Cárdenas, D. I. (2023). Convivir con inteligencias artificiales en la educación superior: Retos y estrategias. *Perfiles Educativos*, 45(Especial), 56-69. <https://doi.org/10.22201/iisue.24486167e.2023.Especial.61691>
- Bonilla-García, M. Á. y López-Suárez, A. D. (2016). Ejemplificación del proceso metodológico de la teoría fundamentada. *Cinta de moebio*, (57), 305-315. <https://doi.org/10.4067/S0717-554X2016000300006>
- Cebrián-Robles, V., Raposo-Rivas, M., Cebrián-de-la-Serna, M. y Sarmiento-Campos, J. A. (2018). Percepción sobre el plagio académico de estudiantes universitarios españoles. *Educación XX1*, 21(2), 105-129. <https://doi.org/10.5944/educxx1.20062>
- Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. *International Journal of Educational Technology in Higher Education*, 20(1), 1-25. <https://doi.org/10.1186/s41239-023-00408-3>
- Chan, C. K. Y. y Tsi, L. H. Y. (2024). Will generative AI replace teachers in higher education? A study of teacher and student perceptions. *Studies in Educational Evaluation*, 83, 101395. <https://doi.org/10.1016/j.stueduc.2024.101395>
- Chui, M., Hazan, E., Roberts, R., Singla, A., Smaje, K., Sukharevsky, A. Yee, L. y Zimmel, R. (2023). *The economic potential of generative AI: The next productivity frontier*. McKinsey Global Institute. <http://dln.jaipuria.ac.in:8080/jspui/bitstream/123456789/14313/1/The-economic-potential-of-generative-ai-the-next-productivity-frontier.pdf>
- Cotton, D. R. E., Cotton, P. A. y Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 61(2), 228-239. <https://doi.org/10.1080/14703297.2023.2190148>
- Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). SAGE Publications.
- Cruz Argudo, F., García Varea, I., Martínez Carrascal, J. A., Ruiz Martínez, A., Ruiz Martínez, P. M., Sánchez Campos, A. y Turró Ribalta, C. (2024). *La Inteligencia Artificial Generativa en la docencia universitaria: Oportunidades, desafíos y recomendaciones*. Crue Universidades Españolas. https://www.crue.org/wp-content/uploads/2024/03/Crue-Digitalizacion_IA-Generativa.pdf
- Farahani, M. y Ghasemi, G. (2024). Artificial Intelligence and Inequality: Challenges and Opportunities. *International Journal of Innovation in Education*, 9, 78-99. <https://doi.org/10.32388/7hwuz2>
- Frey, C. B. y Osborne, M. (2023). Generative AI and the Future of Work: A Reappraisal. *The Brown Journal of World Affairs*, XXX(1). <https://bjwa.brown.edu/30-1/generative-ai-and-the-future-of-work-a-reappraisal/>
- Gallent-Torres, C., Zapata-González, A. y Ortego-Hernando, J. L. (2023). El impacto de la inteligencia artificial generativa en la educación superior: una mirada desde la ética y la integridad académica. *RELIEVE*, 29(2), art. M5. <https://doi.org/10.30827/relieve.v29i2.29134>
- García-Peñalvo, F. J., Llorens-Largo, F. y Vidal, J. (2024). La nueva realidad de la educación ante los avances de la inteligencia artificial generativa. *RIED-Revista Iberoamericana de Educación*

- a Distancia, 27(1), 9-39. <https://doi.org/10.5944/ried.27.1.37716>
- Grassini, S. (2023). Shaping the Future of Education: Exploring the Potential and Consequences of AI and ChatGPT in Educational Settings. *Education Sciences*, 13(7), 692. <https://doi.org/10.3390/educsci13070692>
- Hadi, M. U., Tashi, Qasem Al, Qureshi, R., Shah, A., Muneer, Amgad, Irfan, M., Zafar, A., Shaikh, M. B., Akhtar, N., Wu, J. y Mirjalili, S. (2024). A Survey on Large Language Models: Applications, Challenges, Limitations, and Practical Usage [Preprint]. TechRxiv. <https://doi.org/10.36227/techrxiv.23589741.v1>
- Huang, J., Shen, G. y Xiping, R. (2021). Connotation Analysis and Paradigm Shift of Teaching Design under Artificial Intelligence Technology. *International Journal of Emerging Technologies in Learning*, 16(5), 73-86. <https://doi.org/10.3991/ijet.v16i05.20287>
- Hwang, G. J. y Chen, N. S. (2023). Editorial Position Paper: Exploring the Potential of Generative Artificial Intelligence in Education: Applications, Challenges, and Future Research Directions. *Educational Technology and Society*, 26(2). [https://doi.org/10.30191/ETS.202304_26\(2\).0014](https://doi.org/10.30191/ETS.202304_26(2).0014)
- Jarrahi, M. H., Askay, D., Eshraghi, A. y Smith, P. (2023). Artificial intelligence and knowledge management: A partnership between human and AI. *Business Horizons*, 66(1), 87-99. <https://doi.org/10.1016/j.bushor.2022.03.002>
- Jobin, A., Ienca, M. y Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1, 389-399. <https://doi.org/10.1038/s42256-019-0088-2>
- Johnston, H., Wells, R. F., Shanks, E. M., Boey, T. y Parsons, B. N. (2024). Student perspectives on the use of generative artificial intelligence technologies in higher education. *International Journal for Educational Integrity*, 20(1), 1-21. <https://doi.org/10.1007/s40979-024-00149-4>
- Kajiwar, Y. y Kawabata, K. (2024). AI literacy for ethical use of chatbot: Will students accept AI ethics? *Computers and Education: Artificial Intelligence*, 6, 100251. <https://doi.org/10.1016/j.caeai.2024.100251>
- 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 Special Issue), 18-33. <https://doi.org/10.34190/ejel.22.6.3196>
- Kaya, F., Aydin, F., Schepman, A., Rodway, P., yetişensoy, O. y Demir Kaya, M. (2024). The Roles of Personality Traits, AI Anxiety, and Demographic Factors in Attitudes toward Artificial Intelligence. *International Journal of Human-Computer Interaction*, 40(2), 497-514. <https://doi.org/10.1080/10447318.2022.2151730>
- Kong, S.-C., Cheung, W. M.-Y. y Zhang, G. (2023). Evaluating an Artificial Intelligence Literacy Programme for Developing University Students' Conceptual Understanding, Literacy, Empowerment and Ethical Awareness. *Educational Technology & Society*, 26(1), 16-30. [https://doi.org/10.30191/ETS.202301_26\(1\).0002](https://doi.org/10.30191/ETS.202301_26(1).0002)
- Labuz, M. y Nehring, C. (2024). Information apocalypse or overblown fears-what AI mis- and disinformation is all about? Shifting away from technology toward human reactions. *Politics and Policy*, 52(4), 874-891. <https://doi.org/10.1111/polp.12617>
- Lee, D., Arnold, M., Srivastava, A., Plastow, K., Strelan, P., Ploeckl, F. y Palmer, E. (2024). The impact of generative AI on higher education learning and teaching:

- A study of educators' perspectives. *Computers and Education: Artificial Intelligence*, 6, 100221. <https://doi.org/10.1016/j.caeai.2024.100221>
- Lim, W. M., Gunasekara, A., Pallant, J. L., Pallant, J. I. y Pechenkina, E. (2023). Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators. *The International Journal of Management Education*, 21(2), 1-13. <https://doi.org/10.1016/j.ijme.2023.100790>
- Martínez-Bravo, M. C., Sádaba Chalezquer, C. y Serrano-Puche, J. (2021). Meta-marco de la alfabetización digital: análisis comparado de marcos de competencias del Siglo XXI. *Revista Latina de Comunicación Social*, 79, 76-110. <https://doi.org/10.4185/RLCS-2021-1508>
- Mohamed, A. M., Shaaban, T. S., Bakry, S. H., Guillén-Gámez, F. D. y Strzelecki, A. (2024). Empowering the Faculty of Education Students: Applying AI's Potential for Motivating and Enhancing Learning. *Innovative Higher Education*, 1-23. <https://doi.org/10.1007/s10755-024-09747-z>
- Newell, S. J. (2023). Employing the interactive oral to mitigate threats to academic integrity from ChatGPT. *Scholarship of Teaching and Learning in Psychology*. <https://doi.org/10.1037/stl0000371>
- Parlamento Europeo. (2024). *Reglamento de Inteligencia Artificial (2024/0138(COD))*. https://www.europarl.europa.eu/doceo/document/TA-9-2024-0138_ES.pdf
- Ponce del Castillo, A. (2024). *Exposing generative AI: Human-Dependent, Legally Uncertain, Environmentally Unsustainable* [Preprint]. SSRN Electronic Journal. <https://doi.org/10.2139/ssrn.4975411>
- Pragya, A. (2024). Generative AI and epistemic diversity of its inputs and outputs: call for further scrutiny. *AI and Society*, 1-2. <https://doi.org/10.1007/s00146-024-02097-6>
- Rahman, M. M. y Watanobe, Y. (2023). ChatGPT for Education and Research: Opportunities, Threats, and Strategies. *Applied Sciences*, 13(9), 5783. <https://doi.org/10.3390/app13095783>
- Raman, R., Kumar Nair, V., Nedungadi, P., Kumar Sahu, A., Kowalski, R., Ramanathan, S. y Achuthan, K. (2024). Fake news research trends, linkages to generative artificial intelligence and sustainable development goals. *Heliyon*, 10(3), e24727. <https://doi.org/10.1016/j.heliyon.2024.e24727>
- Rane, N., Choudhary, S. y Rane, J. (2024). *Contribution of ChatGPT and Similar Generative Artificial Intelligence for Enhanced Climate Change Mitigation Strategies* [Preprint]. SSRN Electronic Journal. <https://doi.org/10.2139/ssrn.4681720>
- Roe, J., Perkins, M. y Furze, L. (2024). Deepfakes and Higher Education: A Research Agenda and Scoping Review of Synthetic Media. *Journal of University Teaching and Learning Practice*, 1-22. <https://doi.org/10.53761/2y2np178>
- Romero Carbonell, M., Romeu Fontanillas, T., Guitert Catasús, M. y Baztán Quemada, P. (2024). Validation of the OCPBL model for online collaborative project-based learning [Validación del modelo ABPCL para el aprendizaje basado en proyectos colaborativos en línea]. *RIED-Revista Iberoamericana de Educación a Distancia*, 27(2), 159-181. <https://doi.org/10.5944/ried.27.2.39120>
- Ruiz Mendoza, K. K., Miramontes Arteaga, M. 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>
- Ruiz-Rojas, L. I., Salvador-Ullauri, L. y Acosta-Vargas, P. (2024). Collaborative

- Working and Critical Thinking: Adoption of Generative Artificial Intelligence Tools in Higher Education. *Sustainability*, 16(13), 5367. <https://doi.org/10.3390/su16135367>
- Ruschmeier, H. (2024). Generative AI and Data Protection. En C. Poncibo, M. Ebers, R. Calo, y M. Zou (Eds.), *Handbook on Generative AI and the Law* (pp. 1-18). Cambridge University Press. <https://doi.org/10.1017/cfl.2024.2>
- Sebastià, I. (2024). *Inteligencia artificial generativa*. Universitat Oberta de Catalunya. <https://iag.recursos.uoc.edu/es/>
- Shahzad, M. F., Xu, S. y Asif, M. (2024). Factors affecting generative artificial intelligence, such as ChatGPT, use in higher education: An application of technology acceptance model. *British Educational Research Journal*, 1-25. <https://doi.org/10.1002/berj.4084>
- Shi, Y. (2023). Study on security risks and legal regulations of generative artificial intelligence. *Science of Law Journal*, 2(11), 17-23. <https://doi.org/10.23977/law.2023.021104>
- Shoaib, M. R., Wang, Z., Ahvanooey, M. T. y Zhao, J. (2023). Deepfakes, Misinformation, and Disinformation in the Era of Frontier AI, Generative AI, and Large AI Models. *ICCA 2023 - 2023 5th International Conference on Computer and Applications, Proceedings*. <https://doi.org/10.1109/ICCA59364.2023.10401723>
- Stöhr, C., Ou, A. W. y Malmström, H. (2024). Perceptions and usage of AI chatbots among students in higher education across genders, academic levels and fields of study. *Computers and Education: Artificial Intelligence*, 7, 100259. <https://doi.org/10.1016/j.caeai.2024.100259>
- Strzelecki, A. (2023). To use or not to use ChatGPT in higher education? A study of students' acceptance and use of technology. *Interactive Learning Environments*, 32(9), 5142-5155. <https://doi.org/10.1080/10494820.2023.2209881>
- Sullivan, M., Kelly, A. y McLaughlan, P. (2023). ChatGPT in higher education: Considerations for academic integrity and student learning. *Journal of Applied Learning & Teaching*, 6(1), 1-10. <https://doi.org/10.37074/jalt.2023.6.1.17>
- UNESCO. (2019). *Beijing Consensus on Artificial Intelligence and Education*. <https://unesdoc.unesco.org/ark:/48223/pf0000368303>
- UNESCO. (2021a). *Ethics of Artificial Intelligence: The Recommendation*. <https://www.unesco.org/en/artificial-intelligence/recommendation-ethics?hub=32618>
- UNESCO. (2021b). *Inteligencia artificial y educación: guía para las personas a cargo de formular políticas*. <https://unesdoc.unesco.org/ark:/48223/pf0000379376>
- UNESCO. (2023). *Guidance for generative AI in education and research*. <https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research>
- Wach, K., Duong, C. D., Ejdy, J., Kazlauskaitė, R., Korzynski, P., Mazurek, G., Paliszkievicz, J. y Ziemba, E. (2023). The dark side of generative artificial intelligence: A critical analysis of controversies and risks of ChatGPT. *Entrepreneurial Business and Economics Review*, 11(2), 7-30. <https://doi.org/10.15678/EBER.2023.110201>
- Wang, L. y Li, W. (2024). The Impact of AI Usage on University Students' Willingness for Autonomous Learning. *Behavioral Sciences*, 14(10), 956. <https://doi.org/10.3390/bs14100956>
- Wang, X. y Lei, L. (2024). A Path Study of Generative Artificial Intelligence Enabling Online Education Platforms in Colleges and Universities. *Proceedings of the 2024 International Symposium on Artificial Intelligence*

- for Education, 332-338. <https://doi.org/10.1145/3700297.3700354>
- Wilmers, N. (2024). *Generative AI and the Future of Inequality*. An MIT Exploration of Generative AI. MIT. <https://doi.org/10.21428/e4baedd9.777b7123>
- Zarifhonarvar, A. (2024). Economics of ChatGPT: a labor market view on the occupational impact of artificial intelligence. *Journal of Electronic Business & Digital Economics*, 3(2), 100-116. <https://doi.org/10.1108/jebde-10-2023-0021>
- Zhou, Y., Zhou, Y. y Machtmes, K. (2024). Mixed methods integration strategies used in education: A systematic review. *Methodological Innovations*, 17(1), 41-49. <https://doi.org/10.1177/20597991231217937>

Fecha de recepción del artículo: 1 de diciembre de 2024

Fecha de aceptación del artículo: 11 de marzo de 2025

Fecha de aprobación para maquetación: 28 de marzo de 2025

Fecha de publicación en OnlineFirst: 21 de abril de 2025

Fecha de publicación: 1 de julio de 2025

IA generativa y pensamiento crítico en la educación universitaria a distancia: desafíos y oportunidades

Generative AI and critical thinking in online higher education: challenges and opportunities



- ID César Muñoz Martínez - *Universidad Nacional de Educación a Distancia, UNED (España)*
ID Vanessa Roger-Monzo - *Universitat de València, UV (España)*
ID Fernando Castelló-Sirvent - *Universitat Politècnica de València, UPV (España)*

RESUMEN

La inteligencia artificial generativa (IAG) está transformando la educación universitaria, especialmente en entornos virtuales donde el predominio de actividades asincrónicas exige que los estudiantes gestionen activamente su aprendizaje. Su integración plantea desafíos y oportunidades para los docentes, quienes desempeñan un papel fundamental en el desarrollo del pensamiento crítico y requieren habilidades tecnopedagógicas para garantizar un uso ético y reflexivo de estas herramientas. Este estudio exploratorio analiza la incorporación de la IAG en la educación a distancia desde cinco dimensiones: barreras que limitan el pensamiento crítico, aceleradores que pueden impulsarlo, alternativas tecnológicas, retos sociales y consecuencias de fomentarlo. Se empleó un enfoque cualitativo basado en entrevistas semiestructuradas con once expertos en educación superior, los resultados identifican tres barreras principales: la falta de formación docente en IAG y pensamiento crítico, la resistencia institucional y la ausencia de directrices claras. No obstante, la alfabetización digital, la innovación pedagógica y la adaptación de los sistemas de evaluación pueden mitigar estos obstáculos. Entre las alternativas tecnológicas, se propone el desarrollo de *edu-chatbots* en entornos controlados y la implementación de marcos para analizar sesgos algorítmicos. Sin embargo, persisten retos como garantizar un acceso equitativo y evitar una dependencia acrítica. Como contribución, se proponen cinco vectores de acción que orientan la integración de la IAG y el diseño de políticas pedagógicas. Su implementación requiere una estrategia coordinada entre instituciones, docentes y responsables académicos, de modo que la automatización generada por la IA no solo optimice los procesos educativos, sino que también actúe como catalizador del pensamiento crítico.

Palabras clave: inteligencia artificial en *e-learning* (AIeL); pensamiento crítico; metacognición; educación universitaria a distancia; ética de la IA; brecha digital.

ABSTRACT

Generative artificial intelligence (GAI) is reshaping higher education, particularly in virtual learning environments where the prevalence of asynchronous activities requires students to take an active role in managing their own learning. Its integration presents both challenges and opportunities for educators, who not only support critical thinking but also need techno-pedagogical skills to guide its ethical and reflective use. This exploratory study examines the incorporation of GAI into distance education across five dimensions: barriers that limit critical thinking, factors that can enhance it, available socio-technological alternatives, social challenges and broader implications of strengthening this skill. A qualitative approach was used, based on semi-structured interviews with eleven higher education experts. The findings highlight key obstacles, including limited teacher training in GAI and critical thinking, institutional resistance and a lack of clear guidelines. However, digital literacy, pedagogical innovation and adapted assessment methods can help overcome these barriers. Among the proposed solutions are the development of *edu-chatbots* in controlled university environments and frameworks to assess algorithmic biases. Even so, ensuring equitable access and avoiding an uncritical reliance on AI persist as notable challenges. This study contributes by proposing five action areas to support educators and academic decision-makers in integrating GAI and shaping educational policies. Its implementation requires collaboration between institutions, faculty and policymakers to ensure that AI-driven automation not only enhances educational processes but also fosters critical thinking meaningfully.

Keywords: artificial intelligence in *e-learning* (AIeL); critical thinking; metacognitive awareness; online higher education; AI ethics; digital divide.

INTRODUCCIÓN

La inteligencia artificial generativa (IAG) está transformando los sistemas educativos, impactando principalmente en las áreas de administración, enseñanza y aprendizaje (Chassignol et al., 2018). En entornos de educación a distancia o en línea, estas herramientas optimizan la gestión académica al facilitar tareas como la resolución de preguntas frecuentes y las comunicaciones institucionales. Asimismo, los sistemas de tutorías inteligentes automatizan la corrección de pruebas de evaluación continua, proporcionando retroalimentación personalizada basada en rúbricas diseñadas previamente por el profesorado (Tang et al., 2021).

El uso de la IAG como herramienta pedagógica permite a los docentes personalizar las rutas de aprendizaje mediante la creación de contenidos adaptativos, además de simplificar el diseño y la evaluación de pruebas ajustadas a las necesidades de cada estudiante (Bhutoria, 2022; Romero Alonso et al., 2025; Zhang et al., 2021). Agentes conversacionales como *chatbots* y *cobots* pueden resolver dudas, distribuir materiales y ofrecer retroalimentación personalizada, fomentando entornos más centrados en el estudiante (Adiguzel et al., 2023); complementariamente, su integración con la realidad virtual posibilita simulaciones inmersivas que favorecen la comprensión práctica de conceptos complejos (Timms, 2016).

Estas herramientas permiten identificar deficiencias en el aprendizaje y abordarlas para minimizar su impacto en el desempeño académico a largo plazo (Ocumpaugh et al., 2024). Mediante analíticas detalladas, la IAG puede rastrear patrones, detectar áreas problemáticas y ofrecer recomendaciones personalizadas que optimicen el proceso educativo (Drugova et al., 2024). Asimismo, ofrece a los estudiantes información sobre su desempeño y competencias, ayudándolos a elegir programas formativos que se ajusten a sus intereses y aspiraciones profesionales (Chen et al., 2020).

Desde una perspectiva institucional, la IAG facilita el monitoreo en tiempo real del progreso académico, anticipando necesidades de aprendizaje y fortaleciendo la autorregulación mediante modelos como el *Open Learner Model* (OLM) y las técnicas de *Knowledge Tracing* (KT). Estos enfoques se alinean con los modelos *e-learning*, reconocidos por su flexibilidad y capacidad para ofrecer experiencias educativas más inclusivas y personalizadas (Garcés y Bastías, 2025; Ilic et al., 2023).

No obstante, su integración plantea desafíos significativos. La IAG puede amplificar sesgos algorítmicos, acentuar desigualdades en el acceso a la tecnología y reducir los espacios de reflexión crítica y creatividad estudiantil (Adiguzel et al., 2023). Además, el acceso inmediato a información generada por IA puede desincentivar el análisis autónomo y la formulación de conclusiones propias. El reto no es solo incorporar la IAG en la educación, sino hacerlo de manera crítica y responsable, garantizando la integridad académica (Eke, 2023; Kumar et al., 2024).

En este contexto, el pensamiento crítico es fundamental para analizar rigurosamente la retroalimentación generada por estas herramientas (Barrot,

2023). Mientras que muchas de las denominadas *hard skills* han sido delegadas a las máquinas por su carácter rutinario y técnico, las habilidades humanas más complejas¹ han cobrado mayor relevancia. Esta tendencia se refleja en el *Informe sobre el Futuro de los Empleos* (World Economic Forum, 2023), que destaca el pensamiento analítico y creativo entre las competencias más valoradas por los empleadores.

En los entornos laborales actuales, caracterizados por volatilidad, incertidumbre, complejidad y ambigüedad (entornos *VUCA*), los profesionales deben gestionar su aprendizaje de manera continua y autónoma (Aguilar Vargas et al., 2020). En este escenario, el pensamiento crítico constituye una herramienta imprescindible para reflexionar sobre los propios procesos cognitivos, tomar decisiones estratégicas y afrontar desafíos complejos con enfoques innovadores (Ayyıldız y Yılmaz, 2021).

En la educación universitaria a distancia, la integración de herramientas de IAG puede contribuir al desarrollo del pensamiento crítico; sin embargo, su efectividad está condicionada por las competencias y disposiciones del profesorado, cuya función mediadora resulta fundamental en este proceso. Por ello, aunque diversos *stakeholders* intervienen en los sistemas educativos, este estudio sitúa al docente como eje central de acción.

Partiendo de esta premisa, se formuló la siguiente pregunta de investigación: ¿Qué barreras, impulsores y desafíos éticos enfrentan los docentes en la promoción del pensamiento crítico en entornos mediados por IA? Para abordarla, se adoptó una metodología cualitativa basada en entrevistas semiestructuradas organizadas en torno a cinco dimensiones. Este estudio exploratorio busca identificar los obstáculos que limitan el desarrollo del pensamiento crítico entre los estudiantes, los factores que pueden impulsarlo, los beneficios potenciales de la IAG y las estrategias para mitigar los desafíos sociales y éticos derivados de su uso no guiado.

REVISIÓN DE LA LITERATURA

Diversas investigaciones señalan que, aunque herramientas como *ChatGPT* presentan limitaciones en el pensamiento de orden superior frente a los humanos (Deng y Lin, 2022; Guo et al., 2023), estas brechas podrían reducirse con el avance tecnológico (Liu et al., 2023). Asimismo, estudios recientes han examinado la integración del pensamiento crítico (Cananau et al., 2025) y la alfabetización digital (Ng et al., 2023) en normativas educativas y planes de formación docente y alfabetización digital.

En este contexto, la revisión de la literatura se enfoca en la instrucción docente en entornos mediados por IAG, abordando las bases conceptuales del pensamiento crítico, sus herramientas de evaluación y las intervenciones pedagógicas que integran IA para potenciar su desarrollo. Este enfoque proporciona un marco integral para comprender su implementación en los escenarios educativos contemporáneos.

Muñoz Martínez, C., Roger-Monzo, V., & Castelló-Sirvent, F. (2025). Generative AI and critical thinking in online higher education: challenges and opportunities. [IA generativa y pensamiento crítico en la educación universitaria a distancia: desafíos y oportunidades]. *RIED-Revista Iberoamericana de Educación a Distancia*, 28(2), pp. 233-273. <https://doi.org/10.5944/ried.28.2.43556>

El pensamiento crítico como una habilidad metacognitiva

La psicología cognitiva aporta claves para comprender la progresión del pensamiento crítico, desde procesos básicos hasta niveles superiores. La taxonomía de Bloom et al. (1956, p. 12) estructura los “actos mentales o procesos de pensamiento derivados de las experiencias educativas” en una jerarquía que distingue entre Habilidades de Pensamiento de Orden Inferior (*LOTS*) como Conocimiento, Comprensión y Aplicación, y Habilidades de Orden Superior (*HOTS*) como Análisis, Síntesis y Evaluación.

Aunque la taxonomía no menciona explícitamente la metacognición, sus niveles superiores están estrechamente vinculados con las concepciones actuales del pensamiento metacognitivo (Wegerif, 2002). En este marco, el pensamiento crítico se considera un hábito mental esencial dentro de la metacognición, ya que permite a los individuos reflexionar sobre sus propios procesos cognitivos y regular su aprendizaje. La metacognición, definida como la capacidad de monitorear, evaluar y ajustar el pensamiento, es clave para el desarrollo del pensamiento crítico (Flavell, 1976). Esta relación bidireccional se evidencia en habilidades como la evaluación, que no solo implica analizar la calidad de la información, sino también cuestionar los propios juicios y creencias. En otras palabras, pensar críticamente significa ejercitar la capacidad de “pensar sobre el pensamiento” (Flavell, 1979).

La metacognición abarca la conciencia y el control de los procesos emocionales y motivacionales que influyen en el aprendizaje y la toma de decisiones (Papleontiou-Louca, 2003). Además de facilitar la comprensión sobre cómo se procesan y organizan los conocimientos, también fortalece la autonomía al permitir que los estudiantes regulen y optimicen su aprendizaje de forma continua (Çakıcı, 2018; Choy y Cheah, 2009; Maor et al., 2023). Por consiguiente, el pensamiento crítico está intrínsecamente vinculado con habilidades metacognitivas (Kuhn y Dean, 2004) como la autorregulación y el uso de procesos cognitivos avanzados, tales como identificar sesgos, justificar conclusiones y proponer soluciones innovadoras (Ku y Ho, 2010). En este contexto, el pensamiento crítico puede entenderse como una manifestación avanzada de la metacognición, que permite gestionar el conocimiento de manera efectiva y afrontar problemas complejos con una perspectiva reflexiva y creativa.

Habilidades específicas del pensamiento crítico

Distinguir entre las habilidades específicas del pensamiento crítico y componentes relacionados, como la motivación y la metacognición, permite evitar solapamientos conceptuales y clarificar su alcance. Su estrecha relación con los procesos metacognitivos dificulta su delimitación como fenómenos independientes (Rivas et al., 2022). Desde esta perspectiva, resulta pertinente analizar las capacidades cognitivas que lo conforman.

Definir el pensamiento crítico es un desafío, pues integra múltiples habilidades interconectadas. Pasquinelli et al. (2021, p. 170) lo describen como “la capacidad de evaluar la calidad epistémica de la información disponible y calibrar la propia confianza para actuar en consecuencia”. Este enfoque resalta su carácter multidimensional y su conexión con habilidades cognitivas avanzadas.

Diversos marcos teóricos han identificado subhabilidades esenciales del pensamiento crítico (Halpern, 1998; Pascarella y Terenzini, 2005). No obstante, la falta de herramientas validadas y estandarizadas para su medición e implementación sigue siendo un desafío (Ku, 2009; Plummer et al., 2022). Entre las pruebas más utilizadas destacan la *Prueba de Habilidades de Pensamiento Crítico de California* (Facione, 1990), la *Prueba de Pensamiento Crítico de Cornell* (Ennis y Millman, 1985) y la *Evaluación Watson-Glaser* (Watson y Glaser, 1980). En este contexto, Facione (2023) desarrolló un marco teórico sólido con una rúbrica confiable para evaluar estas habilidades, el cual ha sido seleccionado como base conceptual del presente estudio (ver Apéndice; Tabla 1).

En entornos de aprendizaje en línea, las plataformas virtuales ofrecen un espacio ideal para fomentar el pensamiento crítico mediante estrategias como el cuestionamiento socrático, la argumentación, la resolución colaborativa de problemas y la evaluación por pares. Herramientas como foros, mapas conceptuales y entornos grupales favorecen dinámicas que estimulan la reflexión y el análisis crítico (Ertmer et al., 2011; Goodsett, 2020; MacKnight, 2000; Yang et al., 2008). La combinación de estos enfoques fortalece la aplicación práctica del pensamiento crítico en la educación virtual.

La IAG como impulsora o limitadora del desarrollo del pensamiento crítico

Lipman (1988) sostiene que el pensamiento crítico es una forma de inteligencia que puede enseñarse y aprenderse. En esta línea, Niu et al. (2013), a partir de un metaanálisis en el ámbito universitario, concluyeron que las intervenciones educativas favorecen su desarrollo. Al mismo tiempo, el auge de la tecnología en las aulas ha generado debate sobre su impacto en el aprendizaje. A medida que los docentes integran herramientas digitales, resulta imprescindible evaluar cómo las aplicaciones de IA afectan el desarrollo del pensamiento crítico (Delgado et al., 2015).

Las herramientas de IAG pueden potenciar el pensamiento crítico al generar experiencias de aprendizaje dinámicas e interactivas que fomentan la participación activa (Baidoo-Anu y Owusu Ansah, 2023). No obstante, su efectividad depende de marcos que guíen su implementación pedagógica y permitan analizar el desarrollo real de estas habilidades.

En este contexto, Shanto et al. (2024) propusieron el marco “AI-CRITIQUE” para fomentar el pensamiento crítico en entornos con IAG. Sin embargo, debido a sus limitaciones en flexibilidad y adaptabilidad, el presente estudio adopta el enfoque

de Yusuf et al. (2024), que estructura el aprendizaje en cinco fases interconectadas: familiarización, conceptualización, indagación, evaluación y síntesis (ver Apéndice; Tabla 2). Este modelo destaca la importancia de progresar desde procesos cognitivos básicos hacia niveles superiores, promoviendo un análisis crítico de la información generada por la IA.

Como muestra la Tabla 2, integrar la IAG en la evaluación ofrece una oportunidad para fomentar el pensamiento crítico a través de un aprendizaje práctico y personalizado. Estas herramientas amplían los enfoques sobre temas complejos, proporcionan retroalimentación inmediata e incorporan ejemplos, analogías y escenarios hipotéticos que estimulan el razonamiento crítico (Javaid et al., 2023). Mediante simulaciones y debates guiados, los estudiantes pueden desarrollar habilidades como la evaluación de argumentos, la identificación de falacias y la formulación de respuestas fundamentadas. Ello fortalecería su capacidad para estructurar ideas con coherencia, cuestionar supuestos y considerar perspectivas alternativas.

La evaluación de estas intervenciones combina diseños longitudinales y transversales. Los primeros emplean cuestionarios *ex ante* y *ex post* para medir cambios en las percepciones y habilidades de los estudiantes a lo largo del tiempo, mientras que los segundos incluyen grupos de control para comparar el impacto de la intervención entre estudiantes expuestos y no expuestos, identificando diferencias significativas atribuibles al uso de la IAG.

La evidencia empírica reciente respalda el potencial de estas herramientas en la educación superior. Estudios (Guo y Lee, 2023; Ruiz-Rojas et al., 2024) indican que la integración efectiva de IAG mejora la autopercepción de los estudiantes sobre su competencia en pensamiento crítico, con avances notables en la formulación de preguntas exploratorias, la evaluación rigurosa de información, la construcción de conclusiones lógicas y la comprensión de temas complejos.

No obstante, una integración inadecuada de estas herramientas conlleva riesgos. Fuchs (2023) advierte que una dependencia excesiva de la IAG sin comprender los conceptos subyacentes puede limitar el aprendizaje genuino (Ivanov, 2023). Este riesgo se evidencia en la investigación de Dilekli y Boyraz (2024), donde se solicitó a estudiantes de posgrado realizar una autoevaluación reflexiva comparando sus propios ensayos con los generados por *ChatGPT*. La mayoría aceptó la información proporcionada por la IA sin cuestionar ni verificar su fiabilidad, a pesar de haber cursado una asignatura sobre “Enseñanza de habilidades de pensamiento”.

Estos hallazgos previos refuerzan la necesidad de una supervisión docente activa para guiar a los estudiantes hacia un uso más reflexivo y crítico de la IAG. Sin una orientación adecuada, estas tecnologías pueden limitar el desarrollo del pensamiento crítico y la creatividad, ya que, ante la presión de plazos o la optimización de recursos, los estudiantes podrían asumir la información generada sin validarla, incluso cuando su precisión no esté garantizada.

METODOLOGÍA

Datos

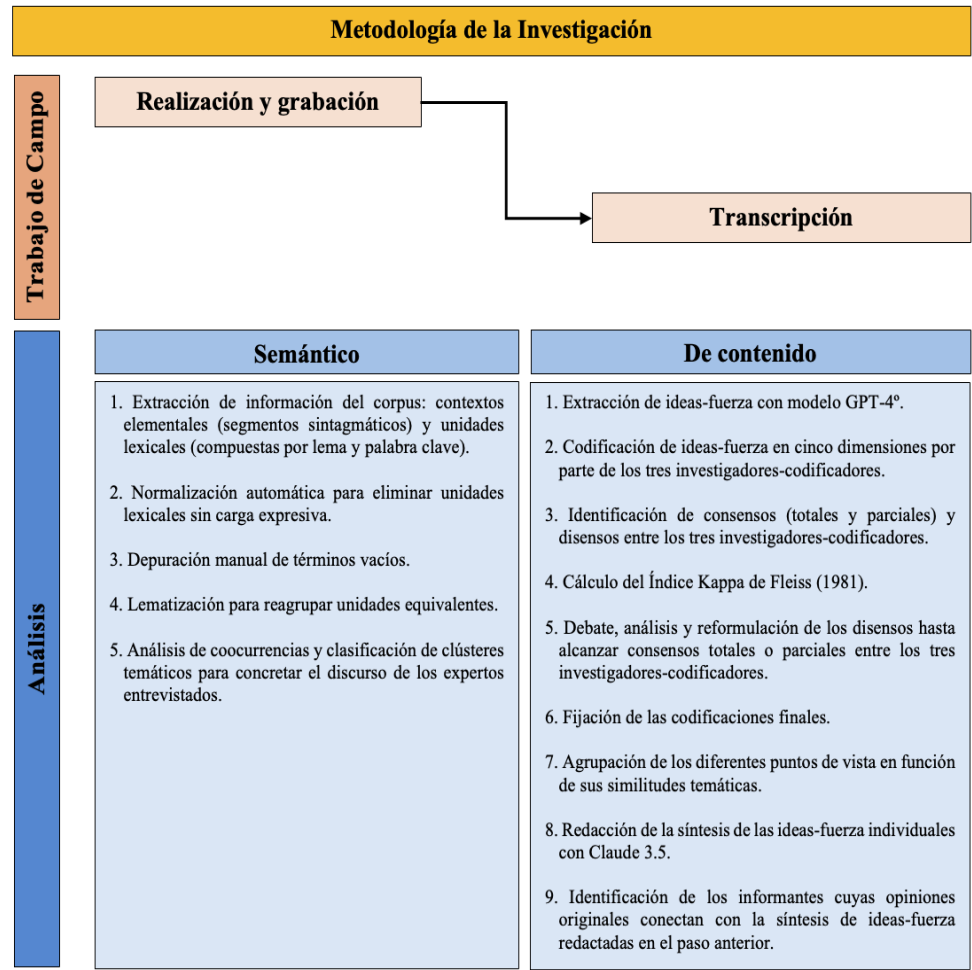
Se realizaron entrevistas semiestructuradas a once pedagogos y docentes de universidades presenciales y a distancia. Aunque se empleó un muestreo no probabilístico por conveniencia con una muestra reducida y de diversidad limitada, la homogeneidad de los participantes y la estructura de las entrevistas refuerzan la validez de los resultados. Young y Casey (2019) sostienen que muestras pequeñas y homogéneas identifican eficazmente códigos y temas, alcanzando representaciones significativas con 6-9 casos, mientras que 7-10 participantes resultan adecuados para temáticas complejas. Por su parte, Almasreh et al. (2019) sugieren un umbral próximo a los 10 expertos. La Tabla 3 (ver Apéndice) presenta el perfil ciego de los informantes utilizados en este estudio.

Las entrevistas siguieron el modelo pentagonal propuesto por De Vicente y Matti (2016) para los procesos de reflexión sistémica. En línea con los objetivos de esta investigación, las entrevistas se centraron en explorar el desarrollo del pensamiento crítico en el contexto del uso de la IAG por parte de estudiantes de universidades a distancia dentro del Espacio Europeo de Educación Superior. Este modelo estructuró las entrevistas en cinco bloques: (1) Barreras para el desarrollo del pensamiento crítico; (2) Aceleradores que podrían impulsarlo; (3) Alternativas socio-tecnológicas disponibles; (4) Retos sociales que se deben abordar; (5) Consecuencias de una mejora generalizada en la adquisición de esta competencia.

Las once entrevistas, realizadas en noviembre de 2024 a través de *Microsoft Teams*, tuvieron una duración promedio de 53 minutos, con una desviación estándar de 20 minutos, sumando un total de 10 horas y 54 minutos. Previamente, los participantes recibieron un formulario de consentimiento informado, elaborado conforme a los modelos del Comité de Ética de la UNED. Las sesiones fueron grabadas para su posterior transcripción y análisis, obteniendo un total de 52.720 palabras transcritas.

La Figura 1 muestra la secuencia-guía de la metodología empleada proporcionando una visión estructurada del proceso seguido para la recopilación y análisis de datos.

Figura 1
Secuencia-guía del proceso metodológico



Fuente: elaboración propia.

Análisis semántico

A partir de la transcripción de las entrevistas, se realizó un análisis textual de tipo cualitativo-cuantitativo mediante el *software T-LAB v.10.2.7*, que permite identificar patrones de palabras a través de aplicaciones estadísticas y gráficas.

Esta metodología ha sido ampliamente utilizada en el análisis de corpus lingüísticos en distintas disciplinas, como el estudio del discurso en medios sociales sobre cuestiones político-sociales (Gil y Guilleumas, 2017), el análisis de contenidos en psicología (Mazzoni et al., 2018) y el ámbito del turismo (Mondo y Gândara, 2017). Su versatilidad la convierte en una herramienta de gran potencial para la investigación científica (Cortini y Tria, 2014).

El *software* extrae información del corpus lingüístico mediante contextos elementales (segmentos textuales en unidades sintagmáticas) y unidades léxicas, compuestas por lemas y palabras clave. Se aplicó una normalización automática para eliminar lemas sin carga expresiva, complementada con una depuración manual de términos vacíos. Posteriormente, se llevó a cabo un proceso de lematización para agrupar unidades equivalentes; por ejemplo, las palabras clave “IA” e “inteligencia artificial” se agruparon bajo el mismo lema “IA”. Finalmente, se identificaron 120 lemas, lo que permitió realizar un análisis de coocurrencias y categorizar clústeres temáticos para estructurar el discurso de los expertos entrevistados.

Análisis de contenido

Además del análisis semántico, se realizó un análisis de contenido de las entrevistas. Se utilizó el modelo de lenguaje *GPT-4o* para asistir a los investigadores en la tarea de extracción de las principales ideas-fuerza expresadas por los expertos. Cada idea-fuerza se transcribió literalmente para conservar su precisión temática mínima. Posteriormente, se codificaron de forma independiente y se asignaron a las dimensiones del modelo pentagonal. Para garantizar la consistencia y fiabilidad del consenso entre los codificadores, se calculó la Kappa de Fleiss (Fleiss, 1981), considerando tres escenarios: (1) consenso total (los tres codificadores coincidieron); (2) consenso parcial (dos de tres coincidieron y se adoptó la opción mayoritaria); y (3) disenso (sin coincidencia). Los casos de disenso se resolvieron mediante discusión entre los investigadores hasta alcanzar consenso parcial o total antes de proceder a la codificación definitiva.

En la etapa siguiente, se empleó el modelo de lenguaje Claude 3.5 para apoyar la reformulación de la redacción de estas ideas-fuerza con el objeto de llevar a cabo una síntesis de los principales consensos identificados en la etapa anterior. Este proceso de búsqueda de consensos permitió condensar puntos de vista compartidos entre los informantes, articulando una sensibilidad común respecto a las cinco dimensiones analizadas: barreras, aceleradores, alternativas, retos sociales y consecuencias visibles. De este modo, se sintetizaron los resultados reflejando un espectro amplio de percepciones compartidas entre los participantes.

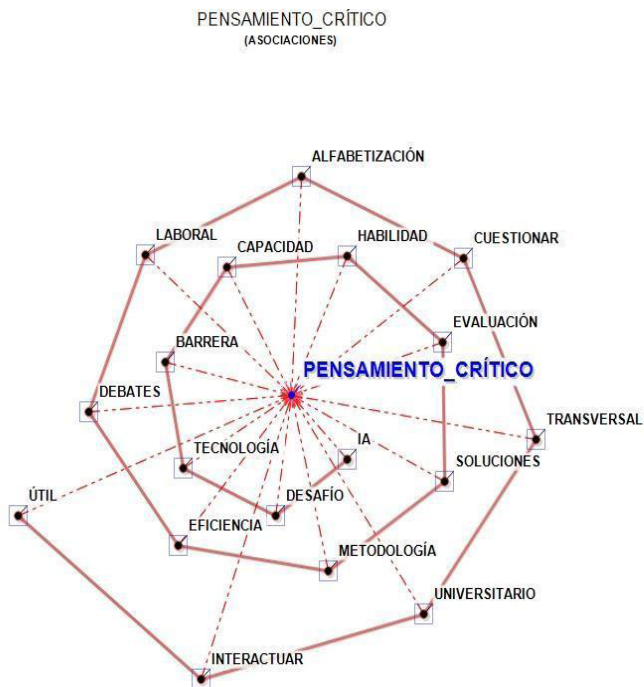
RESULTADOS, DISCUSIÓN E IMPLICACIONES

Análisis semántico de coocurrencias

A partir de los 120 lemas seleccionados del corpus lingüístico se desarrolló un análisis de coocurrencias para identificar asociaciones de palabras y calcular la frecuencia con la que dos o más lemas coincidían en contextos elementales idénticos. La Figura 2 reporta la vinculación entre el lema *pensamiento crítico*, eje del estudio, y otros lemas con los que muestra coocurrencias estadísticamente significativas (test χ^2 , $p < 0,05$).

Figura 2

Diagrama radial de asociación de lemas para pensamiento crítico



Fuente: elaboración propia.

En el diagrama radial, el lema *pensamiento crítico* (lema A) se ubica en el centro, mientras que el resto de lemas (lemas B) se distribuyen alrededor. Los lemas B más

próximos al centro presentan un mayor nivel de coocurrencias y los más alejados presentan asociaciones menos frecuentes. Los lemas con un nivel de coocurrencia más elevado con respecto a *pensamiento crítico* (lema A) son: *IA²*, *desafío*, *tecnología*, *barrera* y *capacidad*. La concentración de coocurrencias en torno a estos lemas indica que el pensamiento crítico está situado en una intersección entre IAG y la innovación educativa.

Por otra parte, los lemas *alfabetización*, *habilidad* y *capacidad* hacen referencia a la importancia de una alfabetización digital y crítica como base para desarrollar competencias avanzadas. Del mismo modo, los lemas *evaluación* y *cuestionar* conectan con la necesidad de integrar el pensamiento crítico en los sistemas de evaluación y en el diseño de actividades educativas que fomenten el análisis y la valoración de información generada por IA.

Las resistencias al cambio que enfrenta el desarrollo del pensamiento crítico en el contexto de la IAG se reflejan en los lemas *desafío* y *barrera*. A su vez, *laboral* y *solución* destacan la necesidad de aplicar el pensamiento crítico no solo como una competencia, sino como una herramienta para resolver problemas en entornos profesionales. Por otro lado, el lema *debates* apunta a la importancia de crear espacios de diálogo, donde estudiantes y docentes puedan discutir y construir conocimientos de manera colaborativa. En la misma línea, el lema *interactuar* alude tanto a la relación entre estudiantes y docentes como a su vínculo con la IAG.

Los datos sugieren que los entrevistados inciden en la comprensión del pensamiento crítico desde un enfoque centrado en las competencias transversales, la alfabetización crítica y los desafíos tecnológicos. La integración de la IAG y la tecnología en los entornos educativos emerge como un eje estratégico, pero requiere cambios metodológicos y una sólida formación para estudiantes y docentes. Además, el papel de la evaluación crítica, los debates y las soluciones aplicadas refuerzan la necesidad de fomentar entornos educativos dinámicos que preparen a los estudiantes para responder a los retos de la sociedad digital y el mercado laboral.

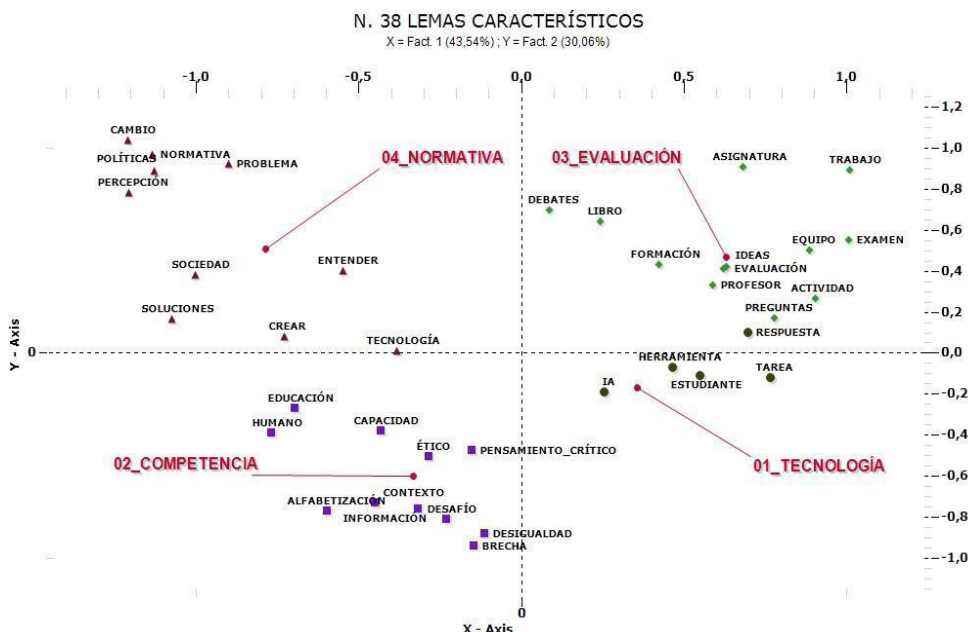
Análisis temáticos

A través del método de *clustering* no supervisado (algoritmo *bisecting k-promedio*) que ofrece *T-LAB*, se categorizó el contenido de las entrevistas en clústeres o grupos temáticos significativos, definidos según el patrón de lemas que los componen. El análisis temático del corpus lingüístico identificó 1.197 contextos elementales, de los cuales se clasificaron 1.006 (84,04 %). Se optó por una partición en cuatro *clústeres*, al presentar la mayor adherencia estadística para la investigación. Estas cuatro dimensiones clave, vinculadas al ámbito educativo y social, presentaron la siguiente distribución: *Clúster 01-Tecnología* (31,9 %); *Clúster 02-Competencias* (25,5 %); *Clúster 03-Evaluación* (19,9 %); y *Clúster 04-Normativa* (22,7 %).

La distribución de los clústeres y sus lemas asociados (Figura 3) facilita la identificación de tendencias en el tema de estudio. Estos cuatro *clústeres* ofrecen

una visión integrada del pensamiento crítico, estructurada en dos ejes: (1) *Eje X* - perspectiva operativa, relacionada con el desarrollo pedagógico de competencias en el aula; y (2) *Eje Y* - perspectiva estratégica, contextualizada en un marco normativo, social y político más amplio (ver Apéndice; Tabla 4).

Figura 3
Clústeres y lemas



Fuente: elaboración propia.

Análisis de contenidos

La transcripción de las entrevistas permitió la extracción de 801 ideas-fuerza, distribuidas según los bloques presentados en la Tabla 5 (ver Apéndice). Para garantizar la consistencia y validez, los investigadores codificaron individualmente estas ideas-fuerza asignándolas a las dimensiones del modelo pentagonal. El índice de Kappa de Fleiss ($\kappa=0,82$) confirma un alto grado de acuerdo entre los codificadores, respaldando la solidez del análisis (Altman, 1990). Por otra parte, las Tabla 6, 7, 8, 9 y 10 recogidas en el Apéndice reportan el detalle de consensos identificados entre los informantes entrevistados para cada una de las cinco dimensiones³. Las barreras, aceleradores, alternativas socio-tecnológicas, retos sociales y consecuencias visibles

incluidas en estas tablas se presentan de manera sintética, asegurando la fidelidad al sentido expresado por los informantes. A continuación, se discuten las ideas que cuentan con mayor consenso⁴ en cada dimensión y se conectan con el debate académico precedente.

Consensos sobre barreras para el desarrollo del pensamiento crítico

El análisis de las opiniones de los informantes respecto a las principales barreras que dificultan el desarrollo del pensamiento crítico de los estudiantes (Apéndice; Tabla 6) consensua un desfase entre el rápido avance de la IAG y la lenta adaptación de los sistemas educativos. Este problema, ampliamente documentado, ha sido señalado por Barrett y Pack (2023), quienes advierten que la ausencia de políticas claras y directrices institucionales genera incertidumbre y preocupaciones éticas, dificultando la integración de la IAG en las aulas. En respuesta, las universidades han comenzado a establecer directrices para el uso ético de la IAG, como el marco propuesto por Chan (2023). Asimismo, se han desarrollado orientaciones estratégicas para responsables de políticas públicas, entre las cuales destacan las contribuciones de Miao et al. (2021). La UNESCO (2024), en su *Marco de Competencia en Inteligencia Artificial para Docentes*, subraya que los sistemas educativos deben ir más allá de la enseñanza técnica (*enseñar sobre la IAG*) y fomentar una comprensión crítica (*enseñar para la IAG*).

Actualmente, la formación en IAG está liderada principalmente por empresas privadas que priorizan el desarrollo de habilidades técnicas. Por ello, resulta fundamental aumentar la conciencia sobre la necesidad de integrar el pensamiento crítico en los contextos educativos donde se emplea la IA. Esta idea también fue respaldada por los informantes, quienes abogan por prácticas de aprendizaje centradas en el proceso, con enfoques que fomenten la reflexión, la exploración y el pensamiento crítico.

La resistencia al cambio entre actores educativos y en todos los niveles institucionales refleja un consenso relevante. La evidencia sugiere que esta barrera sistémica responde a inercias organizacionales que obstaculizan la implementación de innovaciones pedagógicas y tecnológicas necesarias para fomentar el pensamiento crítico en las enseñanzas a distancia. La voluntad y disposición de los docentes para promover un uso crítico de la IAG constituye una condición indispensable para su integración en el ámbito educativo. En este sentido, comprender sus actitudes, creencias y preconcepciones se erige como una palanca fundamental para garantizar una integración efectiva (Choi et al., 2023).

La insuficiente formación docente en pensamiento crítico y en el uso de IAG limita la capacidad de los educadores para diseñar estrategias pedagógicas que integren ambas dimensiones en el proceso de enseñanza-aprendizaje. En esta línea, diversos estudios destacan la importancia de fomentar el pensamiento crítico en los programas de formación inicial del profesorado (Lorencová et al., 2019; Mpofu y

Maphalala, 2017; Ronderos et al., 2024) y de emplear estrategias metacognitivas que fortalezcan las competencias digitales en el aula (Pereles et al., 2024).

El uso acrítico de la IAG y el riesgo latente de plagio por parte de los estudiantes constituye otro consenso entre los expertos entrevistados. Esta problemática subraya la necesidad de desarrollar competencias para un uso ético y reflexivo de estas herramientas (Cotton et al., 2023). La falta de evaluación y priorización de habilidades transversales como el pensamiento crítico, junto con la ausencia de métricas claras, dificulta su seguimiento y desarrollo en los estudiantes. Este desafío se intensifica en la educación a distancia, particularmente en actividades asincrónicas que emplean ensayos como herramienta de evaluación. Según Eke (2023), la integración de estas herramientas genera preocupaciones sobre la integridad académica y los límites de coautoría, comprometiendo la capacidad de los ensayos para reflejar fielmente el proceso cognitivo del estudiante y evaluar su razonamiento crítico de manera efectiva. Para mitigar este riesgo, esta investigación propone la difusión de los marcos presentados en las Tablas 1 y 2, que ofrecen herramientas para analizar críticamente la información generada por la IAG y fortalecer el pensamiento crítico en los estudiantes.

La carencia de alfabetización específica en IAG y en pensamiento crítico limita la capacidad de docentes y estudiantes para aprovechar el potencial educativo de estas herramientas. La literatura académica aborda la necesidad de alfabetización desde diversas áreas formativas: el diseño curricular (Chiu y Chai, 2020), los marcos de alfabetización en IAG (Luckin et al., 2022), las aplicaciones didácticas (Wilton et al., 2022), los programas de desarrollo profesional (Vazhayil et al., 2019) y las consideraciones éticas (Celik, 2023; Gartner y Krašna, 2023).

Un consenso ampliamente reconocido es la percepción, compartida por docentes y autoridades universitarias, de la IAG como un sustituto educativo en lugar de un complemento. Esta visión puede obstaculizar su integración como herramienta de apoyo para el desarrollo del pensamiento crítico. En este sentido, la literatura enfatiza la importancia de crear entornos de aprendizaje equilibrados que prioricen el razonamiento analítico antes de recurrir a la IAG (Malik et al., 2023). En todo caso, dada la rápida evolución de esta tecnología, estrategias como su prohibición o la inversión en métodos de detección del plagio resultan insostenibles (Martín y López, 2023). Por ello, el principal desafío es lograr su integración ética y efectiva en los procesos educativos.

Consensos sobre aceleradores para el desarrollo del pensamiento crítico

Con respecto a los aceleradores (Apéndice; Tabla 7), la actualización de competencias transversales y específicas para el uso crítico de la IAG se presenta como el consenso más amplio entre los entrevistados, con siete coincidencias. Para abordar estas necesidades formativas, algunas investigaciones (Kong et al., 2023) proponen diseñar un curso introductorio de alfabetización en IAG que fortalezca la

comprensión tecnológica y la conciencia ética de los participantes, preparándolos para aplicar y evaluar la IAG de manera crítica en sus futuras carreras.

La evidencia hallada indica la necesidad de reformar los sistemas de evaluación existentes, integrando la IAG para optimizar el *feedback* y la adaptación curricular. La alineación de las políticas educativas con los avances tecnológicos, desde un enfoque ético y responsable, podría actuar como un acelerador para fortalecer el pensamiento crítico. En este contexto, los expertos consultados resaltan la formación y motivación del profesorado (Ayanwale et al., 2022) y una adecuada coordinación de las políticas públicas como factores claves para superar las barreras identificadas.

Otros puntos de consenso entre los informantes incluyen la accesibilidad a la IAG para reducir brechas y el fortalecimiento de sistemas de autoevaluación y evaluación por pares. Estas prácticas constituyen actividades fundamentales para fomentar la metacognición y la reflexión en el aprendizaje, permitiendo que los estudiantes analicen su desempeño y ajusten sus procesos. La integración de estas evaluaciones con herramientas de IAG amplía su alcance al detectar sesgos, inexactitudes y áreas de mejora de manera objetiva (Guàrdia Ortiz et al., 2024). Este enfoque se vincula con el concepto de *learning by comparing* (Longarela-Ares y Rodríguez-Padín, 2023), al estimular el aprendizaje mediante la comparación entre las evaluaciones de la IA, los pares y el propio estudiante. De este modo, se estimula el pensamiento crítico y se equilibra el uso de la tecnología con la intervención humana.

Otro consenso señala que la IAG optimiza tareas automatizables, liberando tiempo que puede ser redirigido hacia actividades más complejas, como el desarrollo del pensamiento crítico. Esta idea está estrechamente asociada al potencial de la IA para transformar el proceso de aprendizaje y trabajo mediante la retroalimentación inmediata (Cavalcanti et al., 2021). Asimismo, se reconoce la necesidad de implementar políticas específicas que refuercen la conexión entre la universidad y el mercado laboral, garantizando que la formación de los futuros egresados responda a las demandas profesionales. No obstante, surge una dicotomía: por un lado, la importancia de que los sistemas educativos trasciendan la enseñanza técnica de la IAG y fomenten una comprensión crítica; y, por otro, la urgencia de adaptarse a un entorno laboral en constante evolución que exige el conocimiento de aplicaciones de IAG especializadas en distintos campos profesionales.

Consensos sobre alternativas socio-tecnológicas para el desarrollo del pensamiento crítico

Entre las alternativas socio-tecnológicas disponibles (Apéndice; Tabla 8), los informantes destacan la necesidad de garantizar un acceso equitativo a las tecnologías de IAG. Advierten que, si no se gestionan adecuadamente, estas herramientas podrían ampliar la brecha digital y productiva entre quienes poseen habilidades para formular *prompts* efectivos y quienes carecen de ellas, subrayando el papel activo que la universidad debe asumir en la democratización de su acceso.

Los informantes resaltan la importancia de capacitar a toda la comunidad universitaria en el uso crítico de la IA, fomentando habilidades para evaluar e interpretar reflexivamente sus resultados. Para ello, consideran fundamental comprender sus límites, sesgos algorítmicos y alucinaciones latentes (Baker y Hawn, 2022). En este contexto, los informantes proponen implementar *chatbots* en entornos seguros y controlados para asegurar un uso ético y responsable. Sin embargo, advierten que restringir estas herramientas exclusivamente a preguntas académicas podría generar sesgos algorítmicos, limitando tanto la diversidad como el alcance de las respuestas, en contraste con las posibilidades más amplias y flexibles de otras aplicaciones de IAG más versátiles.

Por último, los expertos proponen incorporar una dimensión ética que preserve la integridad académica y fomente la capacidad autorreflexiva del alumnado. Si bien esta línea de actuación es decisiva en la educación superior en general, adquiere especial relevancia en la enseñanza a distancia, donde el aprendizaje autónomo y las limitaciones en la supervisión docente dificultan el control del uso de la IAG por parte del estudiantado (Dilekli y Boyraz, 2024).

Consensos sobre retos sociales para el desarrollo del pensamiento crítico

Los informantes destacaron los principales retos sociales que deben abordarse para fortalecer el pensamiento crítico en la educación universitaria a distancia en un contexto marcado por la generalización del uso de la IAG (Apéndice; Tabla 9). Existe un amplio consenso respecto a la necesidad de transformar los roles docentes actuales para mitigar la dependencia acrítica de la IAG y fomentar la autonomía de pensamiento.

Los expertos identifican como principales retos sociales la alfabetización digital y la reducción de brechas tecnológicas, enfatizando la necesidad de desarrollar competencias digitales desde etapas formativas previas al acceso a la universidad. En este contexto, Lin y Van Brummelen (2021) apuntan que los docentes de educación primaria y secundaria requieren un andamiaje adicional en el uso de herramientas de IA y en el diseño curricular, con el fin de facilitar debates sobre ética y datos, fortalecer la evaluación, promover la participación estudiantil, fomentar la colaboración entre pares y estimular la reflexión crítica y el cuestionamiento de la información generada por la IAG. Este desafío vuelve a poner de relieve el imperativo de garantizar un acceso equitativo a las tecnologías emergentes, una prioridad ya contemplada en la *Estrategia de la UNESCO (2021) sobre la Innovación Tecnológica en la Educación*. Este marco recalca la importancia de adoptar un enfoque centrado en el ser humano, en el que la IA contribuya a reducir desigualdades en el acceso al conocimiento, la investigación y la cultura, evitando la ampliación de brechas tecnológicas y asegurando que sus beneficios sean accesibles para todos.

Consenso sobre consecuencias visibles del desarrollo del pensamiento crítico

En opinión de los expertos consultados (Apéndice; Tabla 10), desde una perspectiva pedagógica, desarrollar la capacidad de cuestionar los resultados de la IAG permitiría transformar la evaluación actual -centrada en los resultados- en una nueva dimensión evaluativa -centrada en los procesos- que dieran lugar al desarrollo de nuevas competencias analíticas y creativas.

Según los consensos detectados, los informantes destacan que los riesgos asociados al uso de la IAG podrían mitigarse mediante la alfabetización digital y la transparencia algorítmica. Asimismo, señalan que una de las principales consecuencias del fortalecimiento del pensamiento crítico en el alumnado sería que, al estar apoyado por la IAG, mejoraría la velocidad del proceso de enseñanza-aprendizaje. Sin embargo, alertan que, dependiendo de cómo se implementen las políticas educativas, podrían surgir riesgos significativos de ampliación de las desigualdades sociales y educativas preexistentes. Por último, los entrevistados coinciden en que el desarrollo efectivo del pensamiento crítico fortalece las capacidades de liderazgo y resolución de problemas en los futuros egresados, favoreciendo perfiles más adaptables y reduciendo el riesgo de homogeneización del pensamiento.

Implicaciones

Los hallazgos de esta investigación coinciden con estudios previos sobre la IAG (Castelló-Sirvent y Cortés-Pellicer, 2024), que destacan las preocupaciones de los futuros egresados sobre el potencial de error, la calidad y la imparcialidad de la información, la manipulación de contenido sesgado o falso y la necesidad de formación en IA para mejorar la empleabilidad. El estudio de Rusdin et al. (2023) subraya la percepción que los estudiantes tienen de la IA como una herramienta valiosa para el pensamiento crítico, especialmente en la investigación académica y el análisis teórico. No obstante, advierten sobre riesgos como la falta de personalización, la generación de cámaras de eco y las dificultades en la comprensión matizada del conocimiento.

Además, este trabajo contribuye a orientar el debate académico en torno a las principales implicaciones que se proponen en la Tabla 12. Los cinco vectores de actuación sirven como guía estratégica para docentes, gestores universitarios y responsables del diseño de políticas educativas, facilitando la toma de decisiones en la integración de la IA en los procesos de enseñanza y aprendizaje.

Tabla 12

Vectores de actuación propuestos para el desarrollo del pensamiento crítico en contextos IAG

Estrategia	Acciones
Desarrollar un marco institucional para la integración ética de IAG	• Definir directrices claras para el uso ético de IAG. • Promover la IAG como herramienta complementaria en educación. • Reducir riesgos de plagio con criterios responsables. • Crear comités especializados para supervisar la integración. • Garantizar coherencia supra institucional en las políticas de IAG. • Priorizar la formación docente en el marco institucional. • Establecer estándares para evaluar pensamiento crítico. • Asegurar equidad en el acceso a tecnologías de IAG.
Impulsar programas de capacitación docente en IAG y en pensamiento crítico	• Diseñar capacitaciones en pensamiento crítico e IAG. • Incluir aspectos técnicos, éticos y pedagógicos en los programas. • Crear espacios de mentoría y colaboración docente. • Extender buenas prácticas en enseñanza con IAG. • Incorporar estrategias basadas en evidencia en las capacitaciones. • Promover metodologías reflexivas para integrar pensamiento crítico. • Facilitar la transición hacia pedagogías apoyadas en IAG. • Asegurar la actualización constante de competencias docentes.
Diseñar e implementar programas de alfabetización digital específica en IAG	• Crear materiales accesibles sobre límites y potencial de IAG. • Fomentar la reflexión crítica en el uso de herramientas digitales. • Establecer módulos obligatorios sobre uso ético de IAG. • Adaptar programas a necesidades específicas de estudiantes y docentes. • Prover formación práctica y ética en uso de IAG. • Diseñar recursos para validar información y evitar sesgos. • Enfatizar la autorregulación en el aprendizaje autónomo. • Atender particularidades de educación a distancia sin docencia síncrona.
Actualización de sistemas de evaluación centrados en procesos analíticos	• Priorizar evaluación de razonamiento sobre resultados finales. • Diseñar rúbricas específicas para análisis crítico con IAG. • Promover el uso reflexivo de IAG en procesos evaluativos. • Integrar habilidades metacognitivas en sistemas de evaluación. • Implementar evaluaciones que valoren argumentos desarrollados con IAG. • Alinear evaluación con competencias críticas y analíticas. • Incentivar la colaboración vigilante entre estudiantes y docentes. • Facilitar retroalimentación centrada en procesos reflexivos.
Rediseñar los planes de estudio universitarios	• Transformar currículos para incluir habilidades críticas con IAG. • Coordinar rediseño con niveles meso y macro institucionales. • Garantizar acceso equitativo a licencias y recursos de IAG. • Promover igualdad de oportunidades en desarrollo crítico. • Incluir módulos sobre pensamiento crítico y sesgos algorítmicos. • Fortalecer la atención a diversidad y necesidades especiales. • Asegurar alineación con demandas del mercado laboral. • Incorporar componentes críticos y reflexivos en planes de estudio.

Fuente: elaboración propia.

CONCLUSIONES

La investigación sobre el desarrollo del pensamiento crítico en contextos de uso generalizado de IAG sigue siendo limitada y fragmentaria. Este estudio exploratorio constituye un primer esfuerzo por identificar los factores que inciden en su desarrollo, analizando barreras, aceleradores, alternativas tecnológicas, retos sociales y consecuencias de su integración. Los consensos alcanzados y los vectores de actuación propuestos pueden servir como base para el diseño de instrumentos de medición, como encuestas dirigidas a docentes universitarios. Contar con una muestra suficientemente amplia y diversificada permitirá contrastar las percepciones expertas con datos empíricos, fortaleciendo la validez externa y fiabilidad de los hallazgos. Esta triangulación permitirá evaluar la coherencia entre tendencias cualitativas y mediciones cuantitativas, identificando posibles discrepancias o convergencias en la relación entre IAG y pensamiento crítico, y proporcionando un marco más sólido para la toma de decisiones en el diseño de políticas pedagógicas y estrategias de formación docente.

Aunque este estudio se ha desarrollado en universidades a distancia dentro del Espacio Europeo de Educación Superior (EEES), lo que podría restringir la generalización de los resultados a otros sistemas con normativas y estructuras organizativas distintas, los consensos alcanzados y las líneas de actuación propuestas muestran una alta capacidad de adaptación a distintos entornos educativos digitales, favoreciendo su aplicabilidad en contextos con dinámicas de aprendizaje similares.

Si bien esta investigación ha centrado el análisis en el rol del profesorado en la implementación de la IAG, futuros trabajos deberían ampliar el espectro de actores involucrados, incluyendo la perspectiva de estudiantes, responsables de políticas educativas y representantes del sector productivo. Incorporar estos perfiles permitiría un enfoque más holístico y multidimensional, facilitando una comprensión más completa de los desafíos y oportunidades que la IAG plantea en el desarrollo del pensamiento crítico. Esto enriquecería el debate académico y fortalecería el diseño de estrategias pedagógicas más contextualizadas y ajustadas a las necesidades del ecosistema educativo.

La IAG puede hacer el aprendizaje más inmersivo, dinámico y personalizado, mejorando el rendimiento académico, la motivación y la autorregulación (Huang et al., 2023; Yuan y Liu, 2025). No obstante, sin una mediación adecuada, su uso podría generar una dependencia acrítica, llevando a los estudiantes a aceptar respuestas automatizadas sin analizarlas reflexivamente, lo que limitaría su capacidad de cuestionamiento (Chng et al., 2023). Para mitigar este riesgo, es fundamental equilibrar la integración tecnológica con la intervención docente, asegurando que la IAG no reemplace el proceso cognitivo del estudiante, sino que lo potencie a través de estrategias estructuradas y reflexivas.

El *engagement* en educación a distancia no solo depende del acceso a herramientas innovadoras, sino también de la calidad de la interacción docentes-estudiantes, un

factor determinante para la retención del conocimiento y la satisfacción académica (Bae et al., 2020; Hoi y Le Hang, 2021). Por ello, la planificación pedagógica debe integrar la IAG sin desplazar la labor docente, alineándose con objetivos que promuevan la validación de información, la identificación de sesgos algorítmicos y la reflexión sobre el impacto de la IA en el aprendizaje (Martín y López, 2023).

Para una integración efectiva, los docentes requieren de respaldo institucional. Las instituciones deben establecer directrices que regulen su uso asegurando una implementación ética, accesible y alineada con estándares pedagógicos que estimulen el pensamiento crítico. Asimismo, la formación docente en alfabetización digital e IA resulta imprescindible para diseñar estrategias de evaluación que fomenten un aprendizaje auténtico. En entornos de aprendizaje a distancia, donde el ensayo ha sido la principal herramienta de evaluación, contrastar respuestas generadas por IA con fuentes verificadas, diseñar *prompts* estratégicos y analizar críticamente los *outputs* algorítmicos pueden fomentar una interacción más reflexiva con la tecnología.

Futuros estudios empíricos deberían analizar cómo las intervenciones educativas con herramientas de IA impactan en el pensamiento crítico y en las dimensiones afectiva y conductual del *engagement* estudiantil. Para ello, resultaría útil combinar instrumentos de evaluación de habilidades cognitivas con autoinformes, análisis de interacciones y métricas cuantitativas extraídas de plataformas virtuales. Además, estudios longitudinales podrían examinar su efecto a medio plazo, identificando su influencia en la reducción del abandono, la mitigación de la desconexión y el aislamiento académico, así como su contribución al desarrollo de la autonomía, la autorregulación del aprendizaje y la motivación estudiantil.

Agradecimientos

Este artículo es resultado de la producción académica derivada del Proyecto de Innovación Docente de la Universidad Nacional de Educación a Distancia (UNED) “La mejora del Engagement y el Rendimiento Académico mediante el Uso de IA en las Pruebas de Evaluación Continua” (MERA-IAPEC); del Proyecto de Innovación Docente UV-SFPIE_PIEE 3325253 “Inteligencia Artificial en Comunicación (IAcom)”, concedido por el Vicerectorat de Formació Permanent, Transformació Docent i Ocupació de la Universitat de València; del Proyecto de Innovación y Mejora Educativa PIME/23-24/390 de la Universitat Politècnica de València “La influencia de las metodologías activas en la interacción y en el desarrollo de habilidades metacognitivas”; y del Grupo de Investigación Emergente “Artificial Intelligence Impacts on Society & Economy (AIISE)”. Diferentes fases del trabajo fueron desarrolladas durante la estancia de investigación de César Muñoz Martínez en el Departamento de Organización de Empresas de la Universitat Politècnica de València, coincidiendo con la etapa de revisión final del manuscrito.

NOTAS

1. Algunos autores (Van Laar et al., 2017; Thornhill-Miller et al., 2023) las describen como *21st Century Skills* o las “4 C” (creatividad, pensamiento crítico, comunicación y colaboración).
2. En el debate académico se emplea el término IAG, si bien una práctica totalidad de los informantes utilizó el término IA en un sentido coloquial que hacía referencia a la IAG. Como consecuencia el análisis textual y de contenido emplea ambos términos de manera indistinta en la expresión literal o en la discusión de los resultados, respectivamente.
3. La Tabla 11, recogida en el Apéndice, permite consultar un ejemplo de expresiones literales de los informantes que se vinculan con las ideas-fuerza de cada dimensión.
4. Si bien las Tablas 6 a 10, recogidas en el Apéndice, incluyen consensos iguales o superiores a dos informantes, la siguiente sección de consensos describe los acuerdos iguales o superiores a tres informantes.

REFERENCIAS

- Adiguzel, T., Kaya, M. H. y Cansu, F. K. (2023). Revolutionizing education with AI: Exploring the transformative potential of ChatGPT. *Contemporary Educational Technology*, 15(3), Article ep429. <https://doi.org/10.30935/cedtech/13152>
- Aguilar Vargas, L. R., Alcántara Llanas, I. T. y Braun Mondragón, K. A. (2020). Impacto del pensamiento crítico en las habilidades para el campo laboral. *ACADEMO*, 7(2), 166-174. <https://doi.org/10.30545/academo.2020.jul-dic.7>
- Almanasreh, E., Moles, R. y Chen, T. F. (2019). Evaluation of methods used for estimating content validity. *Research in Social and Administrative Pharmacy*, 15(2), 214-221. <https://doi.org/10.1016/j.sapharm.2018.03.066>
- Altman, D. G. (1990). *Practical statistics for medical research*. Chapman and Hall/CRC. <https://doi.org/10.1201/9780429258589>
- Ayanwale, M. A., Sanusi, I. T., Adelana, O. P., Aruleba, K. D. y Oyelerere, S. S. (2022). Teachers' readiness and intention to teach artificial intelligence in schools. *Computers and Education: Artificial Intelligence*, 3, Article 100099. <https://doi.org/10.1016/j.caeai.2022.100099>
- Ayyıldız, P. y Yılmaz, A. (2021). Moving the kaleidoscope to see the effect of creative personality traits on creative thinking dispositions of preservice teachers: The mediating effect of creative learning environments and teachers' creativity fostering behavior. *Thinking Skills and Creativity*, 41, Article 100879. <https://doi.org/10.1016/j.tsc.2021.100879>
- Bae, C. L., DeBusk-Lane, M. L. y Lester, A. M. (2020). Engagement profiles of elementary students in urban schools. *Contemporary Educational Psychology*, 62, Article 101880. <https://doi.org/10.1016/j.cedpsych.2020.101880>
- Baidoo-Anu, D. y Owusu Ansah, L. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI*, 7(1), 52-62. <https://doi.org/10.61969/jai.1337500>
- Baker, R. S. y Hawn, A. (2022). Algorithmic bias in education. *International Journal of Artificial Intelligence in Education*, 32, 1052-1092. <https://doi.org/10.1007/s40593-021-00285-9>
- Barrett, A. y Pack, A. (2023). Not quite eye to A.I.: Student and teacher perspectives on

- the use of generative artificial intelligence in the writing process. *International Journal of Educational Technology in Higher Education*, 20(1), Article 59. <https://doi.org/10.1186/s41239-023-00427-0>
- Barrot, J. S. (2023). Using ChatGPT for second language writing: Pitfalls and potentials. *Assessing Writing*, 57, Article 100745. <https://doi.org/10.1016/j.asw.2023.100745>
- Bhutoria, A. (2022). Personalized education and artificial intelligence in the United States, China, and India: A systematic review using a Human-In-The-Loop model. *Computers and Education: Artificial Intelligence*, 3, Article 100068. <https://doi.org/10.1016/j.caeai.2022.100068>
- Bloom, B. S., Englehart, M. D., Furst, E. J., Hill, W. H. y Krathwohl, D. R. (Eds.). (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive domain*. Longmans, Green.
- Çakıcı, D. (2018). Metacognitive awareness and critical thinking abilities of pre-service EFL teachers. *Journal of Education and Learning*, 7(5), 116-129. <https://doi.org/10.5539/jel.v7n5p116>
- Cananau, I., Edling, S. y Haglund, B. (2025). Critical thinking in preparation for student teachers' professional practice: A case study of critical thinking conceptions in policy documents framing teaching placement at a Swedish university. *Teaching and Teacher Education*, 153, Article 104816. <https://doi.org/10.1016/j.tate.2024.104816>
- Castelló-Sirvent, F. y Cortés-Pellicer, P. (2024). Using generative AI for strategic analysis? A study on perceived utility among industrial organization engineering students. En *INTED2024 Proceedings* (pp. 3365-3372). IATED. <https://doi.org/10.21125/inted.2024.0894>
- Cavalcanti, A. P., Barbosa, A., Carvalho, R., Freitas, F., Tsai, Y. S., Gašević, D. y Mello, R. F. (2021). Automatic feedback in online learning environments: A systematic literature review. *Computers and Education: Artificial Intelligence*, 2, Article 100027. <https://doi.org/10.1016/j.caeai.2021.100027>
- Celik, I. (2023). Towards Intelligent-TPACK: An empirical study on teachers' professional knowledge to ethically integrate artificial intelligence (AI)-based tools into education. *Computers in Human Behavior*, 138, Article 107468. <https://doi.org/10.1016/j.chb.2022.107468>
- Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. *International Journal of Educational Technology in Higher Education*, 20(1), Article 38. <https://doi.org/10.1186/s41239-023-00408-3>
- Chassignol, M., Khoroshavin, A., Klimova, A. y Bilyatdinova, A. (2018). Artificial intelligence trends in education: A narrative overview. *Procedia Computer Science*, 136, 16-24. <https://doi.org/10.1016/j.procs.2018.08.233>
- Chen, L., Chen, P. y Lin, Z. (2020). Artificial intelligence in education: A review. *IEEE Access*, 8, 75264-75278. <https://doi.org/10.1109/ACCESS.2020.2988510>
- Chiu, T. K. y Chai, C. S. (2020). Sustainable curriculum planning for artificial intelligence education: A self-determination theory perspective. *Sustainability*, 12(14), Article 5568. <https://doi.org/10.3390/su12145568>
- Chng, E., Tan, A. L. y Tan, S. C. (2023). Examining the use of emerging technologies in schools: A review of artificial intelligence and immersive technologies in STEM education. *Journal for STEM Education Research*, 6(3), 385-407. <https://doi.org/10.1007/s41979-023-00092-y>

- Choi, S., Jang, Y. y Kim, H. (2023). Influence of pedagogical beliefs and perceived trust on teachers' acceptance of educational artificial intelligence tools. *International Journal of Human-Computer Interaction*, 39(4), 910-922. <https://doi.org/10.1080/10447318.2022.2049145>
- Choy, S. C. y Cheah, P. K. (2009). Teacher perceptions of critical thinking among students and its influence on higher education. *International Journal of Teaching and Learning in Higher Education*, 20(2), 198-206. <https://files.eric.ed.gov/fulltext/EJ864337.pdf>
- Cortini, M. y Tria, S. (2014). Triangulating qualitative and quantitative approaches for the analysis of textual materials: An introduction to T-Lab. *Social Science Computer Review*, 32(4), 561-568. <https://doi.org/10.1177/0894439313510108>
- Cotton, D. R. E., Cotton, P. A. y Shipway, J. R. (2023). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*, 61(2), 228-239. <https://doi.org/10.1080/14703297.2023.2190148>
- De Vicente, J. y Matti, C. (2016). *Visual toolbox for system innovation. A resource book for practitioners to map, analyse and facilitate sustainability transitions*. Transitions Hub Series. Climate-KIC.
- Delgado, A. J., Wardlow, L., McKnight, K. y O'Malley, K. (2015). Educational technology: A review of the integration, resources, and effectiveness of technology in K-12 classrooms. *Journal of Information Technology Education: Research*, 14, 397-416. <https://doi.org/10.28945/2298>
- Deng, J. y Lin, Y. (2022). The benefits and challenges of ChatGPT: An overview. *Frontiers in Computing and Intelligent Systems*, 2(2), 81-83. <https://doi.org/10.54097/fcis.v2i2.4465>
- Dilekli, Y. y Boyraz, S. (2024). From "Can AI think?" to "Can AI help thinking deeper?": Is use of ChatGPT in higher education a tool of transformation or fraud? *International Journal of Modern Education Studies*, 8(1), 49-71. <https://doi.org/10.51383/ijonmes.2024.316>
- Drugova, E., Zhuravleva, I., Zakharova, U. y Latipov, A. (2024). Learning analytics driven improvements in learning design in higher education: A systematic literature review. *Journal of Computer Assisted Learning*, 40(2), 510-524. <https://doi.org/10.1111/jcal.12894>
- Eke, O. D. (2023). ChatGPT and the rise of generative AI: Threat to academic integrity? *Journal of Responsible Technology*, 13, Article 100060. <https://doi.org/10.1016/j.jrt.2023.100060>
- Ennis, R. H. y Millman, J. (1985). *Cornell critical thinking test, level Z*. Midwest Publications.
- Ertmer, P. A., Sadaf, A. y Ertmer, D. J. (2011). Student-content interactions in online courses: The role of question prompts in facilitating higher-level engagement with course content. *Journal of Computing in Higher Education*, 23, 157-186. <https://doi.org/10.1007/s12528-011-9047-6>
- Facione, P. A. (1990). *Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction. Research findings and recommendations*. ERIC Clearinghouse on Higher Education. <https://eric.ed.gov/?id=ED315423>
- Facione, P. A. (2023). *Critical thinking: What it is and why it counts* (2023 update). Insight Assessment. <https://insightassessment.com/wp-content/uploads/2023/12/Critical-Thinking-What-It-Is-and-Why-It-Counts.pdf>
- Flavell, J. H. (1976). Metacognitive aspects of problem solving. En L. B. Resnick (Ed.), *The nature of intelligence* (pp. 231-235). Lawrence Erlbaum Associates. <https://doi.org/10.4324/9781032646527-16>
- Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area

- of cognitive-developmental inquiry. *American Psychologist*, 34(10), 906-911. <https://doi.org/10.1037/0003-066X.34.10.906>
- Fleiss, J. L. (1981). *Statistical methods for rates and proportions* (2nd ed.). John Wiley & Sons.
- Fuchs, K. (2023). Exploring the opportunities and challenges of NLP models in higher education: Is ChatGPT a blessing or a curse? *Frontiers in Education*, 8, Article 1166682. <https://doi.org/10.3389/feduc.2023.1166682>
- Garcés, G. y Bastías, E. (2025). Competencies model for online learning in higher education: A bibliometric analysis and systematic review. *RIED-Revista Iberoamericana de Educación a Distancia*, 28(1), 37-66. <https://doi.org/10.5944/ried.28.1.41351>
- Gartner, S. y Krašna, M. (2023). Artificial intelligence in education-ethical framework. En *2023 12th Mediterranean Conference on Embedded Computing (MECO)* (pp. 1-7). IEEE. <https://doi.org/10.1109/MECO58584.2023.10155012>
- Gil, H. y Guilleumas, R. M. (2017). Redes de comunicación del movimiento 15M en Twitter. *Redes. Revista Hispana para el Análisis de Redes Sociales*, 28(1), 135-146. <https://doi.org/10.5565/rev/redes.670>
- Goodsett, M. (2020). Best practices for teaching and assessing critical thinking in information literacy online learning objects. *The Journal of Academic Librarianship*, 46(5), Article 102163. <https://doi.org/10.1016/j.acalib.2020.102163>
- Guàrdia Ortiz, L., Maina, M., Cabrera Lanzo, N. y Fernández-Ferrer, M. (2024). La autorregulación del aprendizaje desde un enfoque de feedback entre pares: Perspectivas de la IA generativa. *Revista de Educación a Distancia (RED)*, 24(78). <https://doi.org/10.6018/red.599511>
- Guo, B., Zhang, X., Wang, Z., Jiang, M., Nie, J., Ding, Y., Yue, J. y Wu, Y. (2023). *How close is ChatGPT to human experts? Comparison corpus, evaluation, and detection.* arXiv. <https://doi.org/10.48550/arXiv.2301.07597>
- Guo, Y. y Lee, D. (2023). Leveraging ChatGPT for enhancing critical thinking skills. *Journal of Chemical Education*, 100(12), 4876-4883. <https://doi.org/10.1021/acs.jchemed.3c00505>
- Halpern, D. F. (1998). Teaching critical thinking for transfer across domains: Disposition, skills, structure training, and metacognitive monitoring. *American Psychologist*, 53(4), 449-455. <https://doi.org/10.1037/0003-066X.53.4.449>
- Hoi, V. N. y Le Hang, H. (2021). The structure of student engagement in online learning: A bi-factor exploratory structural equation modelling approach. *Journal of Computer Assisted Learning*, 37(4), 1141-1153. <https://doi.org/10.1111/jcal.12551>
- Huang, A. Y. Q., Lu, O. H. T. y Yang, S. J. H. (2023). Effects of artificial intelligence-enabled personalized recommendations on learners' learning engagement, motivation, and outcomes in a flipped classroom. *Computers and Education*, 194, Article 104684. <https://doi.org/10.1016/j.compedu.2022.104684>
- Ilić, M., Mikić, V., Kopanja, L. y Vesin, B. (2023). Intelligent techniques in e-learning: A literature review. *Artificial Intelligence Review*, 56(12), 14907-14953. <https://doi.org/10.1007/s10462-023-10508-1>
- Ivanov, S. (2023). The dark side of artificial intelligence in higher education. *Service Industries Journal*, 43(15-16), 1055-1082. <https://doi.org/10.1080/02642069.2023.2258799>
- Javaid, M., Haleem, A., Singh, R. P., Khan, S. y Khan, I. H. (2023). Unlocking the opportunities through ChatGPT Tool towards ameliorating the education system. *BenchCouncil Transactions on Benchmarks, Standards and*

- Evaluations*, 3(2), Article 100115. <https://doi.org/10.1016/j.tbench.2023.100115>
- Kong, S. C., Cheung, W. M. Y. y Tsang, O. (2023). Evaluating an artificial intelligence literacy programme for empowering and developing concepts, literacy and ethical awareness in senior secondary students. *Education and Information Technologies*, 28(6), 4703-4724. <https://doi.org/10.1007/s10639-022-11408-7>
- Ku, K. Y. L. (2009). Assessing students' critical thinking performance: Urging for measurements using multi-response format. *Thinking Skills and Creativity*, 4(1), 70-76. <https://doi.org/10.1016/j.tsc.2009.02.001>
- Ku, K. Y. L. y Ho, I. T. (2010). Metacognitive strategies that enhance critical thinking. *Metacognition and Learning*, 5(3), 251-267. <https://doi.org/10.1007/s11409-010-9060-6>
- Kuhn, D. y Dean, D., Jr. (2004). Metacognition: A bridge between cognitive psychology and educational practice. *Theory Into Practice*, 43(4), 268-274. https://doi.org/10.1207/s15430421tip4304_4
- Kumar, R., Eaton, S. E., Mindzak, M. y Morrison, R. (2024). Academic integrity and artificial intelligence: An overview. En S. E. Eaton (Ed.), *Second handbook of academic integrity* (pp. 1583-1596). Springer. https://doi.org/10.1007/978-3-031-54144-5_153
- Lin, P. y Van Brummelen, J. (2021). Engaging teachers to co-design integrated AI curriculum for K-12 classrooms. En *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Article 239, pp. 1-12). Association for Computing Machinery. <https://doi.org/10.1145/3411764.3445377>
- Lipman, M. (1988). Critical thinking—What can it be? *Educational Leadership*, 46(1), 38-43.
- Liu, Y., Han, T., Ma, S., Zhang, J., Yang, Y., Tian, J., He, H., Li, A., He, M., Liu, Z., Wu, Z., Zhao, L., Zhu, D., Li, X., Qiang, N., Shen, D., Liu, T. y Ge, B. (2023). Summary of ChatGPT-related research and perspective towards the future of large language models. *Meta-Radiology*, 1(2), Article 100017. <https://doi.org/10.1016/j.metrad.2023.100017>
- Longarela-Ares, A. M. y Rodríguez-Padín, R. (2023). Aprendizaje colaborativo, learning-by-doing y evaluación entre pares en educación superior. *EDUCA. Revista Internacional para la Calidad Educativa*, 3(2), 275-298. <https://doi.org/10.55040/educa.v3i2.66>
- Lorencová, H., Jarošová, E., Avgitidou, S. y Dimitriadou, C. (2019). Critical thinking practices in teacher education programmes: A systematic review. *Studies in Higher Education*, 44(5), 844-859. <https://doi.org/10.1080/03075079.2019.1586331>
- Luckin, R., Cukurova, M., Kent, C. y Du Boulay, B. (2022). Empowering educators to be AI-ready. *Computers and Education: Artificial Intelligence*, 3, Article 100076. <https://doi.org/10.1016/j.caeai.2022.100076>
- MacKnight, C. B. (2000). Teaching critical thinking through online discussions. *Educause Quarterly*, 23(4), 38-41.
- Malik, A. R., Pratiwi, Y., Andajani, K., Numertayasa, I. W., Suharti, S. y Darwis, A. (2023). Exploring artificial intelligence in academic essay: Higher education student's perspective. *International Journal of Educational Research Open*, 5, Article 100296. <https://doi.org/10.1016/j.ijedro.2023.100296>
- Maor, R., Paz-Baruch, N., Grinshpan, N., Milman, A., Mevarech, Z., Levi, R. y Zion, M. (2023). Relationships between metacognition, creativity, and critical thinking in self-reported teaching performances in project-based learning settings. *Thinking Skills and Creativity*, 50, Article 101425. <https://doi.org/10.1016/j.tsc.2023.101425>

- Martín, S. y López, E. (2023). *Guía para integrar las tecnologías basadas en inteligencia artificial generativa en los procesos de enseñanza y aprendizaje*. Vicerrectorado de Innovación Educativa, UNED.
- Mazzoni, D., Marchetti, L., Albanesi, C. y Cicognani, E. (2018). L'uso di T-LAB in psicologia della salute. Una rassegna della letteratura [The use of T-LAB in health psychology. A literature review]. *Psicologia della Salute*, (2), 91-114. <https://doi.org/10.3280/PDS2018-002009>
- Miao, F., Holmes, W., Huang, R. y Zhang, H. (2021). *AI and education: A guidance for policymakers*. UNESCO Publishing. <https://unesdoc.unesco.org/ark:/48223/pf0000376709>
- Mondo, T. S. y Gândara, J. M. (2017). O turismo experiencial a partir de uma perspectiva socioeconômica mercadológica [Experiential tourism from a socioeconomic and market perspective]. *Journal of Tourism Analysis*, 24, 26-40.
- Mpofu, N. y Maphalala, M. C. (2017). Fostering critical thinking in initial teacher education curriculums: A comprehensive literature review. *Gender and Behaviour*, 15(2), 9226-9236. <https://hdl.handle.net/10520/EJC-b41be80a7>
- Ng, D. T. K., Leung, J. K. L., Su, J., Ng, R. C. W. y Chu, S. K. W. (2023). Teachers' AI digital competencies and twenty-first century skills in the post-pandemic world. *Educational Technology Research and Development*, 71(1), 137-161. <https://doi.org/10.1007/s11423-023-10203-6>
- Niu, L., Behar-Horenstein, L. S. y Garvan, C. W. (2013). Do instructional interventions influence college students' critical thinking skills? A meta-analysis. *Educational Research Review*, 9, 114-128. <https://doi.org/10.1016/j.edurev.2012.12.002>
- Ocumpaugh, J., Roscoe, R. D., Baker, R. S., Hutt, S. y Aguilar, S. J. (2024). Toward asset-based instruction and assessment in artificial intelligence in education. *International Journal of Artificial Intelligence in Education*, 34, 1559-1598. <https://doi.org/10.1007/s40593-023-00382-x>
- Papleontiou-Louca, E. (2003). The concept and instruction of metacognition. *Teacher Development*, 7(1), 9-30. <https://doi.org/10.1080/13664530300200184>
- Pascarella, E. T. y Terenzini, P. T. (2005). *How college affects students: A third decade of research*. Jossey-Bass.
- Pasquinelli, E., Farina, M., Bedel, A. y Casati, R. (2021). Naturalizing critical thinking: Consequences for education, blueprint for future research in cognitive science. *Mind, Brain, and Education*, 15(2), 168-176. <https://doi.org/10.1111/mbe.12286>
- Pereles, A., Ortega-Ruipérez, B. y Lázaro, M. (2024). A digital world toolkit: Enhancing teachers' metacognitive strategies for student digital literacy development. *RIED-Revista Iberoamericana de Educación a Distancia*, 27(2), 267-294. <https://doi.org/10.5944/ried.27.2.38798>
- Plummer, K. J., Kebritchi, M., Leary, H. M. y Halverson, D. M. (2022). Enhancing critical thinking skills through decision-based learning. *Innovative Higher Education*, 47(4), 711-734. <https://doi.org/10.1007/s10755-022-09595-9>
- Rivas, S. F., Saiz, C. y Ossa, C. (2022). Metacognitive strategies and development of critical thinking in higher education. *Frontiers in Psychology*, 13, Article 913219. <https://doi.org/10.3389/fpsyg.2022.913219>
- Romero Alonso, R., Araya Carvajal, K. y Reyes Acevedo, N. (2025). Role of artificial intelligence in the personalization of distance education: A systematic review. *RIED-Revista Iberoamericana de Educación a Distancia*, 28(1), 9-36. <https://doi.org/10.5944/ried.28.1.41538>
- Ronderos, N., Foster-Heinzer, S., Flick-Holtsch, D., Shavelson, R. J., Mariño, J. P., Solano-Flores, G. y Perfetti, M.

- C. (2024). Construct overlap in cross-national assessment: Critical thinking in the teacher education curricula of two countries. *Journal of Curriculum Studies*, 56(4), 514-535. <https://doi.org/10.1080/00220272.2024.2312392>
- Ruiz-Rojas, L. I., Salvador-Ullauri, L. y Acosta-Vargas, P. (2024). Collaborative working and critical thinking: Adoption of generative artificial intelligence tools in higher education. *Sustainability*, 16(13), Article 5367. <https://doi.org/10.3390/sui16135367>
- Rusdin, D., Mukminatien, N., Suryati, N. y Laksmi, E. D. (2023). Critical thinking in the AI era: An exploration of EFL students' perceptions, benefits, and limitations. *Cogent Education*, 11(1), Article 2290342. <https://doi.org/10.1080/2331186X.2023.2290342>
- Shanto, S. S., Ahmed, Z. y Jony, A. I. (2024). Enriching learning process with generative AI: A proposed framework to cultivate critical thinking in higher education using ChatGPT. *Tuijin Jishu/ Journal of Propulsion Technology*, 45(1), 3019-3029. <https://www.propulsiontechjournal.com/index.php/journal/article/view/4680/3181>
- Tang, K. Y., Chang, C. Y. y Hwang, G. J. (2021). Trends in artificial intelligence-supported e-learning: A systematic review and co-citation network analysis (1998-2019). *Interactive Learning Environments*, 31(4), 2134-2152. <https://doi.org/10.1080/10494820.2021.1875001>
- Thornhill-Miller, B., Camarda, A., Mercier, M., Burkhardt, J. M., Morisseau, T., Bourgeois-Bougrine, S., Vinchon, F., El Hayek, S., Augereau-Landais, M., Mourey, F., Feybesse, C., Sundquist, D. y Lubart, T. (2023). Creativity, critical thinking, communication, and collaboration: Assessment, certification, and promotion of 21st century skills for the future of work and education. *Journal of Intelligence*, 11(3), Article 54. <https://doi.org/10.3390/jintelligence11030054>
- Timms, M. J. (2016). Letting artificial intelligence in education out of the box: Educational cobots and smart classrooms. *International Journal of Artificial Intelligence in Education*, 26, 701-712. <https://doi.org/10.1007/s40593-016-0095-y>
- UNESCO. (2021). *Estrategia de la UNESCO sobre la innovación tecnológica en la educación (2022-2025)*. https://unesdoc.unesco.org/ark:/48223/pf0000378847_spa
- UNESCO. (2024). *AI competency framework for teachers*. <https://unesdoc.unesco.org/ark:/48223/pf0000391104>
- Van Laar, E., Van Deursen, A. J. A. M., Van Dijk, J. A. G. M. y De Haan, J. (2017). The relation between 21st-century skills and digital skills: A systematic literature review. *Computers in Human Behavior*, 72, 577-588. <https://doi.org/10.1016/j.chb.2017.03.010>
- Vazhayil, A., Shetty, R., Bhavani, R. R. y Akshay, N. (2019). Focusing on teacher education to introduce AI in schools: Perspectives and illustrative findings. En 2019 IEEE Tenth International Conference on Technology for Education (T4E) (pp. 71-77). IEEE. <https://doi.org/10.1109/T4E.2019.00021>
- Watson, G. y Glaser, E. M. (1980). *Watson-Glaser Critical Thinking Appraisal, forms A and B*. Psychological Corporation.
- Wegerif, R. (2002). *Literature review in thinking skills, technology and learning* (Report No. 2). NESTA Futurelab. <https://hal.archives-ouvertes.fr/hal-00190219>
- Wilton, L., Ip, S., Sharma, M. y Fan, F. (2022). Where is the AI? AI literacy for educators. En M. M. Rodrigo, N. Matsuda, A. I. Cristea, y V. Dimitrova (Eds.), *Artificial intelligence in education. Posters and late breaking results, workshops and tutorials, industry and innovation tracks, practitioners' and doctoral consortium*

- (pp. 180-188). Springer. https://doi.org/10.1007/978-3-031-11647-6_31
- World Economic Forum. (2023). *Future of jobs report 2023*. https://www3.weforum.org/docs/WEF_Future_of_Jobs_2023.pdf
- Yang, Y. T. C., Newby, T. J. y Bill, R. L. (2008). Facilitating interactions through structured web-based bulletin boards: A quasi-experimental study on promoting learners' critical thinking skills. *Computers & Education*, 50(4), 1572-1585. <https://doi.org/10.1016/j.compedu.2007.04.006>
- Young, D. S. y Casey, E. A. (2019). An examination of the sufficiency of small qualitative samples. *Social Work Research*, 43(1), 53-58. <https://doi.org/10.1093/swr/svy026>
- Yuan, L. y Liu, X. (2025). The effect of artificial intelligence tools on EFL learners' engagement, enjoyment, and motivation. *Computers in Human Behavior*, 162, Article 108474. <https://doi.org/10.1016/j.chb.2024.108474>
- Yusuf, A., Bello, S., Pervin, N. y Tukur, A. K. (2024). Implementing a proposed framework for enhancing critical thinking skills in synthesizing AI-generated texts. *Thinking Skills and Creativity*, 53, Article 101619. <https://doi.org/10.1016/j.tsc.2024.101619>
- Zhang, Q., Lu, J. y Zhang, G. (2021). Recommender systems in E-learning. *Journal of Smart Environment and Green Computing*, 1, 76-89. <https://doi.org/10.20517/jsegc.2020.06>

Fecha de recepción del artículo: 1 de diciembre de 2024

Fecha de aceptación del artículo: 11 de marzo de 2025

Fecha de aprobación para maquetación: 26 de marzo de 2025

Fecha de publicación en OnlineFirst: 9 de abril de 2025

Fecha de publicación: 1 de julio de 2025

APÉNDICE

Tabla 1

Habilidades fundamentales del pensamiento crítico y preguntas para estimularlas

Habilidad	Descripción por consenso de expertos	Sub-habilidades	Preguntas para estimular	Actividades de Evaluación
Interpretación	Comprender y expresar el significado o la relevancia de experiencias, situaciones, datos, eventos, juicios, reglas, procedimientos o criterios.	Categorizar. Decodificar significados. Aclarar conceptos.	¿Qué significa esto? ¿Qué está sucediendo? ¿Cómo deberíamos interpretar esta afirmación o situación? ¿Qué factores influyen en esta interpretación?	Identificar problemas y describirlos sin sesgos. Distinguir las ideas principales de las secundarias en un texto. Clarificar gráficos o esquemas.
Análisis	Identificar las relaciones inferenciales, intencionadas o reales entre afirmaciones, conceptos y representaciones de ideas o juicios.	Examinar ideas. Identificar argumentos. Detectar premisas.	¿Cuáles son las razones para esta afirmación? ¿Qué está concluyendo? ¿Cuáles son los argumentos a favor o en contra? ¿Qué supuestos subyacen a esta postura?	Identificar similitudes, diferencias y supuestos implícitos en argumentos. Representar gráficamente relaciones entre ideas clave.
Inferencia	Identificar los elementos necesarios para formular conclusiones razonables, construir hipótesis y analizar información relevante para prever consecuencias.	Consultar evidencias. Formular alternativas. Extraer conclusiones válidas.	¿Qué conclusiones podemos extraer con esta información? ¿Qué implica esta evidencia? ¿Qué consecuencias podemos prever? ¿Qué información adicional necesitamos?	Formular hipótesis, anticipar implicaciones y diseñar estrategias para abordarlas. Validar hipótesis mediante experimentos.

Habilidad	Descripción por consenso de expertos	Sub-habilidades	Preguntas para estimular	Actividades de Evaluación
Evaluación	Evaluar la credibilidad de afirmaciones y la coherencia lógica de sus relaciones.	Evaluar credibilidad. Evaluar argumentos deductivos/ inductivos.	¿Qué criterios nos permiten determinar la calidad de la evidencia? ¿Cuán sólidos son estos argumentos? ¿Qué falacias pueden estar presentes en estos razonamientos? ¿La evidencia presentada es confiable?	Evaluar la credibilidad de las fuentes, la solidez de los argumentos y la coherencia del razonamiento, identificando posibles contradicciones o sesgos.
Explicación	Justificar el razonamiento basado en evidencia, conceptos, métodos y criterios, y presentarlo de manera estructurada y coherente.	Comunicar resultados. Justificar procedimientos. Presentar argumentos.	¿Cuáles son los hallazgos específicos? ¿Qué razones justifican esta conclusión? ¿Qué criterios respaldan esta evaluación? ¿Cómo justificaría esta decisión?	Comunicar hallazgos, justificar decisiones con evidencia y diseñar gráficos para estructurar conceptos jerárquicos.
Autorregulación	Monitorear y evaluar de manera consciente los propios procesos de pensamiento para cuestionar, validar o corregir razonamientos y resultados.	Cuestionar juicios propios. Confirmar resultados. Validar o corregir razonamientos.	¿Estamos siendo lo suficientemente precisos? ¿Cuestionamos nuestras suposiciones? ¿Qué consecuencias podemos prever? ¿Qué estamos pasando por alto?	Revisar juicios propios considerando sesgos. Verificar comprensión y corregir errores en cálculos, interpretaciones o conclusiones.

Fuente: adaptado de Facione (2023).

Tabla 2
Marco para analizar críticamente la información generada por la IA

Fase	Propósito	Acciones Principales con IA	Resultado Esperado
1. Familiarización	Explorar y comprender el contenido generado por IA.	Leer el texto generado por IA de manera crítica. Identificar sesgos, supuestos o patrones recurrentes en el contenido. Reflexionar sobre el contexto en el que se produce la información.	Comprensión inicial sólida del contenido y detección de posibles limitaciones inherentes al uso de IA.
2. Conceptualización	Organizar las ideas clave identificadas en el contenido generado por IA.	Extraer conceptos principales del contenido generado. Diseñar mapas visuales que conecten ideas y utilizarlos como referencia para análisis posteriores.	Construcción de una estructura sólida para profundizar en el contenido y fomentar un análisis crítico.
3. Indagación	Cuestionar y analizar de manera crítica los conceptos generados por la IA.	Formular preguntas dirigidas a la IA, usando <i>prompts</i> específicos (“¿por qué?”, “¿cómo?”). Evaluar la relevancia y profundidad de las respuestas generadas, considerando enfoques alternativos al contenido proporcionado por la IA.	Desarrollo de habilidades críticas a través del cuestionamiento activo y la exploración de diversas perspectivas.

Fase	Propósito	Acciones Principales con IA	Resultado Esperado
4. Evaluación	Validar la credibilidad, relevancia y coherencia de las respuestas generadas por IA.	Contrastar las respuestas de la IA con fuentes externas confiables. Evaluar la consistencia de las respuestas frente a evidencias adicionales. Identificar y corregir posibles errores o desinformación generada por la IA.	Desarrollar la capacidad de evaluar críticamente la validez de los argumentos, distinguiendo entre aquellos con fundamento sólido y los que carecen de respaldo.
5. Síntesis	Integrar el contenido generado por IA con fuentes externas y conocimientos previos.	Combinar información de la IA con fuentes externas y conocimientos previos.	Fomentar una comprensión estructurada y conectada del contenido, facilitando su aplicación a distintos contextos.

Fuente: adaptado de Yusuf et al. (2024).

Tabla 3
Perfil ciego de los informantes

Infor- mante	Figura contractual	Sexo	Área de conocimiento	Experiencia en gestión e innovación educativa
1	Profesor Titular de Universidad	Hombre	Didáctica	Especialista en innovación educativa, virtualización y diseño instruccional.
2	Catedrático de Universidad	Hombre	Estadística e Investigación Operativa	Experto en IA y ética tecnológica, con experiencia en liderar estrategias de transferencia del conocimiento y alianzas universidad-empresa.
3	Profesora Ayudante Doctora	Mujer	Didáctica y Tecnología Educativa	Especialista en tecnología educativa, con liderazgo en innovación, planificación de metodologías, formación docente y diseño de modelos de aprendizaje autónomo.

Informante	Figura contractual	Sexo	Área de conocimiento	Experiencia en gestión e innovación educativa
4	Profesor Titular de Universidad	Hombre	Historia Económica	Docente con liderazgo institucional en innovación educativa y tecnologías aplicadas a la enseñanza, responsable de la planificación y adaptación de marcos metodológicos en entornos virtuales.
5	Técnico superior	Mujer	Pedagogía	Profesional en innovación educativa y formación pedagógica, con experiencia en asesoramiento docente.
6	Profesor Titular de Universidad	Hombre	Tecnología Electrónica	Ejerce liderazgo institucional en innovación educativa y es responsable de observatorios académicos especializados en tecnología aplicada a la enseñanza.
7	Catedrático de Universidad	Hombre	Economía Financiera y Contabilidad	Investigador especializado en redes neuronales artificiales y en la divulgación del impacto de las TIC.
8	Profesora Agregada	Mujer	Dirección de Empresas	Docente con experiencia en gestión estratégica e innovación educativa, especializada en la aplicación de IAG.
9	Profesor Titular de Universidad	Hombre	Ingeniería de Sistemas y Automática	Director académico con experiencia en la integración de tecnologías digitales e IAG en entornos de aprendizaje; lidera proyectos de desarrollo tecnológico y transformación digital institucional.
10	Profesora Asociada	Mujer	Filosofía	Profesional dedicada a promover el pensamiento crítico y la ética en la toma de decisiones en entornos laborales.
11	Profesor de Universidad Privada	Hombre	Filosofía	Docente especializado en fortalecer el pensamiento crítico en entornos digitales y analizar las implicaciones filosóficas y educativas de la singularidad tecnológica.

Fuente: elaboración propia.

Tabla 4
Descriptores sobre los clústeres de la Figura 3

Clúster	Observaciones
Clúster 01: Tecnologías (31,9 %)	Situado en la parte inferior derecha, reúne lemas como IA, herramienta, tarea y estudiante, lo que apunta a una reflexión sobre el impacto de las tecnologías en los procesos educativos. Además, destaca el papel de los estudiantes como actores principales en un entorno educativo transformado por la tecnología.
Clúster 02: Competencias (25,5 %)	En la parte inferior izquierda, este clúster se enfoca en lemas como alfabetización, educación, ético, capacidad e información, reforzando la idea de que el pensamiento crítico debe integrarse en los programas educativos como una competencia esencial que abarque aspectos técnicos y éticos. Los lemas desigualdad, brecha y contexto sugieren una preocupación por el impacto diferencial de estas tecnologías y subraya la necesidad de alfabetización digital y equidad en el acceso a recursos. Representa el núcleo conceptual del pensamiento crítico como objetivo educativo.
Clúster 03: Evaluación (19,9 %)	Ubicado en la parte superior derecha, este grupo temático agrupa lemas como formación, evaluación, preguntas, respuesta y actividad, reflejando una aproximación metodológica al pensamiento crítico. De hecho, se vincula con la importancia de diseñar metodologías innovadoras que fomenten la participación y el análisis reflexivo en el aula. Asimismo, hace referencia al papel del docente en la implementación de actividades que desarrollen habilidades críticas.
Clúster 04: Normativa (22,7 %)	Localizado en la parte superior izquierda incluye lemas como políticas, normativa, sociedad y soluciones, lo que subraya la dimensión estructural y política del pensamiento crítico e indican que su implementación en la educación requiere de marcos regulatorios sólidos y un enfoque estratégico que permita la creación de condiciones favorables para la ejecución de enfoques éticos e inclusivos en el aprendizaje. La distribución de los clústeres sugiere que el pensamiento crítico depende de actividades educativas específicas y de factores normativos que enmarcan su implementación.

Fuente: elaboración propia.

Tabla 5

Distribución de las ideas-fuerza por dimensiones incluidas en las entrevistas

Dimensión	Ideas – Fuerza	Cuota
Barreras	172	21 %
Aceleradores	162	20 %
Alternativas socio-tecnológicas	159	20 %
Retos sociales	153	19 %
Consecuencias visibles	155	19 %

Fuente: elaboración propia.

Tabla 6

Consensos sobre las barreras expresados por los informantes participantes en el estudio

Barreras para el desarrollo del pensamiento crítico	1	2	3	4	5	6	7	8	9	10	11
Desfase entre el rápido avance de la IA y la lenta adaptación de los sistemas educativos.	✓	✓	✓								
Resistencia al cambio en todos los niveles institucionales y actores educativos.	✓	✓	✓								
Formación insuficiente del profesorado en pensamiento crítico y uso de IA.	✓			✓			✓		✓		
Falta de evaluación y priorización del pensamiento crítico en los sistemas educativos.		✓		✓					✓	✓	
Uso acrítico y riesgo de plagio en la utilización de IA por parte de los estudiantes.						✓	✓	✓			✓
Ausencia de un marco normativo y legal claro para el uso educativo de la IA.			✓						✓		
Carencia de alfabetización específica en IA y pensamiento crítico.								✓		✓	✓

Barreras para el desarrollo del pensamiento crítico	1	2	3	4	5	6	7	8	9	10	11
Contexto social y mediático desfavorable para el desarrollo del pensamiento crítico.				✓							✓
Falta de integración entre IA y actividades de pensamiento crítico.					✓			✓			✓
Percepción errónea de la IA como sustituto en lugar de complemento educativo.		✓							✓		✓

Fuente: elaboración propia.

Tabla 7
Consensos sobre los aceleradores expresados por los informantes participantes en el estudio

Aceleradores para el desarrollo del pensamiento crítico	1	2	3	4	5	6	7	8	9	10	11
Actualización de competencias transversales y específicas en el uso crítico de la IA.	✓	✓	✓	✓	✓	✓	✓				
Impulso a la formación y motivación del profesorado como líderes de la transformación educativa.		✓					✓			✓	
Transformación de los sistemas evaluativos hacia la valoración de procesos con IA.			✓	✓		✓	✓				
Promoción de la educación en el uso ético y responsable de la IA.			✓	✓		✓		✓			
Implementación prioritaria de alfabetización en IA desde el inicio universitario.								✓			✓
Alineación de políticas educativas con los avances tecnológicos.		✓	✓			✓					
Fortalecimiento de la conexión academia-mercado laboral en competencias IA.		✓				✓			✓		

Aceleradores para el desarrollo del pensamiento crítico	1	2	3	4	5	6	7	8	9	10	11
Optimización de procesos y gestión del tiempo mediante IA supervisada.					✓		✓		✓		
Facilitación del acceso equitativo a la IA para reducir brechas.								✓	✓		✓
Impulso de sistemas de autoevaluación y evaluación por pares.				✓	✓		✓				

Fuente: elaboración propia.

Tabla 8

Consensos sobre las alternativas socio-tecnológicas expresados por los informantes participantes en el estudio

Alternativas socio-tecnológicas para el desarrollo del pensamiento crítico	1	2	3	4	5	6	7	8	9	10	11
Desarrollo de habilidades de evaluación y validación de herramientas IA.	✓	✓						✓			
Implementación de sistemas de interacción crítica con IA que fomenten el cuestionamiento.	✓			✓							
Creación de <i>frameworks</i> para comprender limitaciones y sesgos algorítmicos.	✓	✓						✓			
Integración de valores éticos en el uso educativo de la IA.	✓		✓		✓						
Establecimiento de programas de capacitación docente en uso crítico de IA.		✓								✓	
Incorporación temprana de alfabetización digital crítica en IA.		✓	✓								
Democratización del acceso a herramientas de IA en educación.				✓		✓	✓				✓
Desarrollo de modelos de uso de IA como herramienta complementaria.					✓					✓	

Alternativas socio-tecnológicas para el desarrollo del pensamiento crítico	1	2	3	4	5	6	7	8	9	10	11
Creación de entornos universitarios seguros para experimentación con IA.							✓		✓		
Implementación de sistemas de debate socrático asistidos por IA.								✓	✓		
Establecimiento de protocolos de transparencia y documentación en uso de IA.								✓	✓		

Fuente: elaboración propia.

Tabla 9

Consensos sobre retos sociales expresados por los informantes participantes en el estudio

Retos sociales para el desarrollo del pensamiento crítico	1	2	3	4	5	6	7	8	9	10	11
Integración efectiva de la alfabetización digital y ética en IA en la educación básica.	✓	✓		✓	✓	✓	✓	✓			
Reducción de brechas tecnológicas y desigualdades en el acceso a la IA.		✓		✓		✓	✓	✓	✓		
Superación de la dependencia acrítica y fomento de la autonomía de pensamiento.			✓				✓			✓	
Desarrollo de capacidades de evaluación y cuestionamiento de información IA.			✓	✓		✓	✓		✓	✓	
Transformación del rol docente hacia la integración pedagógica de IA.				✓	✓				✓		
Establecimiento de marcos regulatorios y éticos para el uso de IA.						✓			✓		
Implementación de educación ética para la gobernanza global de IA.						✓					✓
Garantía de acceso equitativo y adaptativo a herramientas IA.								✓	✓		

Fuente: elaboración propia.

Tabla 10

Consensos sobre las consecuencias visibles expresados por los informantes participantes en el estudio

Consecuencias visibles del desarrollo del pensamiento crítico	1	2	3	4	5	6	7	8	9	10	11
Prevención del pensamiento homogéneo y desarrollo de capacidad de cuestionamiento de la IA.	✓		✓			✓	✓				✓
Transformación hacia evaluaciones basadas en competencias analíticas y procesos creativos.	✓	✓			✓			✓			
Mitigación de riesgos educativos mediante alfabetización y transparencia en el uso de IA.	✓					✓	✓				
Revalorización del factor humano en el proceso educativo.			✓					✓			✓
Mayor competitividad y adaptabilidad profesional en entornos automatizados.				✓				✓	✓	✓	
Mejora de la calidad del debate público y fortalecimiento democrático.				✓							✓
Optimización del tiempo educativo para análisis crítico.					✓		✓	✓	✓		
Potencial amplificación de desigualdades educativas y sociales.						✓	✓	✓	✓		
Personalización efectiva de la educación mediante IA.					✓				✓		
Fortalecimiento de la toma de decisiones éticas en contextos tecnológicos.			✓		✓	✓					
Mejora en resultados y motivación mediante proyectos prácticos con IA.		✓					✓				
Desarrollo de una sociedad más madura y exigente.										✓	✓

Fuente: elaboración propia.

Tabla 11

Ejemplos de expresiones literales indicadas por los informantes para cada dimensión

Dimensión	Idea-Fuerza	Ejemplos de expresiones literales	Informante
Barreras	Desfase entre el rápido avance de la IA y la lenta adaptación de los sistemas educativos	<i>“Hay una barrera, y es que no hemos educado, no estamos educando y además es que ahora encima vivimos en una etapa de auténtico tsunami con esto de la IA, pero sin embargo hablando de educación, pues claro, nuestros alumnos llegan a la universidad y ¿qué pasa? Que se está viviendo un negacionismo digital, prohibiendo; profesores innovadores que quieren hacer algo con pedagogías emergentes, por ejemplo, que tienen que ver con las tecnologías emergentes de cada momento y se capan”.</i>	3
Aceleradores	Actualización de competencias transversales y específicas en el uso crítico de la IA	<i>“¿Qué queremos que aprendan nuestros estudiantes? ¿Qué queremos que sea un filósofo? ¿Qué queremos que sea un filólogo? Pues habrá que repensar las competencias incluyendo estas nuevas herramientas e incluyendo competencias que en ese caso yo creo que es el pensamiento crítico seguramente que sería de la naturaleza de lo que llamamos competencias transversales, que son en inglés las soft skills, es decir que competir, igual que eso la IA, pues también es interesante saber utilizar medios, saber organizarse, saber buscar referencias, saber colaborar con otros, etcétera, la autonomía personal. Bueno, pues tal vez habrá que revisar. Esto es un tipo de competencias que a veces no somos conscientes de que debemos revisar y aquí hay un ejemplo de que toca revisar, toca revisar porque llaman a la puerta, entonces ahora ya no lo podemos ignorar”.</i>	1
Alternativas sociotecnológicas	Desarrollo de habilidades de evaluación y validación de herramientas IA	<i>“Nosotros, por ejemplo, si vamos desde el punto de vista de evaluación de competencias, por una parte, para el uso de IA, dividido en dos partes; por una parte, los procesos de evaluación, una parte que es el desarrollo de un trabajo sobre una serie de temas que les expongo previamente, en la parte de la que estoy son temas de la inteligencia artificial o para la buena inteligencia artificial, que ellos tienen que hacer una aplicación, desarrollarla y defenderla”.</i>	2

Dimensión	Idea-Fuerza	Ejemplos de expresiones literales	Informante
Retos sociales	Integración efectiva de la alfabetización digital y ética en IA en la educación básica	<i>“Es fundamental que realmente aprendamos a utilizar la IA, aprendamos a utilizarla como una herramienta para trabajar con ella [...] también es verdad que, sobre todo en un entorno profesional, lo que se va a valorar no es realmente el proceso, sino es lo que tú ofrezcas”.</i>	4
Consecuencias visibles	Prevención del pensamiento homogéneo y desarrollo de capacidad de cuestionamiento de la IA	<i>“Podríamos coger un problema y preguntar a dos IAs y probablemente las respuestas que te dan las dos IAs, como tú me decías, ya sean distintas, entonces que el estudiante lo que tiene que hacer es compararlo [...] El riesgo, y a mí me preocupa, es cómo van a enjuiciar algo cuando la visión que tienen del todo es un todo muy sesgado y muy parcial, y en realidad muchas veces, por ahora mismo, las alucinaciones que genera [la IA] son probabilísticas”.</i>	4

Fuente: elaboración propia.

El estudiantado de secundaria ante la RV en materias STEM. Efecto de la variable de género

Perceptions of secondary school students towards virtual reality in STEM subjects. Effect of the gender variable



- David García-Marín - Universidad Rey Juan Carlos, URJC (España)
Ricardo Roncero Palomar - Universidad Rey Juan Carlos, URJC (España)
Marina Santín - Universidad Rey Juan Carlos, URJC (España)
Victoria– Mora de la Torre - Universidad Rey Juan Carlos, URJC (España)

RESUMEN

El presente estudio pretende medir las percepciones y actitudes del estudiantado de Secundaria hacia el uso de la realidad virtual (RV) en materias científicas y tecnológicas, así como analizar los posibles sesgos de género en la valoración de este recurso. Este segundo objetivo se justifica en la escasez de trabajos que aúnan el uso de la RV para la formación STEM con la variable de género. Se llevó a cabo un estudio cuasi-experimental ($n = 510$) basado en la aplicación en el aula y utilización por parte del estudiantado de cuatro lecciones de asignaturas STEM en RV elaboradas *ad hoc* para esta investigación en tres centros de Secundaria españoles de diferentes entornos poblacionales y con distintos niveles de experiencia de uso de esta tecnología. Se utilizó para ello el test Instructional Material Motivational Survey (IMMS) validado en anteriores estudios. Los datos resultantes fueron analizados mediante estadística descriptiva e inferencial. Nuestros resultados evidencian que los aspectos mejor valorados de la RV son los relativos a la estructura y diseño de las lecciones, así como su capacidad para facilitar la atención en el contenido. Se observa un notable efecto de la variable de género. Las mujeres perciben de forma significativa una mayor dificultad en la usabilidad de las lecciones y afirman que la experiencia con RV les ayuda menos a mantener la atención. Manifiestan haber aprendido menos que sus compañeros varones y se sienten menos confiadas en su aprendizaje durante el uso de estas tecnologías inmersivas.

Palabras clave: realidad virtual; STEM; educación secundaria; sesgo de género; estudio cuasi-experimental.

ABSTRACT

This paper aims to examine the perceptions and attitudes of students in the secondary educational cycle towards the use of VR in scientific and technological subjects, as well as to analyze possible gender biases in the valuation of this technology. This second objective is justified by the scarcity of research that combines the use of VR for STEM education with the gender variable. A quasi-experimental study was carried out ($n = 510$) based on the application in the classroom and use by students of four VR STEM lessons developed *ad hoc* for this research in three Spanish secondary schools located in different population environments and with different levels of experience in the use of this technology. The Instructional Material Motivational Survey (IMMS) test –validated in previous studies– was used for this purpose. The resulting data were analyzed using descriptive and inferential statistics based on a predictive factor analysis using ordinal logistic regression. The results show that the most highly rated aspects of VR are those related to the structure and design of the lessons, as well as their potential to facilitate attention to the content. A notable effect of the gender variable is detected. Women significantly perceive greater difficulty in the usability of the lessons and report that the VR experience helps them less to maintain attention. They claim to have learned less than their male peers. They also feel less confident in their learning while using these immersive technologies.

Keywords: virtual reality; STEM; secondary school; gender bias; quasi-experimental study.

INTRODUCCIÓN

La formación sobre las disciplinas STEM (Ciencia, Tecnología, Ingeniería y Matemáticas) se ha demostrado esencial para dar respuesta a la creciente demanda de habilidades en estas materias en el mundo laboral. En toda la etapa educativa y concretamente en el ciclo de Bachillerato, la formación STEM resulta fundamental y puede considerarse una inversión estratégica para el progreso personal de los estudiantes y, por extensión, de la sociedad en su conjunto, pues en esa etapa se asientan las bases para el desarrollo de habilidades críticas como el pensamiento lógico, la resolución de problemas complejos y la innovación.

En torno a este campo y al estudio integrado de estas disciplinas, los trabajos académicos han prestado atención a la capacitación docente del profesorado (Boel et al., 2023), los métodos para evaluar el rendimiento y las competencias (Aguilera et al., 2022), la aplicación de métodos pedagógicos innovadores (Fuentes Hurtado y González Martínez, 2017) o el uso de la tecnología en el aula (Marrero Galván y Hernández Padrón, 2022) sin dejar de lado los análisis sobre la equidad (Archer et al., 2020; Martín Carrasquilla et al., 2022).

A la obligación de formar a los más jóvenes en esas materias se une la necesidad de prestar mayor atención al alumnado femenino. La literatura científica ha identificado que existe una brecha de género en las áreas STEM en todos los niveles de la educación y, por ende, en el mercado laboral. El desequilibrio se percibe desde edades tempranas. Martín Carrasquilla et al. (2022) ponen en evidencia que, entre los 10 y los 14 años, los estudiantes varones manifiestan actitudes más favorables hacia la ciencia. Lejos de minimizarse, la brecha aumenta a lo largo de la Educación Secundaria. Las causas que parecen estar detrás de este sesgo responden a razones psicológicas (creencias individuales sobre la propia habilidad y cómo se valora), socioculturales (influencia parental y socialización, discriminación, estereotipos y roles de género) e incluso biológicas, según Morales Inga y Morales Tristán (2020).

Para reducir esta brecha, las administraciones públicas buscan con distintos programas promover la inclusión de las mujeres en el campo de las ciencias y la tecnología. En sintonía con esas acciones, la tecnología se ha integrado en el sistema educativo como una vía con la que aumentar el interés y la motivación hacia materias STEM entre el alumnado. La implementación tecnológica ha llevado al aula el uso de la realidad virtual (RV en adelante), la realidad aumentada (RA), la robótica o la inteligencia artificial. De estas tecnologías emergentes en el campo educativo, la RV parece ocupar la primera posición como instrumento más aplicado, seguida de la robótica, “considerándose ambas tecnologías como las que muestran más recorrido temporal y experiencia en su aplicabilidad” (Palacios Ortega et al., 2022, p. 14).

Posiblemente por ello, en los últimos años, la literatura científica sobre RV y educación ha sido crecientemente prolífica. Los trabajos que analizan las experiencias en las aulas señalan que la RV combinada con el aprendizaje basado en problemas o la indagación resulta de gran utilidad para la formación en materias STEM (Marrero

Galván y Hernández Padrón, 2022). El elemento motivacional aparece destacado en todas las experiencias al margen de que los experimentos versen o no sobre asignaturas de ciencias. La mayoría de las conclusiones recogidas en los artículos son, en general, positivas y señalan un mayor aprendizaje significativo, un aumento de la eficacia del proceso de enseñanza-aprendizaje (Cabero-Almenara et al., 2023; Bermejo et al., 2023) y una mejora en la motivación (Marrero Galván y Hernández Padrón, 2022). La mayoría de los autores señalan logros cognitivos (adquisición de conocimientos) y afectivos (estimulación, cooperación, etc.) con estas tecnologías, lo que justifica la utilización de la RV para motivar a los estudiantes.

Asimismo, se incide en otros aspectos positivos que permite la RV como son el incremento del compromiso, el acceso a entornos inaccesibles, una mejora en la comprensión de conceptos complejos (Spiriev et al., 2022), el perfeccionamiento de la educación a distancia o incluso el entrenamiento de la empatía (Marrero Galván y Hernández Padrón, 2022). En la misma línea, trabajos como el de Giakoni-Ramírez et al. (2023) demuestran que tanto la RV como la RA tienen prometedores beneficios como apoyo al proceso de enseñanza. Los trabajos de Marín-Díaz y Sampedro-Requena (2023) revelan que, en lo que respecta a aspectos relacionados con la metodología de enseñanza, la RV promueve la autonomía y la iniciativa de los estudiantes en Educación Secundaria.

Desde el punto de vista de las estrategias de elaboración de las producciones en RV para la educación, la investigación en diferentes países concluye que los componentes gamificados en esta tecnología apoyan a los estudiantes en su aprendizaje (Lucena-Antón et al., 2022; Castellano et al., 2023). Numerosos estudios evidencian que los grupos que trabajan con laboratorios virtuales y programas de simulación en RV obtienen resultados significativamente mejores que aquellos que lo hacen de forma tradicional (Arce et al., 2022; Tatiana Cox et al., 2022; Pontes Pedrajas, 2022). Estos estudios destacan también el hecho de que el uso de estos espacios virtuales inmersivos fomenta la motivación del estudiantado y son una herramienta que debe considerarse para el trabajo en clase, confirmando así trabajos anteriores como el de Mercado et al. (2019).

En cuanto a la valoración de los docentes, estudios como el de Castaño-Calle et al. (2022) muestran una opinión favorable hacia estas tecnologías, aunque evidencian el escaso conocimiento que tiene el profesorado de estas. Los resultados apoyan la relevancia de proporcionar conocimientos y potenciar habilidades en el profesorado, fomentando así actitudes positivas hacia estos instrumentos. Sin una capacitación adecuada, muchos educadores pueden sentirse intimidados o incapaces de utilizar estas nuevas tecnologías en sus aulas (De Moraes Rossetto et al., 2023).

En el contexto español, son escasos los trabajos que aúnan el uso de la RV para la formación STEM con la variable de género. Para paliar este déficit, este trabajo pretende no solo medir la percepción del estudiantado de Secundaria hacia el uso de la RV educativa en materias científicas y tecnológicas, sino indagar en los posibles sesgos de género en la valoración de este recurso educativo. En

concreto, nuestro estudio se centra en cuatro dimensiones esenciales para el proceso de enseñanza-aprendizaje: (1) atención, (2) relevancia, (3) confianza y (4) satisfacción del estudiante con la tecnología o metodología utilizada. La elección de estas dimensiones se fundamenta en el modelo ARCS (Attention, Relevance, Confidence, Satisfaction) definido por Keller (1987, 2010), que basa la activación de la motivación en el proceso educativo en un constructo constituido por estos cuatro aspectos. Con respecto a la atención, nuestro trabajo pretende medir la capacidad que tiene la herramienta para hacer que el alumnado atienda exclusivamente a las tareas a realizar dentro del entorno en RV. La relevancia se define como el grado de utilidad percibida de los aprendizajes. La confianza se relaciona con la seguridad del estudiante en el aprendizaje durante la interacción con los materiales, se encuentra estrechamente conectada con las experiencias pasadas (Cózar Gutiérrez et al., 2019) y resulta la dimensión más complicada de mejorar porque requiere de un mayor tiempo de dedicación (Huett et al., 2008). Estas tres dimensiones confluyen en la cuarta, la satisfacción, “resultando en la predicción de buenos resultados durante la realización de la tarea” (Cózar Gutiérrez et al., 2019, p. 2005). De acuerdo con Cabero et al. (2017), estas cuatro dimensiones tienen una relación encadenada de forma que la atención resulta esencial para la relevancia que, a su vez, conduce a la confianza en el aprendizaje y, como enunciábamos anteriormente, el conjunto de estas tres dimensiones determina la satisfacción.

A las cuatro dimensiones anteriormente definidas se añadieron otras dos específicas de la tecnología en RV: (1) el diseño, jugabilidad e interactividad de la herramienta y (2) su facilidad de uso y navegación.

Objetivos e hipótesis

Teniendo todo esto en cuenta, nuestro trabajo persigue los siguientes objetivos:

- O1. Medir la percepción y actitud del estudiantado en Secundaria hacia la aplicación de RV educativa en materias STEM, en especial tomando como referencia las dimensiones de (1) diseño e interactividad, (2) facilidad de uso, (3) atención, (4) relevancia, (5) confianza y (6) satisfacción del estudiantado con esta tecnología.
- O2. Examinar las posibles diferencias de género en la percepción y actitud hacia esta tecnología en su aplicación a la enseñanza STEM. Este objetivo se despliega en un subobjetivo (O2a), condicionado a la existencia de diferencias de género estadísticamente significativas en las percepciones y actitudes, consistente en la medición de la variable de género como factor predictor de las dimensiones expuestas en el O1.

La hipótesis de partida establece que no deberían detectarse sesgos de género estadísticamente significativos en ninguna de las seis dimensiones del estudio. Aunque existen numerosos trabajos que evidencian que las docentes de sexo femenino muestran una actitud más positiva y una mayor inclinación que sus compañeros varones hacia el uso de la RV y de otras tecnologías avanzadas con fines educativos (Ateş y Kölemen, 2024), los estudios específicos practicados sobre estudiantes no observan diferencias de género en la percepción sobre la RV educativa ni en el ciclo de Educación Secundaria (Cózar Gutiérrez et al., 2019) ni a nivel universitario (Alsalleeh et al., 2024). En esta línea, investigaciones recientes realizadas con estudiantes de Secundaria descartan la influencia del género en la aceptación de las tecnologías de vanguardia con fines pedagógicos (An et al., 2024).

Método

Diseño de la investigación

Se llevó a cabo un estudio cuasi-experimental basado en la aplicación en el aula y utilización por parte del estudiantado de cuatro lecciones o píldoras educativas de asignaturas STEM en RV con contenido propio del currículo de Educación Secundaria en España para las asignaturas de Física, Química, Matemáticas y Biología. El trabajo de campo se ejecutó en tres centros con casuísticas diferentes situados en la provincia de Toledo y en dos ciudades de la comunidad de Madrid. El centro de Toledo pertenece a un entorno rural (Centro Rural, CR) con un tamaño poblacional pequeño (10.000 habitantes) y rentas bajas. Uno de los centros de Madrid se sitúa en un municipio urbano (72.500 habitantes con rentas muy altas) y no cuenta con especialización en RV (Centro Urbano No Especializado, CUNE) y el restante se emplaza también en contexto urbano (171.700 habitantes con rentas medias) y tiene especialización en esta tecnología (Centro Urbano Especializado).

Una de las principales fortalezas del trabajo de campo reside en que las cuatro píldoras educativas fueron elaboradas *ad hoc* de forma específica y exclusiva para este estudio (no se trabajó con material educativo en RV preexistente). La información sobre estas lecciones está disponible en: <https://bit.ly/3Ye9AUg>

La intervención se dividió en las siguientes fases:

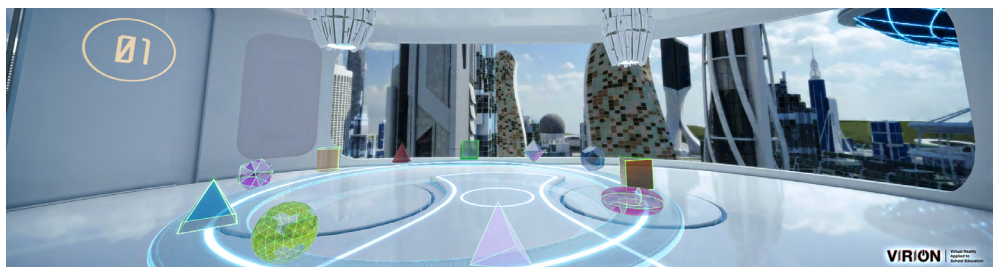
- Fase 1: Elección del contenido de las píldoras educativas. El equipo de investigación y la empresa especializada en RV contaron con el asesoramiento de los docentes del ámbito STEM de los tres centros participantes, quienes en esta fase preliminar del estudio eligieron los contenidos específicos de cada una de las cuatro píldoras educativas. Los contenidos elegidos fueron: (1) la geometría y visualización espacial (Matemáticas), (2) la estructura

y las funciones celulares (Biología), (3) la tabla periódica y sus elementos (Química) y (4) conceptos de electricidad y magnetismo (Física).

- Fase 2: Diseño y elaboración de las píldoras educativas por la empresa encargada de la realización del software educativo en RV, en coordinación con el equipo de investigación (Figura 1).

Figura 1

Captura de la lección en RV de Matemáticas



- Fase 3: Formación de los docentes de los tres centros implicados en el uso de la RV. Una vez elaboradas las cuatro lecciones en RV, varios docentes de los tres institutos de Secundaria participantes en el proyecto asistieron durante una semana a un taller formativo (Figura 2) organizado en la universidad del equipo investigador para familiarizarse con las piezas resultantes y recibir instrucción sobre su aplicación durante las pruebas con su alumnado (distancia de seguridad entre los estudiantes, necesidad de requerimientos informáticos en el aula, incompatibilidad de esta tecnología con determinados problemas de salud, etc.).

Figura 2

Taller formativo en RV impartido a los profesores participantes



- Fase 4. Aplicación del software educativo en RV. Cada centro contó con el equipamiento para la realización de las pruebas durante dos semanas (Figura 3). De forma previa a su utilización por parte del alumnado, los investigadores, con el apoyo de los docentes, explicaron las actividades a realizar en cada lección, así como el funcionamiento de los equipos con el fin de nivelar las posibles diferencias de experiencia previa en el uso de esta tecnología. El equipamiento utilizado consistió en las gafas Meta Quest 2, uno de los dispositivos de RV con mayor penetración en el mercado. La elección de este equipo se basó en su seguridad a la hora de ser utilizadas por adolescentes. El dispositivo es ajustable y no oprime la cara ni la cabeza. Es adaptable para las personas que utilizan gafas. No producen ningún daño a la vista, funcionan con batería recargable (como un *smartphone*) y no se utilizan conectadas a la corriente eléctrica, aspecto que facilitó su integración en el aula.

Figura 3

Trabajo con los materiales en RV en uno de los centros educativos



- Fase 5. Realización de un cuestionario por parte del alumnado. Justo al término del trabajo con cada una de las píldoras, el alumnado debía cumplimentar un cuestionario (Tabla 1) cuyos resultados se exponen en este artículo.

Muestra e instrumentos

Se utilizó una muestra por conveniencia procurando que los centros participantes tuvieran carácter público, procedieran de contextos con diferentes tamaños poblacionales y entornos socioeconómicos y tuvieran diferentes niveles de especialización tecnológica, tal como se explicó en el punto anterior. Dado que se utilizaron los mismos equipos de RV en los tres centros, resultaba necesario que los centros participantes quedaran próximos en términos geográficos y con respecto a la universidad del equipo investigador a fin de trasladar el material utilizado en el trabajo de campo. Se decidió que un total de tres centros participantes resultaba ideal a fin de tener un volumen suficiente de respuestas y no extender demasiado el tiempo de realización del experimento, que se habría alargado en demasía en caso de haber incorporado a algún centro más.

La muestra estuvo compuesta por jóvenes estudiantes de Secundaria (itinerario de ciencias) entre los 14 y los 16 años. Se obtuvieron un total de 510 respuestas de los tres centros. El promedio de edad de los participantes se sitúa en 15.10 años ($DT = .822$). La distribución por géneros es la siguiente: mujeres ($n = 188$; 36.9 %), hombres ($n = 296$; 58 %), prefiero no decirlo ($n = 26$; 5.1 %). Por centros, la procedencia de la muestra es: CR ($n = 106$; 20.8 %), CUNE ($n = 171$; 33.5 %), CUE ($n = 233$; 45.7 %). El 52,7 % de los participantes no tenía experiencia previa en el uso de la RV. Casi 6 de cada 10 participantes (56 %) se declaró jugador habitual de videojuegos.

Como se indicó anteriormente, el instrumento utilizado fue un cuestionario basado en el test IMMS (Instructional Material Motivational Survey) propuesto por Keller (2010) y validado por Loorbach et al. (2015). En su versión original, el instrumento establece 12 ítems en escala de Likert con 7 opciones de respuesta. Para nuestro estudio, el instrumento sufrió las siguientes adaptaciones: (1) se incluyeron nuevos ítems para incluir las dimensiones de diseño / jugabilidad / interactividad y la facilidad de uso y navegación, (2) se simplificó la redacción de algunas cuestiones para hacerlas más comprensibles al nivel del alumnado y (3) se redujeron las opciones de respuesta a un total de 5 (1=Totalmente en desacuerdo; 5=Totalmente de acuerdo). Numerosos trabajos que han utilizado este mismo instrumento han realizado la misma adaptación de la escala a 5 puntos (Cardoso-Júnior y Delbone de Faria, 2021; Huang y Hew, 2016). Esta reducción de los tramos se justifica en que las escalas de 5 puntos son más fáciles de comprender y responder para los participantes, por lo que reduce la carga cognitiva durante su cumplimentación. El equipo investigador consideró este aspecto fundamental dado que los estudiantes

debían responder el test justo después de haber probado el material en RV, con el consiguiente cansancio que esta actividad podría comportar en algunos estudiantes. La juventud del alumnado fue otro de los motivos para reducir el número de puntos, dado que con menos opciones los participantes podrían tener menos dudas sobre qué opciones elegir. En determinados contextos, como es el caso de audiencias poco familiarizadas con la realización de encuestas, una escala de 5 puntos puede resultar más adecuada.

Para medir la fiabilidad del instrumento, se ejecutaron pruebas de consistencia interna por medio del coeficiente alfa de Cronbach ($\alpha = .946$) y utilizando el Método de Dos Mitades (coeficiente de Spearman-Brown = .896), que revelaron una alta consistencia del cuestionario. Aunque se intentó, no resultó posible realizar una prueba de fiabilidad mediante test-retest por cuestiones operativas de los centros, necesidad de trasladar el equipo al resto de institutos participantes y dificultad de contar exactamente con los mismos estudiantes en ambas pruebas. En todo caso, los resultados de los coeficientes de consistencia interna (altamente satisfactorios), junto con el hecho de que el cuestionario ya había sido previamente refrendado en numerosas investigaciones (Cook y Pankratz, 2024) (incluso reducido a escalas de 5 puntos), llevaron al equipo investigador a validar el instrumento.

La Tabla 1 recoge el cuestionario resultante.

Tabla 1
Cuestionario

Ítem (pregunta) del cuestionario	Categorías
P1. Género	Mujer
	Hombre
	Prefiero no decirlo
P2. Edad	14
	15
	16
P3. Centro	Rural (CR)
	Urbano No Especializado (CUNE)
	Urbano Especializado (CUE)
P4. Asignatura (píldora)	Matemáticas
	Biología
	Química
	Física
P5. Me gusta la asignatura	Likert (1-5)

Ítem (pregunta) del cuestionario	Categorías
P6. ¿Tienes gafas de RV o las has utilizado con anterioridad?	Sí
	No
P7. Tengo experiencia con videojuegos y juego con ellos habitualmente.	Likert (1-5)
Instrumento IMMS	
Dimensión 1. Diseño-Jugabilidad-Interactividad	
P8. La lección parece un juego.	Likert (1-5)
P9. Me ha gustado el diseño de la aplicación.	Likert (1-5)
P10. La lección es interactiva.	Likert (1-5)
Dimensión 2. Facilidad de uso y navegación	
P11. La pieza es fácil de usar.	Likert (1-5)
P12. Las instrucciones de la lección son fáciles de entender.	Likert (1-5)
P13. La navegación es sencilla.	Likert (1-5)
P14. No percibo sensación de mareo al navegar por la lección.	Likert (1-5)
Dimensión 3. Atención	
P15. La experiencia en realidad virtual me ayuda a mantener la atención.	Likert (1-5)
P16. El diseño de la app me ayuda a mantener la atención.	Likert (1-5)
P17. Tener muchos ejercicios dentro de la app me ayuda a mantener la atención.	Likert (1-5)
Dimensión 4. Relevancia	
P18. Esta clase está relacionada con cosas que ya había estudiado.	Likert (1-5)
P19. Después de esta experiencia, tengo ganas de conocer más sobre la lección.	Likert (1-5)
P20. La lección es útil desde el punto de vista educativo.	Likert (1-5)
P21. He aprendido cosas nuevas gracias a la lección.	Likert (1-5)
Dimensión 5. Confianza / seguridad en el aprendizaje	
P22. Mientras trabajaba en esta lección, estaba seguro de que iba a aprender los contenidos.	Likert (1-5)
P23. Después de trabajar en esta lección, me siento seguro de que sería capaz de aprobar un examen sobre el tema.	Likert (1-5)
P24. El diseño de la lección me asegura que voy a aprender los contenidos.	Likert (1-5)

Ítem (pregunta) del cuestionario	Categorías
Dimensión 6. Satisfacción	
P25. He disfrutado tanto en esta clase con realidad virtual que me gustaría saber más sobre su contenido.	Likert (1-5)
P26. Me ha gustado esta lección con realidad virtual.	Likert (1-5)

Fuente: elaboración propia.

Téngase en cuenta que, para alcanzar los objetivos propuestos, las preguntas P8-P26 (integradas en el instrumento IMMS), funcionan como variables dependientes del estudio, mientras que la independiente es la relativa al género del estudiantado.

Análisis de los datos

Los datos resultantes fueron analizados mediante estadística descriptiva e inferencial. Antes de decidir la realización de cálculos paramétricos o no paramétricos, se ejecutaron pruebas de normalidad mediante test de Kolmogorov-Smirnov, que observó ausencia de normalidad ($p < .001$) en todos los ítems del cuestionario (ver resultados completos de las pruebas de normalidad en <https://bit.ly/4gRmD5r>). Por tanto, se optó por aplicar estudios no paramétricos (prueba U de Mann-Whitney) para determinar la existencia de diferencias significativas en los resultados de las respuestas en función del género del estudiantado. Nótese que la variable de género se despliega en tres categorías (Mujer / Hombre / Prefiero no decirlo), por lo que para los cálculos estadísticos se redujeron las opciones solo a Mujer / Hombre considerando la tercera opción (Prefiero no decirlo) como valor perdido.

Una vez detectados los ítems donde se establecen tales diferencias estadísticamente significativas, se analizó si la variable de género resulta factor predictor de los resultados de tales ítems. Para ello, se ejecutaron regresiones logísticas ordinales. Para medir el tamaño del efecto, se utiliza el R^2 (en esta prueba, el pseudo R^2 de Nagelkerke es el más habitual), que es el más recomendado en los estudios de este tipo (Domínguez-Lara, 2017). Todo el trabajo estadístico fue realizado con el programa SPSS v.26 (base de datos completa en <https://bit.ly/4dzW3uK>).

El estudio obtuvo la aprobación del Comité de Ética de la universidad del equipo investigador (número de registro interno: 200420242632024).

Resultados

Datos generales

De acuerdo con los resultados obtenidos (y en cumplimiento del O1), los participantes se muestran generalmente satisfechos con la experiencia educativa en RV ($M = 3.84$) (Tabla 3). Casi 6 de cada 10 (58.6 %) están de acuerdo con la

afirmación «He disfrutado tanto en esta clase con RV que me gustaría saber más sobre su contenido» y prácticamente 3 de cada 4 (73.2 %) afirman que les ha gustado la experiencia educativa en RV (Tabla 2). A pesar de esta satisfacción general, las mujeres evalúan todos los ítems con puntuaciones inferiores con respecto a sus compañeros varones.

Tabla 2
Frecuencia de valores en cada ítem

Ítem (variable dependiente)	1	2	3	4	5
Dimensión 1. Diseño-Jugabilidad-Interactividad					
P8. La lección parece un juego.	6.1 %	6.5 %	14.5 %	26.1 %	46.9 %
P9. Me ha gustado el diseño de la aplicación.	7.1 %	3.9 %	13.1 %	27.8 %	48 %
P10. La lección es interactiva.	5.1 %	3.3 %	12.2 %	26.3 %	53.1 %
Dimensión 2. Facilidad de uso y navegación					
P11. La pieza es fácil de usar.	6.3 %	8.8 %	20.6 %	29.8 %	34.5 %
P12. Las instrucciones de la lección son fáciles de entender.	7.1 %	8.4 %	18 %	31.6 %	34.9 %
P13. La navegación es sencilla.	8 %	9 %	18,8 %	33,1 %	31 %
P14. No percibo sensación de mareo al navegar por la lección.	27.8 %	14.7 %	15.3 %	15.5 %	26.7 %
Dimensión 3. Atención					
P15. La experiencia en realidad virtual me ayuda a mantener la atención.	9.4 %	7.5 %	21 %	27.3 %	34.9 %
P16. El diseño de la app me ayuda a mantener la atención.	9 %	8.6 %	23.3 %	27.6 %	31.4 %
P17. Tener muchos ejercicios dentro de la app me ayuda a mantener la atención.	8.4 %	7.8 %	26.3 %	29.8 %	27.6 %
Dimensión 4. Relevancia					
P18. Esta clase está relacionada con cosas que ya había estudiado.	9.4 %	5.3 %	14.9 %	22.5 %	47.8 %
P19. Después de esta experiencia, tengo ganas de conocer más sobre la lección.	13.7 %	10.4 %	21 %	25.1 %	29.8 %
P20. La lección es útil desde el punto de vista educativo.	8.2 %	10.4 %	16.9 %	28.6 %	35.9 %
P21. He aprendido cosas nuevas gracias a la lección.	20.4 %	14.7 %	23.5 %	18.4 %	22.9 %
Dimensión 5. Confianza / seguridad en el aprendizaje					
P22. Mientras trabajaba en esta lección, estaba seguro de que iba a aprender los contenidos.	15.1 %	14.9 %	30 %	19.4 %	20.6 %
P23. Después de trabajar en esta lección, me siento seguro de que sería capaz de aprobar un examen sobre el tema.	20 %	15.5 %	23.9 %	16.5 %	24.1 %
P24. El diseño de la lección me asegura que voy a aprender los contenidos.	14.7 %	11.8 %	28.2 %	21.2 %	24.1 %
Dimensión 6. Satisfacción					
P25. He disfrutado tanto en esta clase con realidad virtual que me gustaría saber más sobre su contenido.	10.4 %	9.2 %	21.8 %	20.6 %	38 %
P26. Me ha gustado esta lección con realidad virtual.	7.5 %	7.8 %	11.6 %	22.2 %	51 %

Fuente: elaboración propia.

La dimensión más valorada es la relativa al diseño de las piezas como juegos dotados de una alta interactividad ($M = 4.08$). El 73 % de los participantes está de acuerdo con el carácter gamificado de las lecciones («La lección parece un juego»), mientras que casi 8 de cada 10 (79.4 %) consideran que las piezas tienen carácter interactivo. En el polo negativo, resulta destacable que la confianza en el aprendizaje durante el uso de las diferentes lecciones es la dimensión que registra los valores más bajos ($M = 3.17$).

Tabla 3

Promedio y desviación típica generales y desagregados por género

Ítem (variable dependiente)	Media (DT)	Media Mujeres	Media Hombres	pa
Dimensión 1. Diseño-Jugabilidad-Interactividad	4.08			
P8. La lección parece un juego.	4.01 (1.19)	3.97	4.05	.477
P9. Me ha gustado el diseño de la aplicación.	4.06 (1.18)	4.03	4.12	.986
P10. La lección es interactiva.	4.19 (1.10)	4.13	4.23	.528
Dimensión 2. Facilidad de uso y navegación	3.56			
P11. La pieza es fácil de usar.	3.77 (1.19)	3.58	3.91	.002*
P12. Las instrucciones de la lección son fáciles de entender.	3.79 (1.20)	3.61	3.92	.003*
P13. La navegación es sencilla.	3.70 (1.22)	3.54	3.79	.015*
P14. No percibo sensación de mareo al navegar por la lección.	2.98 (1.57)	2.64	3.21	<.001*
Dimensión 3. Atención	3.65			
P15. La experiencia en realidad virtual me ayuda a mantener la atención.	3.71 (1.27)	3.55	3.83	.016*
P16. El diseño de la app me ayuda a mantener la atención.	3.64 (1.25)	3.57	3.72	.218
P17. Tener muchos ejercicios dentro de la app me ayuda a mantener la atención.	3.60 (1.20)	3.39	3.77	<.001*
Dimensión 4. Relevancia	3.56			
P18. Esta clase está relacionada con cosas que ya había estudiado.	3.94 (1.29)	3.91	3.97	.858
P19. Después de esta experiencia, tengo ganas de conocer más sobre la lección.	3.47 (1.37)	3.37	3.56	.181

Ítem (variable dependiente)	Media (DT)	Media Mujeres	Media Hombres	pa
P20. La lección es útil desde el punto de vista educativo.	3.74 (1.27)	3.65	3.80	.128
P21. He aprendido cosas nuevas gracias a la lección	3.09 (1.43)	2.85	3.26	.002*
Dimensión 5. Confianza / seguridad en el aprendizaje	3.17			
P22. Mientras trabajaba en esta lección, estaba seguro de que iba a aprender los contenidos.	3.15 (1.32)	2.95	3.32	.003*
P23. Después de trabajar en esta lección, me siento seguro de que sería capaz de aprobar un examen sobre el tema.	3.09 (1.44)	2.76	3.33	<.001*
P24. El diseño de la lección me asegura que voy a aprender los contenidos.	3.28 (1.34)	3.12	3.41	.017*
Dimensión 6. Satisfacción	3.84			
P25. He disfrutado tanto en esta clase con realidad virtual que me gustaría saber más sobre su contenido.	3.67 (1.34)	3.46	3.79	.008*
P26. Me ha gustado esta lección con realidad virtual.	4.01 (1.27)	3.88	4.10	.268

Fuente: elaboración propia.

*a*Prueba U de Mann-Whitney.

**Nivel de significación:* < .05.

Efecto de la variable de género

Como también recoge la Tabla 3, se observan diferencias notables en las respuestas en función del género de los participantes en 11 de las 19 preguntas formuladas (O2). Este dato obliga a rechazar la hipótesis de partida del estudio.

La dimensión donde se observan las mayores diferencias de género es la relativa a la facilidad de uso. Se registran tales desviaciones en todos los ítems de esta dimensión. Las mujeres perciben de forma significativa una mayor dificultad en la usabilidad de las lecciones ($U = 32180.50$; $p = .002$). Asimismo, tienden a considerar más complicadas las instrucciones para utilizarlas ($U = 32127.50$; $p = .003$), a la vez que consideran que la navegación por el software en RV les resulta menos sencilla en comparación con la percepción de sus compañeros varones ($U = 31318.00$; $p = .015$). Sin embargo, donde se detectan las mayores diferencias –dentro de esta dimensión–

es en la percepción de mareo: las chicas manifiestan sentir una mayor sensación de mareo durante el uso de las piezas en RV ($U = 33423.00$; $p < .001$).

Estas desviaciones en la facilidad de uso percibida podrían estar afectadas por la existencia de diferencias de género significativas en la experiencia previa con tecnologías de RV (P6) y el hábito de jugar a videojuegos (P7). Para comprobar tales diferencias en la pregunta sobre la experiencia previa con tecnologías de RV («P6. ¿Tienes gafas de RV o las habías utilizado con anterioridad?»), se ejecutó un análisis por tablas bivariadas, que observó que los hombres tienen una experiencia previa con esta tecnología casi 8 puntos superior a la manifestada por las mujeres (50.3 % frente a 42.6 %). Por otro lado, a fin de observar las desviaciones de género en el hábito de jugar a videojuegos («P7. Tengo experiencia con videojuegos y juego con ellos habitualmente»), se llevó a cabo la prueba U de Mann-Whitney, cuyos resultados observan que los estudiantes varones tienen un hábito en el uso de videojuegos significativamente superior al reconocido por sus compañeras ($U = 43985.50$; $p < .001$). Estas desviaciones de género relativas al hábito y experiencia con tecnologías de RV y videojuegos podrían, en parte, explicar estas diferencias entre hombres y mujeres en cuanto a la percepción de facilidad de uso de la tecnología educativa en RV testada en nuestro trabajo.

Al igual que en la facilidad del uso, se registran desviaciones notables en todos los ítems de la dimensión relativa a la confianza en el aprendizaje. Las mujeres perciben una menor seguridad de aprendizaje durante el uso de la RV ($U = 32245.00$; $p = .003$), a la vez que manifiestan menor confianza en aprobar un examen tras el trabajo con las lecciones ($U = 34061.00$; $p < .001$). También tienden a percibir en menor medida que el diseño de las piezas en RV asegura el aprendizaje de los contenidos ($U = 31314.00$; $p = .017$).

La atención es la tercera dimensión donde se registran las desviaciones de género más notables. Por un lado, las alumnas participantes en el estudio afirman que la experiencia en RV les ayuda menos a mantener la atención en el contenido ($U = 31301.00$; $p = .016$). También son menos proclives a afirmar que el hecho de realizar muchas actividades durante la experiencia en RV les ayuda a mantener su atención en los contenidos ($U = 32823.00$; $p < .001$).

Finalmente, la relevancia educativa y los aspectos formales de diseño / jugabilidad / interacción son las dimensiones donde el género tiene menor impacto.

Estudio de factores predictivos

Una vez detectados cuáles son los ítems donde se registran diferencias estadísticamente significativas entre las y los jóvenes estudiantes participantes en el estudio (y en cumplimiento del O2a), se ejecutó un estudio de factores predictivos – mediante regresiones logísticas ordinales– para establecer relaciones de causalidad entre los 11 ítems del cuestionario donde se observan tales desviaciones (que funcionan como variables dependientes) y el género del estudiantado (variable

independiente). Para determinar la idoneidad del estudio, se realizó previamente la prueba de líneas paralelas, que resultó satisfactoria ($p > .05$ en todas las variables).

Se observa que el género es factor predictor de las 11 variables dependientes, lo que refuerza la asociación entre la variable de género y los ítems medidos. La Tabla 4 recoge los resultados del efecto del género en todas las variables en orden decreciente.

Tabla 4
Estudio de factores predictivos (regresión ordinal)

Ítem (variable dependiente)	Valor p del modelo (pseudo R^2 Nagelkerke)	β (ES)	Wald	p
P23. Después de trabajar en esta lección, me siento seguro de que sería capaz de aprobar un examen sobre el tema.	<.001 (.039)	.710	17.922	<.001*
P14. No percibo sensación de mareo al navegar por la lección.	<.001 (.031)	-.638	14.428	<.001*
P17. Tener muchos ejercicios dentro de la app me ayuda a mantener la atención.	<.001 (.026)	.596	12.422	<.001*
P21. He aprendido cosas nuevas gracias a la lección.	.002 (.021)	.515	9.570	.002*
P11. La pieza es fácil de usar.	.002 (.020)	.510	9.119	.003*
P22. Mientras trabajaba en esta lección, estaba seguro de que iba a aprender los contenidos.	.002 (.020)	.506	9.179	.002*
P12. Las instrucciones de la lección son fáciles de entender.	.003 (.019)	.505	8.881	.003*
P25. He disfrutado tanto en esta clase con realidad virtual que me gustaría saber más sobre su contenido.	.007 (.015)	.450	7.155	.007*
P13. La navegación es sencilla.	.015 (.013)	.412	5.970	.015*
P15. La experiencia en realidad virtual me ayuda a mantener la atención.	.015 (.013)	.408	5.866	.016*
P24. El diseño de la lección me asegura que voy a aprender los contenidos.	.016 (.012)	.399	5.748	.017*

Fuente: elaboración propia.

*Nivel de significación: < .05.

DISCUSIÓN Y CONCLUSIONES

En términos generales, los estudiantes participantes en el estudio valoran positivamente su experiencia con la RV educativa para materias STEM, si bien existen diferencias relevantes entre las puntuaciones obtenidas en las diferentes dimensiones medidas. Los aspectos mejor valorados son los relativos a la estructura y diseño de las lecciones (interactividad y gamificación), así como su capacidad para ayudarles a mantener la atención en el contenido. Los buenos resultados obtenidos en la cuestión relativa a la utilidad educativa percibida se alinean con la evidencia aportada por trabajos anteriores como los de Cabero-Almenara et al. (2023) y Bermejo et al. (2023). En el polo negativo, la facilidad de uso, la relevancia del contenido percibida y, sobre todo, la confianza en el aprendizaje a través del uso de esta tecnología son las dimensiones peor valoradas. Aun así, las puntuaciones generales en todos los aspectos medidos resultan positivas.

De acuerdo con nuestros resultados, podemos extraer las siguientes conclusiones:

1. Sesgo de género. En línea con recientes estudios como el de Pan et al. (2024), nuestro trabajo evidencia un notable efecto de la variable de género en las actitudes hacia la RV educativa, lo que nos lleva a rechazar la hipótesis de partida del estudio. Trabajos previos ya habían mostrado que la variable de género influye en los resultados de aprendizaje obtenidos en los recursos de aprendizaje dinámicos como la RV (Saha y Halder, 2016; Castro-Alonso et al., 2019). Sin embargo, también existe evidencia empírica que contradice este sesgo de género, tanto en España (Cózar Gutiérrez et al., 2019) como en otros contextos (Ateş y Kölemen, 2024; Alsalleeh et al., 2024).
2. Diferencias en la percepción de facilidad de uso. Estas desviaciones de género son especialmente significativas en las cuestiones de manejo y uso de la aplicación. Quizá las diferencias de género encontradas en la experiencia previa con esta tecnología y en el uso de videojuegos (Halpern, 2000) puedan explicar las desviaciones observadas no solo en la facilidad de uso sino también en otras dimensiones como la confianza en el aprendizaje, la atención o la motivación para seguir conociendo más aspectos de las materias STEM tras la experiencia con RV. Que el género sea un factor predictor de la sensación de mareo refuerza esta hipótesis. En el mismo sentido, estudios previos han evidenciado que los estudiantes varones encuentran mayor facilidad que las mujeres en las tareas de habilidad espacial (Heo y Toomey, 2020; Roach et al., 2021), aspecto clave en el trabajo con recursos educativos en RV. Por este motivo, las alumnas podrían enfrentarse a una carga cognitiva excesiva cuando interactúan con tecnologías que exigen una elevada habilidad espacial (como es el caso de la RV), lo que podría traducirse en un menor rendimiento en el aprendizaje (Jian y Abu Bakar, 2024).

3. Alumnado femenino y materias STEM. En la misma línea, nuestros datos apuntan hacia una posible ineficacia de esta tecnología a la hora de fomentar vocaciones STEM entre el alumnado femenino, al menos en el ciclo de Secundaria y en el contexto español. Por ello, en futuros trabajos, la variable de género debería tenerse en cuenta para que todo el estudiantado pueda obtener experiencias similares en su interacción con los recursos de aprendizaje (Castro-Alonso et al., 2019). En el caso de los materiales inmersivos, se debería reducir la carga cognitiva para que las personas menos acostumbradas al uso de esta tecnología puedan experimentar un adecuado entorno de aprendizaje (Heo y Toomey, 2020).
4. Recomendaciones prácticas. Aunque no es el aspecto central de nuestro estudio, algunas consideraciones prácticas pueden ser tenidas en cuenta a la hora de reducir el trabajo cognitivo con RV a fin de resolver las brechas de género anteriormente señaladas:
 - Se recomienda implementar tutoriales introductorios que expliquen el manejo de las gafas y los controles.
 - El diseño de interfaces debe ser claro e intuitivo a fin de evitar la sobrecarga de información. Se deben establecer controles básicos y una navegación clara y sencilla.
 - Limitar las sesiones a 10-15 minutos por grupo, con intervalos de descanso entre usos.
 - Incorporar un sistema de alertas para recordar a los estudiantes cuándo deben hacer pausas.
 - Priorizar sesiones breves para conceptos concretos, evitando actividades prolongadas y complejas en una sola sesión.
 - Diseñar actividades colaborativas donde los estudiantes trabajen en parejas para optimizar el aprendizaje.
 - Establecer roles alternos: un estudiante usa las gafas y otro guía desde una pantalla o tableta.
 - Implementar un ciclo de retroalimentación para evaluar el impacto pedagógico.
5. Falsa dicotomía entre nativos e inmigrantes digitales. Asimismo, nuestros resultados también ponen de manifiesto la imposibilidad de atribuir a los jóvenes y adolescentes unas percepciones y actitudes homogéneas hacia la tecnología, lo que nos debería llevar a abandonar la dicotomía entre nativos e inmigrantes digitales. Resulta evidente que existen otras variables, al margen de la etaria, que influyen en la facilidad de uso percibida de las tecnologías, lo que sugiere la adopción de un nuevo sistema de distinciones en términos de sabiduría digital (Prensky, 2009), que puede distribuirse de manera disímil dentro del mismo grupo etario en función de otros factores sociales y educativos.

Este trabajo tiene las limitaciones habituales de este tipo de estudios, que habitualmente se centran en contextos locales o nacionales. La investigación con los mismos materiales elaborados en RV en diferentes países (o en zonas muy alejadas del mismo país) resulta ciertamente compleja, de ahí la ausencia de este tipo de estudios. La limitación de la muestra a un solo país podría condicionar los resultados obtenidos. La literatura científica evidencia que las diferencias culturales influyen significativamente en las creencias, actitudes e intenciones individuales hacia la tecnología, lo que podría repercutir profundamente en los resultados de las investigaciones (Ateş y Kölemen, 2024). La posible variación de los datos en los distintos contextos culturales subraya la necesidad de tener en cuenta estas diferencias a la hora de interpretar y analizar los hallazgos obtenidos (Bedenlier et al., 2020; Lynch et al., 2024). En el mundo actual, donde la tecnología trasciende las fronteras geográficas, cada vez es más importante tener en cuenta el contexto cultural en el que se realizan los estudios (Arpaci et al., 2020). De ahí la necesidad de poner nuestros resultados a la luz de los obtenidos en otros países con mayor y menor penetración de la RV a fin de medir el efecto de los factores contextuales en las actitudes hacia la aplicación de la RV en entornos educativos. Otras futuras vías de investigación relevantes son la aplicación de estudios longitudinales replicando los mismos instrumentos aquí utilizados y, sobre todo, la triangulación de los datos cuantitativos con técnicas cualitativas que, sin duda, podrían profundizar y matizar los resultados aquí obtenidos.

Financiación

Esta investigación está financiada por el proyecto VIRION (Virtual Reality Applied to School Education) cofinanciado por el Programa *Erasmus +* de la Comisión Europea con número de referencia ESO1-KA220-SCH-000089414.

REFERENCIAS

- Aguilera, D., García-Yeguas, A., Perales Palacios, F. y Vélchez-González, J. (2022). Diseño y validación de una rúbrica para la evaluación de propuestas didácticas STEM (RUBESTEM). *Revista Interuniversitaria de Formación del Profesorado*, 97(36.1), 11-34. <https://doi.org/10.47553/rifop.v97i36.1.92409>
- Alsalleeh, F., Okazaki, K., Alkahtany, S., Alrwais, F., Bendahmash, M. y Al Sadhan, R. (2024). Augmented reality improved knowledge and efficiency of root canal anatomy learning: A comparative study. *Applied Sciences*, 14(15), 6813. <https://doi.org/10.3390/app14156813>
- An, F., Xi, L. y Yu, J. (2024). The relationship between technology acceptance and self-regulated learning: The mediation roles of intrinsic motivation and learning engagement. *Education and Information Technologies*, 29(3), 2605-2623. <https://doi.org/10.1007/s10639-023-11959-3>
- Arce, E., Zayas-Gato, F., Suárez-García, A., Michelena, Á., Jove, E., Castelleiro-Roca,

- J.-L., Quintián, H. y Calvo-Rolle, J. L. (2022). Experiencia blended learning apoyada en un laboratorio virtual para educación de materias STEM. *Bordón. Revista de Pedagogía*, 74(4), 125-143. <https://doi.org/10.13042/Bordon.2022.95592>
- Archer, L., Moote, J., MacLeod, E., Francis, B. y DeWitt, J. (2020). *ASPIRES 2: Young people's science and career aspirations, age 10-19*. UCL Institute of Education.
- Arpaci, I., Al-Emran, M. y Al-Sharafi, M. A. (2020). The impact of knowledge management practices on the acceptance of massive open online courses (MOOCs) by engineering students: A cross-cultural comparison. *Telematics and Informatics*, 54, 101468. <https://doi.org/10.1016/j.tele.2020.101468>
- Ateş, H. y Kölemen, C. Ş. (2024). Integrating theories for insight: An amalgamated model for gamified virtual reality adoption by science teachers. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-024-12892-9>
- Bedenlier, S., Bond, M., Buntins, K., Zawacki-Richter, O. y Kerres, M. (2020). Facilitating student engagement through educational technology in higher education: A systematic review in the field of arts and humanities. *Australasian Journal of Educational Technology*, 36(4), 126-150. <https://doi.org/10.14742/ajet.5477>
- Bermejo, B., Juiz, C., Cortes, D., Oskam, J., Moilanen, T., Loijas, J., Govender, P., Hussey, J., Schmidt, A. L. y Burbach, R. (2023). AR/VR teaching-learning experiences in higher education institutions (HEI): A systematic literature review. *Informatics*, 10(2), 45. <https://doi.org/10.3390/informatics10020045>
- Boel, C., Rotsaert, T., Valcke, M., Rosseel, Y., Struyf, D. y Schellens, T. (2023). Are teachers ready to immerse? Acceptance of mobile immersive virtual reality in secondary education teachers. *Research in Learning Technology*, 31. <https://doi.org/10.25304/rlt.v31.2855>
- Cabero, J., Fernández, B. y Marín, V. (2017). Dispositivos móviles y realidad aumentada en el aprendizaje del alumnado universitario. *RIED-revista Iberoamericana de Educación a Distancia*, 20(2). <https://doi.org/10.5944/ried.20.2.17245>
- Cabero-Almenara, J., De-La-Portilla-De-Juan, F., Barroso-Osuna, J. y Palacios-Rodríguez, A. (2023). Technology-enhanced learning in health sciences: Improving the motivation and performance of medical students with immersive reality. *Applied Sciences*, 13(14), 8420. <https://doi.org/10.3390/app13148420>
- Cardoso-Júnior, A. y Delbone de Faria, R. M. (2021). Psychometric assessment of the instructional materials motivation survey (IMMS) instrument in a remote learning environment. *Revista Brasileira de Educação Médica*, 45(4), e197. <https://doi.org/10.1590/1981-5271v45.4-20210066.ing>
- Castañó-Calle, R., Jiménez-Vivas, A., Poy Castro, R., Calvo Álvarez, M. I. y Jenaro, C. (2022). Perceived benefits of future teachers on the usefulness of virtual and augmented reality in the teaching-learning process. *Education Sciences*, 12(12), 855. <https://doi.org/10.3390/educsci12120855>
- Castellano, M. S., Contreras-McKay, I., Neyem, A., Farfán, E., Inzunza, O., Ottone, N. E., del Sol, M., Alario-Hoyos, C., Alvarado, M. S. y Tubbs, R. S. (2023). Empowering human anatomy education through gamification and artificial intelligence: An innovative approach to knowledge appropriation. *Clinical Anatomy*, 36(8), 1171-1183. <https://doi.org/10.1002/ca.24074>

- Castro-Alonso, J. C., Wong, M. y Adesope, O. O. (2019). Gender imbalance in instructional dynamic versus static visualizations: A meta-analysis. *Educational Psychology Review*, 31(2), 361-387. <https://doi.org/10.1007/s10648-019-09469-1>
- Cook, D. A. y Pankratz, V. S. (2024). Validation of the instructional materials motivation survey among medicine resident physicians. *MedEdPublish*, 14(58). <https://doi.org/10.12688/mep.20408.1>
- Cózar Gutiérrez, R., González-Calero Somoza, J. A., Villena Taranilla, R. y Merino Armero, J. M. (2019). Análisis de la motivación ante el uso de la realidad virtual en la enseñanza de la historia en futuros maestros. *Edutec. Revista Electrónica de Tecnología Educativa*, (68), 1-14. <https://doi.org/10.21556/edutec.2019.68.1315>
- De Moraes Rossetto, A. G., Martins, T. C., Silva, L. A., Leithardt, D. R. F., Bermejo-Gil, B. M. y Leithardt, V. R. Q. (2023). An analysis of the use of augmented reality and virtual reality as educational resources. *Computer Applications in Engineering Education*, 31(6), 2091-2105. <https://doi.org/10.1002/cae.22671>
- Domínguez-Lara, S. (2017). Magnitud del efecto en análisis de regresión. *Interacciones*, 3(1), 3-5. <https://doi.org/10.24016/2017.v3n1.46>
- Fuentes Hurtado, M. y González Martínez, J. (2017). Necesidades formativas del profesorado de secundaria para la implementación de experiencias gamificadas en STEM. *Revista de Educación a Distancia (RED)*, 17(54). <https://doi.org/10.6018/red/54/8>
- Giakoni-Ramírez, F., Godoy-Cumillaf, A., Fuentes-Merino, P., Farías-Valenzuela, C., Duclos-Bastias, D., Bruneau-Chávez, J., Merellano-Navarro, E. y Velásquez-Olavarria, R. (2023). Intensity of a physical exercise programme executed through immersive virtual reality. *Healthcare*, 11(17), 2399. <https://doi.org/10.3390/healthcare11172399>
- Halpern, D. F. (2000). *Sex differences in cognitive abilities*. Psychology Press. <https://doi.org/10.4324/9781410605290>
- Heo, M. y Toomey, N. (2020). Learning with multimedia: The effects of gender, type of multimedia learning resources, and spatial ability. *Computers & Education*, 146, 103747. <https://doi.org/10.1016/j.compedu.2019.103747>
- Huang, B. y Hew, K. F. (2016). Measuring learners' motivation level in massive open online courses. *International Journal of Information and Education Technology*, 6(10), 759-764. <https://doi.org/10.7763/IJJET.2016.V6.788>
- Huett, J. B., Moller, L., Young, J., Bray, M. y Huett, K. C. (2008). The effect of ARCS-based strategies on confidence and performance. *The Quarterly Review of Distance Education*, 9(2), 113-126.
- Jian, Y. y Abu Bakar, J. A. (2024). Comparing cognitive load in learning spatial ability: Immersive learning environment vs. digital learning media. *Discover Sustainability*, 5(1), 111. <https://doi.org/10.1007/s43621-024-00310-6>
- Keller, J. M. (1987). Development and use of the ARCS model of instructional design. *Journal of Instructional Development*, 10(3), 2-10. <https://doi.org/10.1007/BF02905780>
- Keller, J. M. (2010). *Motivational design for learning and performance: The ARCS model approach*. Springer. <https://doi.org/10.1007/978-1-4419-1250-3>
- Loorbach, N., Peters, O., Karreman, J. y Steehouder, M. (2015). Validation of the instructional materials motivation survey (IMMS) in a self-directed instructional setting aimed at working with technology. *British Journal of Educational Technology*, 46(1), 204-218. <https://doi.org/10.1111/bjet.12138>

- Lucena-Antón, D., Fernández-López, J. C., Pacheco-Serrano, A. I., García-Muñoz, C. y Moral-Muñoz, J. A. (2022). Virtual and augmented reality versus traditional methods for teaching physiotherapy: A systematic review. *European Journal of Investigation in Health Psychology and Education*, 12(12), 1780-1792. <https://doi.org/10.3390/ejihpe12120125>
- Lynch, P., Singal, N. y Francis, G. A. (2024). Educational technology for learners with disabilities in primary school settings in low- and middle-income countries: A systematic literature review. *Educational Review*, 76(2), 405-431. <https://doi.org/10.1080/00131911.2022.2035685>
- Marín-Díaz, V. y Sampredo-Requena, B. E. (2023). Views of secondary education teachers on the use of mixed reality. *Frontiers in Education*, 7. <https://doi.org/10.3389/feduc.2022.1035003>
- Martín Carrasquilla, O., Santaolalla Pascual, E. y Muñoz San Roque, I. (2022). La brecha de género en la educación STEM. *Revista de Educación*, (396), 15-39. <https://doi.org/10.4438/1988-592X-RE-2022-396-533>
- Marrero Galván, J. J. y Hernández Padrón, M. (2022). La trascendencia de la realidad virtual en la educación STEM: Una revisión sistemática desde el punto de vista de la experimentación en el aula. *Bordón. Revista de Pedagogía*, 74(4), 45-63. <https://doi.org/10.13042/Bordon.2022.94179>
- Mercado, A. E., Sánchez, E. y Rodríguez, A. V. (2019). Estrategias de motivación en ambientes virtuales para el autoaprendizaje en matemáticas. *Revista Espacios*, 40(12), 14-21.
- Morales Inga, S. y Morales Tristán, O. (2020). ¿Por qué hay pocas mujeres científicas? Una revisión de literatura sobre la brecha de género en carreras STEM. *Revista Internacional de Investigación en Comunicación aDResearch*, (22), 118-133. <https://doi.org/10.7263/adresic-022-06>
- Palacios Ortega, A., Pascual López, V. y Moreno Mediavilla, D. (2022). El papel de las nuevas tecnologías en la educación STEM. *Bordón. Revista de Pedagogía*, 74(4), 11-21. <https://doi.org/10.13042/Bordon.2022.96550>
- Pan, J., Liu, Y. J. y Yang, C. (2024). Comparison of intelligent virtual reality first-aid training outcomes among individuals with different demographic characteristics. *Journal of Evaluation in Clinical Practice*, 30(3), 723-729. <https://doi.org/10.1111/jep.14135>
- Pontes Pedrajas, A. (2022). Uso didáctico de un laboratorio virtual para favorecer la progresión de los modelos mentales de los estudiantes sobre circuitos de corriente eléctrica. *Bordón. Revista de Pedagogía*, 74(4), 145-160. <https://doi.org/10.13042/Bordon.2022.93290>
- Prensky, M. (2009). H. sapiens digital: From digital immigrants and digital natives to digital wisdom. *Journal of Online Education*, 5(3).
- Roach, V. A., Mi, M., Mussell, J., Van Nuland, S. E., Lufner, R. S., DeVea, K. M., Dunham, S. M., Husmann, P., Herriott, H. L., Edwards, D. N., Doubleday, A. F., Wilson, B. M. y Wilson, A. B. (2021). Correlating spatial ability with anatomy assessment performance: A meta-analysis. *Anatomical Sciences Education*, 14(3), 317-329. <https://doi.org/10.1002/ase.2029>
- Saha, S. y Halder, S. (2016). He or she: Does gender affect various modes of instructional visual design? *Journal of Research on Women and Gender*, 7(1), 47-58.
- Spiriev, T., Mitev, A., Stoykov, V., Dimitrov, N., Maslarski, I. y Nakov, V. (2022). Three-dimensional immersive photorealistic layered dissection of superficial and deep back muscles: Anatomical study. *Cureus*, 14(7), e26727. <https://doi.org/10.7759/cureus.26727>

Tatiana Cox, F., González, D., Magreñán, Ángel A. y Orcos, L. (2022). Enseñanza de estadística descriptiva mediante el uso de simuladores y laboratorios virtuales en

la etapa universitaria. *Bordón. Revista de Pedagogía*, 74(4), 103-123. <https://doi.org/10.13042/Bordon.2022.94121>

Fecha de recepción del artículo: 1 de diciembre de 2024

Fecha de aceptación del artículo: 4 de marzo de 2025

Fecha de aprobación para maquetación: 18 de marzo de 2025

Fecha de publicación en OnlineFirst: 3 de abril de 2025

Fecha de publicación: 1 de julio de 2025

Impacto de variables sociodemográficas en la percepción del alumnado universitario sobre la educación en línea post-COVID-19

Impact of socio-demographic variables on university students' perception of online education post-COVID-19



- Laura Alexandra Arteaga-Briones - *Universidad Técnica de Manabí, UTM (Ecuador)*
- Luis Ángel Guamán-Lazo - *Universidad Técnica de Manabí, UTM (Ecuador)*
- Andrea Alexandra Azúa-Arteaga - *Universidad Técnica de Manabí, UTM (Ecuador)*
- Alexander López-Padrón - *Universidad Técnica de Manabí, UTM (Ecuador)*

RESUMEN

La transformación digital de la sociedad actual, sumada a los impactos de la COVID-19, han generado preferencias del alumnado universitario hacia la educación en línea, emergiendo la necesidad de reconocer los factores que influyen en su eficacia y calidad percibida. El objetivo del estudio fue determinar el efecto de las variables sociodemográficas nivel académico, sexo, área de residencia y edad en la percepción del alumnado de carreras en línea de la Universidad Técnica de Manabí sobre los componentes “organización pedagógica, tecnológica y social del proceso de enseñanza aprendizaje en línea” (CP1), “logros de aprendizaje” (CP2), “evaluación y retroalimentación” (CP3), y el “diseño del curso o la asignatura” (CP4) en el contexto Post-COVID-19. Se desarrolló una investigación de tipo descriptivo-correlacional con enfoque cuantitativo, donde 545 estudiantes respondieron un cuestionario adaptado y validado para el contexto ecuatoriano. Los resultados del *PERMANOVA* y los efectos por pares evidenciaron el impacto estadísticamente significativo de las variables sociodemográficas como nivel académico, sexo y edad en las percepciones estudiantiles sobre los componentes principales de la educación en línea. Específicamente, el nivel académico influyó en el CP1 y el CP3, mientras el sexo afectó el CP4. Por último, el principal hallazgo de este estudio evidenció que la edad impactó en los cuatro componentes, demostrando que estudiantes de 26 años o más poseen una percepción más favorable, en comparación con aquellos de menor edad. Los resultados obtenidos brindan una base empírica para mejorar la educación en línea en la era Post-COVID-19, haciéndola más eficaz, inclusiva, equitativa y adaptada.

Palabras clave: variables sociodemográficas; educación en línea; percepción estudiantil; ANOVA de permutación; educación superior.

ABSTRACT

The digital transformation of present-day society has, along with the impacts of COVID-19, led university students to prefer online education, and the need to recognize the factors that influence its effectiveness and perceived quality has, therefore, emerged. The objective of the present study was to determine the effect that the socio-demographic variables of academic level, gender, area of residence and age have on the perceptions of online Bachelor's degree students at the Technical University of Manabí with regards to the components “pedagogical, technological and social organization of the online teaching-learning process” (PC1), “learning achievements” (PC2), “assessment and feedback” (PC3), and the “design of the course or subject” (PC4) in the Post-COVID-19 context. Descriptive-correlational-type research with a quantitative focus was developed, in which 545 students responded to a questionnaire adapted to and validated for the context of Ecuador. The results of the *PERMANOVA* and the pairwise effects evidenced the statistically significant impact of socio-demographic variables such as academic level, gender and age on the students' perceptions of the principle components of online education. Specifically, academic level had an influence on PC1 and PC3, while gender affected PC4. Finally, the main finding of this study was that students of 26 years of age and over have a more favorable perception when compared to that of those who are younger. The results obtained provide an empirical basis on which to improve online education in the Post-COVID-19 era, making it more effective, inclusive, equitable and adapted.

Keywords: sociodemographic variables; online education; student perceptions; ANOVA permutation; higher education.

INTRODUCCIÓN

La transformación digital de la sociedad contemporánea ha promovido modalidades de formación en línea, flexibles y con acceso a variedad de recursos y actividades (Pelletier et al., 2023), incrementando la demanda de una educación a distancia virtual por el alumnado universitario (Ramírez García et al., 2021).

Lo antes expresado, unido a la prioridad surgida con la COVID-19 de implementar una educación en línea eficaz y atractiva adaptada a las preferencias de los estudiantes (Wang et al., 2022; Yan et al., 2022), suscitó el desarrollo de investigaciones sobre la educación mediada por tecnologías digitales a nivel mundial. En Iberoamérica, los países que lideraron estas investigaciones fueron España, Ecuador, Colombia, Perú, México, Venezuela y Chile. La producción científica enfatizó temas relacionados con la enseñanza, el aprendizaje, el rendimiento académico, la motivación y la satisfacción del alumnado con la educación (Cruz, 2023) y los servicios universitarios (Colvin et al., 2024).

La educación mediada por tecnologías digitales ha sido abordada desde diferentes enfoques y perspectivas en la educación superior, asignándole diversas denominaciones en la literatura publicada. Entre las más usadas se destacan “educación a distancia” (Carbonell et al., 2021; García Aretio, 2020), “educación en línea” (Arslan et al., 2023; Colvin et al., 2024; Yan et al., 2022) y “educación virtual” (Crisol-Moya et al., 2020; Gómez-Arteta et al., 2024; Ramírez García et al., 2021). Sin embargo, a los efectos del presente estudio, se asume la denominación “educación en línea” para abordar el proceso educativo de carácter no presencial a distancia, que se desarrolla como un diálogo didáctico mediado por conexión telemática, en el cual docentes y estudiantes no comparten un lugar físico para la construcción de nuevos aprendizajes de forma independiente o grupal (García Aretio, 2004, 2020).

Esta modalidad de estudio integra el uso intensivo de tecnologías interactivas multimedia, entornos virtuales de aprendizaje y recursos educativos digitales (Consejo de Educación Superior, 2023; Crisol-Moya et al., 2020), con estrategias didácticas adecuadamente planificadas, ejecutadas y autorreguladas mediante controles (Rivera Tejada et al., 2023). Los resultados de dicha integración potencian la flexibilidad espacio temporal del estudio, la autonomía y autodirección para el aprendizaje, la accesibilidad a la diversidad de recursos educativos digitales y fuentes de información las 24 horas del día, el ahorro de tiempo y dinero, el impacto ambiental positivo por menor uso de transporte e infraestructura, el alcance en cuanto a cantidad de estudiantes, la motivación, la comunicación, el trabajo colaborativo y la innovación en la evaluación de los aprendizajes (Arslan et al., 2023; Barrientos et al., 2022; Cramarenco et al., 2023; Dinu et al., 2022).

A pesar de todas las bondades y ventajas antes descritas, existen retos y limitaciones a superar para el desarrollo de una educación en línea de calidad. Entre ellos se destacan los problemas de conectividad a Internet, la insuficiente interactividad profesor-alumno y alumno-alumno, la ausencia de equipos e

infraestructura tecnológica y recursos educativos para la formación, el escaso desarrollo de competencias digitales en docentes y estudiantes, el alto costo de la infraestructura que amplía la brecha digital y la desigualdad educativa y socioeconómica, la falta de credibilidad en la evaluación y la eficacia de la modalidad formativa, y los problemas de concentración para el aprendizaje (Amaning, 2024; Cabero-Almenara et al., 2023; Carbonell et al., 2021; Martín-Párraga et al., 2024; Vásquez Peñafiel et al., 2023).

Un criterio esencial para la evaluación de la calidad de la educación en línea lo constituye la percepción y niveles de satisfacción del alumnado con el diseño, organización y evaluación del aprendizaje en línea (Rahmatpour et al., 2024). En el contexto latinoamericano, se han establecido dimensiones para la agrupación de las variables que contribuyen a la evaluación de la calidad de la educación en línea. Entre ellas se destacan:

- Diseño del curso o la materia, características de las clases, calidad de la enseñanza, capacidades tecnológicas y sociales del docente, aprendizaje (Avendaño et al., 2021).
- Experiencia de aprendizaje, utilidad percibida, instrumentalidad del curso, facilidad de uso, calidad del servicio (Estrada y Paz Delgado, 2022).
- Compromiso institucional, calidad de los procesos de enseñanza aprendizaje, calidad de la planificación de la enseñanza, calidad de la relación con pares, calidad de la relación con el docente, autoeficacia percibida para la educación virtual (Lobos et al., 2022).
- Servicios educativos generales, recursos tecnológicos, desempeño y actitud del docente, proceso virtual de enseñanza aprendizaje (Gómez-Arteta et al., 2024).

Como se puede apreciar en el párrafo anterior, existe una tendencia a agrupar las variables para la evaluación de la percepción y satisfacción de los estudiantes con la calidad de la educación en línea en dimensiones relacionadas con el diseño del curso o asignatura, la organización pedagógico-tecnológica y social del proceso de enseñanza aprendizaje en línea, la evaluación y retroalimentación y los logros de aprendizaje alcanzados por los estudiantes. Estos aspectos se toman en cuenta en la presente investigación, unidos al criterio de la existencia de asociación entre las percepciones del alumnado sobre la calidad de la educación en línea y las variables sociodemográficas, académicas y de contexto tales como sexo, edad, área de residencia y nivel académico (Avendaño et al., 2021; Estrada-Araoz et al., 2023).

La educación superior ecuatoriana no es ajena a esta necesidad de evaluar la calidad de la educación en línea, lo cual se evidencia en el estudio desarrollado por Remesal y Villarroel (2023), donde compararon las percepciones de docentes y estudiantes de Chile, México y Ecuador sobre aspectos pedagógicos y emocionales relacionados con la calidad de la educación en línea durante la COVID-19. En sus resultados destacan la importancia de valorar el diseño instruccional, el vínculo

docente-estudiante, las evaluaciones significativas y realistas y las oportunidades de retroalimentación en la educación en línea, así como su asociación con el sexo, lugar de residencia y nivel académico en la era Post-COVID-19.

Por lo antes expresado, el presente estudio se diseñó con la finalidad de determinar el efecto de las variables sociodemográficas nivel académico, sexo, área de residencia y edad en la percepción del alumnado de carreras en línea de la Universidad Técnica de Manabí sobre los componentes “organización pedagógica, tecnológica y social del proceso de enseñanza aprendizaje en línea” (CP1), “logros de aprendizaje” (CP2), “evaluación y retroalimentación” (CP3), y el “diseño del curso o la asignatura” (CP4) en el contexto Post-COVID-19.

METODOLOGÍA

El diseño de investigación empleado fue no experimental transversal, sustentado en un enfoque metodológico cuantitativo para el desarrollo de un estudio de tipo descriptivo-correlacional (Hernández y Mendoza, 2018).

Contexto y participantes

La investigación se llevó a cabo en el contexto de las ocho carreras en línea que ofertaba la Universidad Técnica de Manabí en 2024. La educación en línea en dichas carreras se desarrolla alineada a las directrices establecidas en el Reglamento de Régimen Académico (Consejo de Educación Superior, 2023), las cuales establecen que la totalidad de los componentes de aprendizaje estarán mediados por el uso de tecnologías interactivas multimedia y entornos virtuales de aprendizaje. Para ello, se creó un entorno virtual de enseñanza aprendizaje usando la plataforma Moodle, cuyo diseño instruccional integra los recursos educativos digitales y las actividades de aprendizaje, de forma secuencial y planificada, para el desarrollo de los componentes de aprendizaje autónomo, práctico experimental y en contacto con el docente de forma asincrónica (tutorías). El componente de aprendizaje en contacto con el docente de forma sincrónica, se implementa mediante sesiones telemáticas de una hora de duración usando la plataforma de videoconferencia Zoom.

La muestra del estudio surge de la población total de 9127 estudiantes matriculados en la oferta académica en línea para el período 2023-2024. El tamaño muestral calculado fue de 370 estudiantes, sin embargo, para el presente estudio se logró encuestar a 545. La selección de los participantes se realizó de forma aleatoria, a través de los vicedecanos de carrera, a quienes el equipo investigador remitió el link del cuestionario construido en *Google Forms*. Los criterios de inclusión fueron: ser estudiante matriculado en las carreras en línea en el período objeto de estudio y tener más de 15 años de edad, similares a los establecidos por Avendaño et al. (2021). Desde el punto de vista sociodemográfico, el 69,9 % (381) de la muestra eran del sexo masculino y el 30,1 % (164) del femenino, así como el tramo etario predominante fue

de 21-25 años con un 44,6 % (243). Los datos fueron recogidos con el consentimiento informado del alumnado, durante el mes de mayo de 2024, posterior a la aprobación del estudio por el Comité de Ética de la Universidad Técnica de Manabí (CEISH-UTM-INT_24-04-22_ALP).

La representatividad de la muestra seleccionada para las Instituciones de Educación Superior que desarrollan procesos de educación en línea en Ecuador, se sustenta tomando en cuenta tres aspectos fundamentales. El primero, está dado por la flexibilidad espacio temporal, lo cual posibilitó que las 24 provincias del país estén representadas con estudiantes matriculados en estas carreras. En segundo lugar, tomando en cuenta que la Universidad Técnica de Manabí se ubica entre las tres primeras del país con mayor oferta de carreras en línea (Senescyt, 2024), se garantiza que la proporción de estudiantes matriculados de las diversas regiones, representan de una forma más adecuada los datos a nivel de país, a pesar de ser obtenidos en una sola universidad. En tercer lugar, cabe destacar que la ubicación de la Universidad Técnica de Manabí en *QS World University Rankings* y *Times Higher Education World University Rankings*, la convierte en una institución de referencia para la educación en línea ecuatoriana.

Instrumento

El instrumento empleado fue el cuestionario diseñado por Avendaño et al. (2021) para el análisis de la educación virtual a partir de las percepciones de los estudiantes universitarios colombianos. Para la recogida de los datos en el presente estudio, el instrumento quedó estructurado en una sección de variables sociodemográficas (nivel académico, sexo, área de residencia y edad) y una sección que recoge el constructo de la percepción del alumnado sobre la calidad de la educación en línea, representada por 24 ítems distribuidos en cinco factores: “diseño del curso o materia” (5 ítems), “características de las clases” (5 ítems), “calidad de la enseñanza” (6 ítems), “capacidades tecnológicas y sociales del docente” (3 ítems) y “aprendizaje” (4 ítems). La medición de los ítems se realizó mediante la escala Likert de cinco opciones de respuestas establecidas para el cuestionario original, asociando 1= muy en desacuerdo y 5= muy de acuerdo.

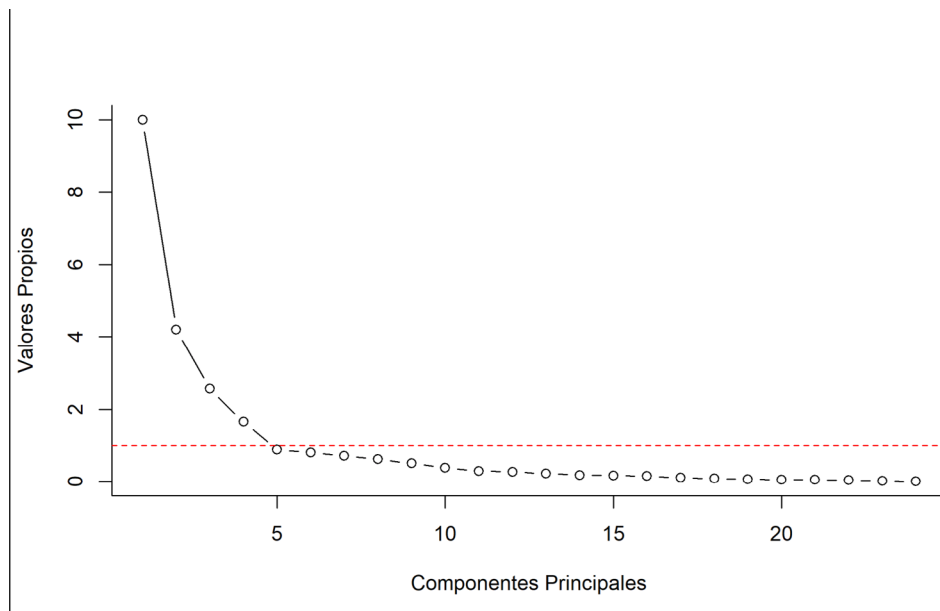
Validez y fiabilidad del instrumento para el contexto ecuatoriano

A pesar de que el instrumento fue validado con estudiantes universitarios, mediante juicio de expertos y un pilotaje que arrojó una alta fiabilidad y validez (Avendaño et al., 2021), considerando que la estructura de la educación en línea en Ecuador podría diferir de la colombiana, se procedió a validar el instrumento mediante el método de análisis de componentes principales (ACP). El índice KMO fue de 0,97 superando el valor de 0,6 y la esfericidad de Bartlett sobre la matriz de correlaciones policórica del instrumento fue significativa con un p-valor < 0.05. En

consecuencia, el ACP se realizó considerando la rotación Varimax con la finalidad de maximizar la varianza de los elementos cuadrados de cada factor. Como resultado, siguiendo el criterio de Kaiser, se concluyó que el número de componentes principales son cuatro (Figura 1).

Figura 1

Número de componentes principales con valores propios > 1



Fuente: elaboración propia.

De la construcción de los componentes principales, se puede observar que la contribución que cada ítem tiene sobre sus componentes es superior a 0,50, además, la primera componente está representada por el 63 % de los ítems, seguido por la segunda componente con el 21 % y el restante 16 % lo representan las componentes 3 y 4, en partes iguales (Tabla 1).

Tabla 1*Máximas cargas factoriales de los componentes principales, por ítems*

Ítem	CP1	CP2	CP3	CP4
Organización y secuencias				0,78334
Ritmos				0,67948
Nivel de exigencia académica			0,73685	
Compromisos / tareas			0,74937	
Bibliografía	0,56674			
Explicaciones del docente	0,75840			
Interés del estudiante	0,76850			
Motivación	0,74320			
Participación	0,65253			
Recursos y medios	0,68096			
Actitud docente	0,71018			
Empatía	0,70494			
Contenidos e información	0,71847			
Procedimientos organizados	0,65782			
Discurso docente	0,72809			
Manejo general de la clase	0,63373			
Manejo de plataformas	0,60548			
Uso de redes	0,63920			
Recursos virtuales y digitales	0,58784			
Autonomía del estudiante		0,71390		
Aprendizaje efectivo		0,77202		
Habilidades y competencias desarrolladas		0,77767		
Emociones		0,73844		
Satisfacción general		0,71265		
Proporción	63 %	21 %	8 %	8 %

Fuente: elaboración propia

Por lo tanto, se redujo la dimensionalidad a cuatro componentes principales, en correspondencia a las características de los ítems que los representan, denominándose estos: “organización pedagógica, tecnológica y social del proceso de enseñanza aprendizaje en línea” (CP1), “logros de aprendizaje” (CP2), “evaluación y retroalimentación” (CP3) y “diseño del curso o la asignatura” (CP4). Además, se

identificó la existencia de *outliers* que se los imputó a los extremos que congregaban el 50 % de los datos alrededor de la mediana del diagrama de cajas, con la finalidad de no perder observaciones de la muestra.

La consistencia del instrumento se validó usando los coeficientes Alfa de Cronbach (α) y Omega de Macdonald (ω). Los resultados de α para cada factor fueron: 0,97, 0,93, 0,78 y 0,70 respectivamente; y los de ω fueron: 0,97 y 0,94 para los factores CP1 y CP2 respectivamente. Para los factores CP3 y CP4 no se presentan valores de ω , ya que este coeficiente exige un mínimo de tres ítems por factor para su determinación. Al CP3 y CP4 contener dos ítems cada uno, su fiabilidad se evaluó solamente aplicando el α (Roco-Videla et al., 2024). Los hallazgos señalan una fiabilidad que oscila entre aceptable y alta para los factores examinados.

Análisis de los datos

Los datos fueron procesados mediante análisis estadísticos descriptivos e inferenciales utilizando el software RStudio versión 4.3.2.

Para evaluar la influencia de las variables sociodemográficas (independientes) sobre los cuatro componentes principales extraídos (dependientes), se aplicó un ANOVA de permutaciones, también conocido como *permutational multivariate analysis of variance* (*PERMANOVA*). Este enfoque no paramétrico fue elegido debido a que no se cumplen completamente las suposiciones de normalidad requeridas por el ANOVA tradicional. Las pruebas de Shapiro-Wilk y Jarque-Bera corroboraron dicha elección, arrojando evidencia suficiente para rechazar la hipótesis nula de normalidad de los datos, al obtener un p-valor menor a 0,05 (Harpe, 2015; Norman, 2010; Sullivan y Artino, 2013).

El *PERMANOVA* es una prueba robusta frente a violaciones de supuestos, como la no normalidad, y es adecuada para analizar datos complejos. Además, es flexible para analizar el efecto de múltiples variables cualitativas independientes e, incluso, considerar sus interacciones dentro del modelo cuando sea pertinente (Anderson, 2001; Good, 2000; Higgins, 2004).

Es importante considerar que para la estimación del *PERMANOVA* se realizaron 5000 permutaciones para generar una distribución nula bajo la hipótesis de no influencia de las variables sociodemográficas en los componentes principales. El p-valor resultante de cada análisis permitió evaluar si las diferencias observadas en los componentes principales entre los diferentes niveles de las variables sociodemográficas eran significativas.

Por lo tanto, en el análisis de varianza no paramétrico es fundamental que en una primera instancia se genere la selección de un modelo óptimo para asegurar interpretaciones robustas y prevenir la sobreparametrización. Esto se realiza mediante la utilización de criterios de información como el Criterio de Información de Akaike y el Criterio de Información Bayesiano (*AIC* y *BIC* por sus siglas en inglés,

respectivamente), los cuales son ampliamente utilizados en el modelado estadístico (Burnham y Anderson, 2004).

Teniendo en cuenta lo antes planteado, se estimaron los *PERMANOVA* para cada variable dependiente con todas las variables independientes antes descritas, considerando que no existe y que existe interacción entre ellas, y consecuentemente, se evaluaron los criterios *AIC* y *BIC* para determinar qué modelo se ajusta mejor a cada componente principal. En la Tabla 2 se puede observar que, en los componentes principales, los modelos sin interacciones presentan de forma consistente los valores más bajos de *AIC* y *BIC*. Esto indica que incorporar interacciones en los modelos genera una complejidad innecesaria sin representar mejora significativa en su ajuste.

Tabla 2

Criterios AIC y BIC para los modelos de cada componente principal sobre las variables sociodemográficas, considerando sin y con interacciones entre ellas

Criterios / Modelo	CP1		CP2		CP3		CP4	
	Sin	Con	Sin	Con	Sin	Con	Sin	Con
<i>AIC</i>	1444,30	1476,04	1346,21	1382,68	1421,44	1439,17	1406,69	1444,13
<i>BIC</i>	1500,61	1675,29	1402,52	1581,93	1477,75	1638,42	1463,00	1643,38

Fuente: elaboración propia.

Dado que el *PERMANOVA* sin interacciones fue superior, se debe evaluar los criterios *AIC* y *BIC* para definir qué variables sociodemográficas se ajustan mejor a los modelos. Por lo tanto, se especificó de manera general el *PERMANOVA* de la siguiente manera:

$$Y_{ijk} = \mu + A_i + B_j + C_k + D_l + e_{ijkl}$$

Donde:

Y_{ijk} representa el valor del componente principal.

μ es la media global.

A_i es el efecto del nivel de educación.

B_j es el efecto del sexo.

C_k es el efecto del área de residencia.

D_l es el efecto de la edad.

e_{ijkl} es el término de error.

Para determinar cuáles de las cuatro variables sociodemográficas se ajustan mejor a los modelos, se estimaron todas las combinaciones posibles de *PERMANOVA*

para cada componente principal, obteniendo un total de 60 modelos (15 por cada componente). Para facilitar el análisis, en la Tabla 3 se presentan los modelos óptimos para cada componente principal estudiado.

Tabla 3

Criterios AIC y BIC para los modelos de cada componente principal sobre las variables sociodemográficas relevantes, sin interacciones entre ellas

Componentes	Modelo	AIC	BIC
CP1	CP1 ~ Nivel académico + Sexo + Edad	1443,41	1495,39
CP2	CP2 ~ Sexo + Edad	1344,92	1370,91
CP3	CP3 ~ Nivel académico + Edad	1417,84	1465,49
CP4	CP4 ~ Sexo + Área de residencia + Edad	1402,09	1432,41

Fuente: elaboración propia.

Adicionalmente, se llevó a cabo un análisis de efectos por pares (*pairwise permutation*) para identificar qué grupos específicos dentro de cada variable sociodemográfica mostraban diferencias significativas en los componentes principales. Este análisis se realizó utilizando el método de permutación por pares, comparando individualmente todos los niveles de las variables independientes (por ejemplo, diferentes niveles académicos entre sí). Como los datos no cumplen el supuesto de normalidad, esta metodología es ideal (Ernst, 2004).

Dado al elevado número de comparaciones entre los niveles de las variables independientes (nivel académico, sexo, área de residencia y edad), se empleó el ajuste de *False Discovery Rate (FDR)* para controlar el error de tipo I y evitar falsos positivos en las pruebas múltiples (Benjamini y Hochberg, 1995; García y Herrera, 2008).

Para cada par de combinaciones, se realizaron 5000 permutaciones con el fin de obtener una distribución nula y estimar los p-valores correspondientes. De este modo, se identificaron qué combinaciones de niveles presentaban diferencias significativas en los componentes principales, proporcionando una evaluación detallada de las variaciones entre pares de cada variable sociodemográfica.

RESULTADOS

En esta sección se presentan los resultados estadísticamente significativos y más importantes de los modelos propuestos. Se divide en dos acápites: el primero describe los resultados del *PERMANOVA*, mientras que el segundo presenta los resultados del efecto pares de permutaciones. Ambos enfoques permiten evaluar la influencia

de las variables sociodemográficas (nivel académico, sexo, área de residencia y edad) sobre los componentes principales obtenidos a partir del cuestionario.

Resultados del PERMANOVA

Los resultados del método *PERMANOVA* se muestran en la Tabla 4. Los mismos demuestran que, bajo un nivel de significancia del 5 %, la variable nivel académico tiene efecto significativo sobre la percepción de los estudiantes en la “organización pedagógica, tecnológica y social del proceso de enseñanza aprendizaje en línea” y la “evaluación y retroalimentación”; mientras la variable sexo tiene efecto significativo sobre la percepción del “diseño del curso o la asignatura”.

Tabla 4

P-valores de las variables independientes por modelo

Modelo	Nivel académico	Sexo	Área de residencia	Edad
CP1	0.002 **	0.17		0 ***
CP2		0.0848		0 ***
CP3	0.0318 *			0.0018 **
CP4		0.016 *	0.2213	0 ***

Nota: Códigos de significancia: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘.’

Fuente: elaboración propia

Por su parte, la variable edad muestra efecto significativo sobre todos los componentes analizados. En consecuencia, uno de los hallazgos más relevantes de este estudio es la evidencia de efectos marginales positivos en la percepción de los estudiantes de 26 años o más, en comparación con aquellos de menor edad, como se observa en las Figuras 2, 3, 4 y 5.

Los efectos marginales estimados para las variables nivel académico, sexo y área de residencia, se presentan a continuación por cada componente. Estos resultados evidencian posibles efectos significativos que serán confirmados en el siguiente acápite del apartado de resultados.

Componente 1: Percepción de la “organización pedagógica, tecnológica y social del proceso de enseñanza aprendizaje en línea”

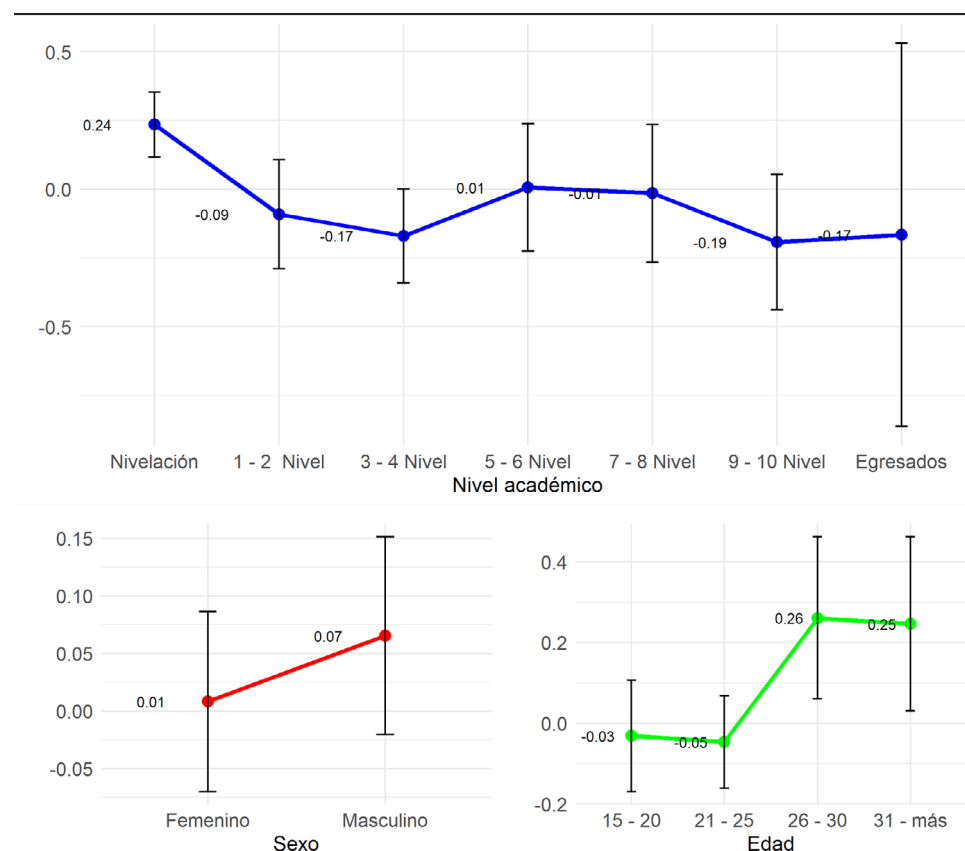
Cuando se analizan los efectos marginales de cada factor sobre la “organización pedagógica, tecnológica y social del proceso de enseñanza aprendizaje en línea”, se puede observar en la Figura 2, aunque no de manera concluyente, que la percepción que tienen los estudiantes que están en el nivel académico nivelación, es mayor y

positiva comparada con la de los niveles posteriores a este. Aún no se puede tener claridad de si todos estos efectos son significativos, ya que existen intervalos de confianza de los factores que se traslapan.

Por otro lado, la variable sexo se mantiene bastante pareja entre los factores que la representan, evidenciando además que sus intervalos de confianza se superponen, razón por la cual no es significativa en el *PERMANOVA*.

Figura 2

Efectos marginales de las variables sociodemográficas sobre la percepción de la “organización pedagógica, tecnológica y social del proceso de enseñanza aprendizaje en línea”



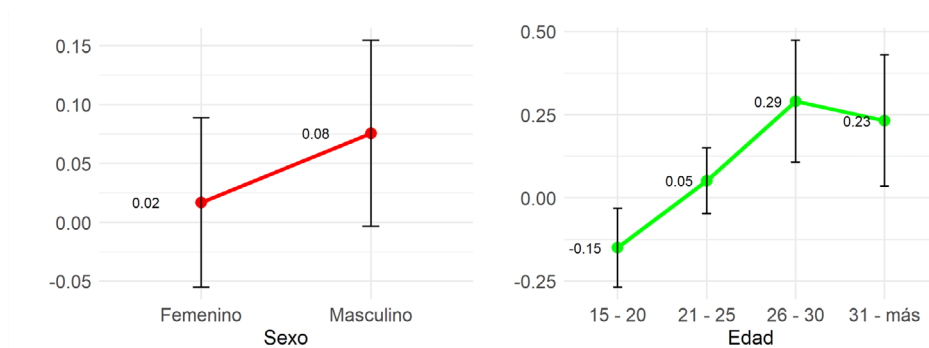
Fuente: elaboración propia.

Componente 2: Percepción de los “logros de aprendizaje”

En la Figura 3, se observa que la percepción que tienen los estudiantes del sexo masculino sobre los “logros de aprendizaje” es mejor que la del femenino, aunque al comparar entre pares podría ser no significativo, porque sus intervalos de confianza se solapan parcialmente.

Figura 3

Efectos marginales de las variables sociodemográficas sobre la percepción de los “logros de aprendizaje”



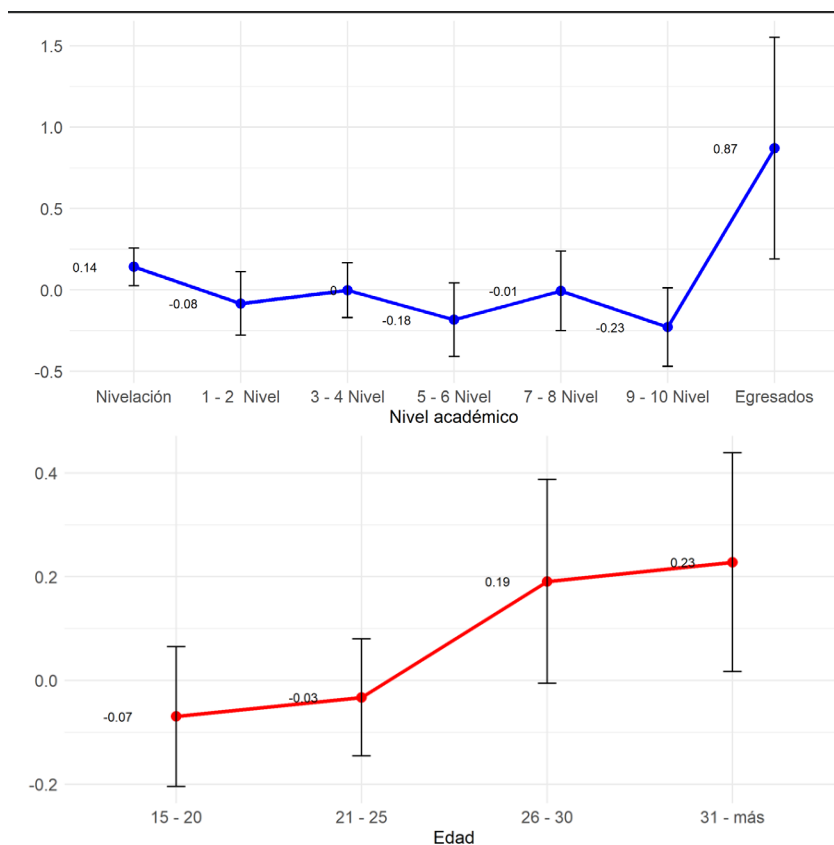
Fuente: elaboración propia.

Componente 3: Percepción de la “evaluación y retroalimentación”

En la Figura 4, aunque de manera no concluyente, se observa que los estudiantes con nivel académico egresado tienen una percepción positiva y presumiblemente significativa sobre la “evaluación y retroalimentación” en comparación con los restantes niveles académicos, ya que sus intervalos de confianza no se traslapan con ningún otro nivel académico. Así también, se puede observar que los estudiantes en el nivel académico de nivelación presentan efectos relativamente significativos por sobre los niveles de 1 – 2, 5 – 6 y 9 – 10.

Figura 4

Efectos marginales de las variables sociodemográficas sobre la percepción de la “evaluación y retroalimentación”



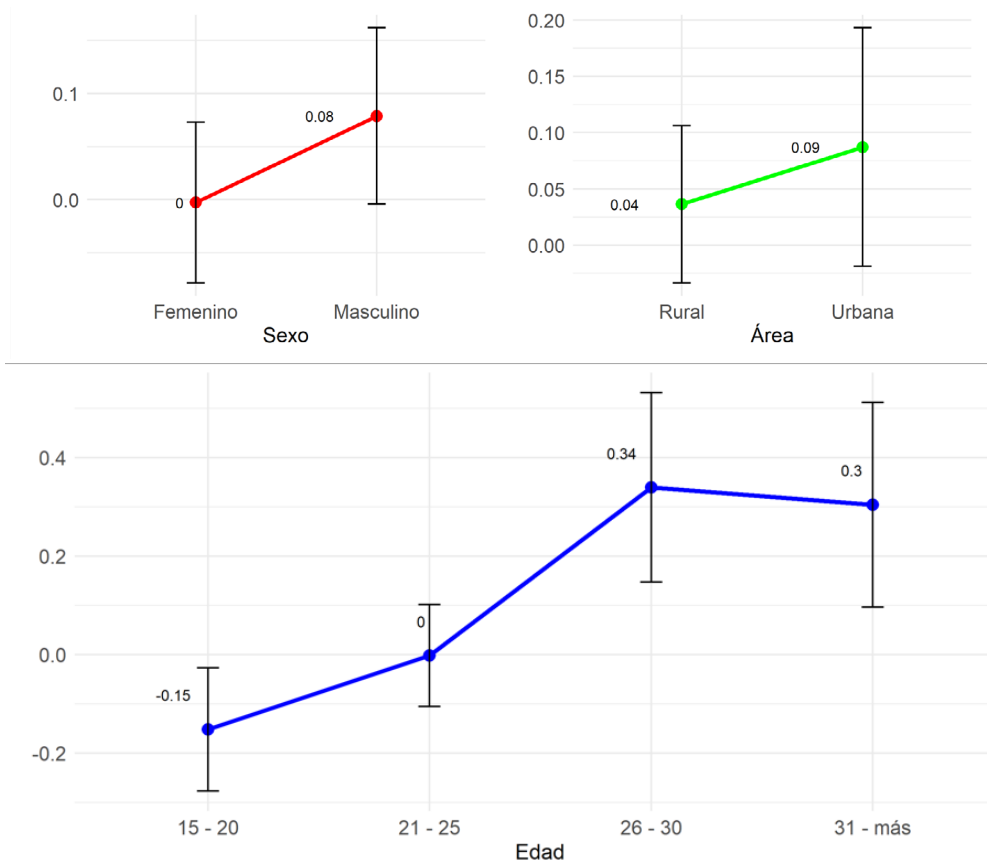
Fuente: elaboración propia.

Componente 4: Percepción del “diseño del curso o la asignatura”

Por último, de la percepción del “diseño del curso o la asignatura”, se puede observar en la Figura 5, que tan solo los efectos obtenidos de la variable sexo no se traslapan totalmente en sus intervalos de confianza, y aunque el *PERMANOVA* declara que esta variable presenta diferencias significativas entre los grupos, no es tan claro que esto sea así. Sin embargo, la variable área de residencia se traslapa significativamente entre los intervalos de confianza de cada factor, razón por la cual no son significativos.

Figura 5

Efectos marginales de las variables sociodemográficas sobre la percepción del “diseño del curso o la asignatura”



Fuente: elaboración propia.

Resultados de efectos pares de permutaciones

A continuación, se presentan los resultados obtenidos de los pares que fueron significativos bajo 5 % de significancia, rechazando así la hipótesis nula de que no existía diferencia sobre los promedios de las variables dependientes analizadas dada ciertas características sociodemográficas del alumnado encuestado.

Componente 1: “organización pedagógica, tecnológica y social del proceso de enseñanza aprendizaje en línea”

En la Tabla 5, como resultado de la comparación de los pares de nivel académico se obtuvo que la magnitud de la diferencia promedio en la percepción de los estudiantes sobre la “organización pedagógica, tecnológica y social del proceso de enseñanza aprendizaje en línea” es mayor en los estudiantes de nivelación comparados con los de niveles 1 - 2, 3 - 4 y 9 - 10, con magnitud de diferencia de 0.33, 0.41 y 0.43, respectivamente.

Por otro lado, al examinar la variable edad, se evidencia que los grupos etarios superiores a 26 años presentan una percepción más favorable en comparación con los estudiantes de 21 y 25 años.

Tabla 5

Resultados significativos al 0.05 en la percepción de la “organización pedagógica, tecnológica y social del proceso de enseñanza aprendizaje en línea”

Comparación	Magnitud de la diferencia de medias (ΔM)	Estadístico	P Valor ajustado	Significancia
<i>Nivel académico</i>				
Nivelación - 1 - 2 Nivel = 0	0.33	-3.36	0.00539	***
Nivelación - 3 - 4 Nivel = 0	0.41	-4.09	0.00046	***
Nivelación - 9 - 10 Nivel = 0	0.43	-4.10	0.00046	***
<i>Edad</i>				
[21 - 25] - [26 - 30] = 0	-0.3074568	3.36	0.00474	***
[21 - 25] - [31 - más] = 0	-0.2926446	2.87	0.01238	**

Nota: Códigos de significancia: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘.’

Fuente: elaboración propia.

Componente 2: “logros de aprendizaje”

Para el componente 2, se evaluaron los pares de la variable sexo. Aunque el *PERMANOVA* rechazó la hipótesis nula de igualdad entre grupos, el análisis por pares reveló que la diferencia entre los niveles femenino y masculino no es estadísticamente significativa ($p = 0.0735$), aunque está cerca al umbral típico ($p < 0.05$). Esto indica que la diferencia observada entre grupos podría deberse al azar, faltando suficiente evidencia para rechazar la hipótesis nula de que tienen efectos similares sobre la percepción de los “logros de aprendizaje” (Tabla 6).

No obstante, la edad revela que la diferencia de medias en el rango etario de 15 – 20 años (edad mínima admisible) es significativa en comparación con todos sus pares, además de indicar que los rangos etarios superiores presentan, en promedio, una mayor magnitud de diferencias de medias en la percepción de los “logros de aprendizaje”.

Tabla 6

Resultados significativos al 0.05 de significancia en la percepción de los “logros de aprendizaje”

Comparación	Magnitud de la diferencia de medias (ΔM)	Estadístico	P Valor ajustado	Significancia
<i>Edad</i>				
[15 - 20] - [21 - 25] = 0	-0.2011647	2.51	0.02422	**
[15 - 20] - [26 - 30] = 0	-0.4395201	3.75	0.00108	***
[15 - 20] - [31 - más] = 0	-0.3815715	3.03	0.00743	***
[21 - 25] - [26 - 30] = 0	-0.2383554	2.38	0.02577	**

Nota: Códigos de significancia: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘’

Fuente: elaboración propia.

Componente 3: “evaluación y retroalimentación”

La percepción del componente 3, al evaluar sus pares, arrojó resultados similares a los del componente 1. La magnitud de las diferencias promedio es significativa, en especial para los estudiantes en los niveles académicos de nivelación y egresados, y para los rangos etarios de 15-20 y 21-25 años.

Al analizar los resultados del post hoc del CP3 con el nivel académico, los estudiantes egresados mostraron diferencias significativas con todos sus pares. En particular, los valores significativamente superiores fueron con los niveles 1 - 2 ($\Delta M = -0.95$) y 5 - 6 ($\Delta M = -1.05$).

Respecto a la variable edad, se evidencia que los estudiantes de 26 a 30 años exhibieron valores más altos de CP3 en comparación con los más jóvenes de 15 a 20 años ($\Delta M = -0.26$), y otra tendencia análoga se observó en el grupo de 31 años o más ($\Delta M = -0.30$). Asimismo, el grupo de 21 a 25 años presentó valores significativamente inferiores en comparación con los mayores de 26 años. Los valores de las diferencias entre los pares se pueden observar en la Tabla 7.

Tabla 7

Resultados significativos al 0.05 de significancia en la percepción de la “evaluación y retroalimentación”

Comparación	Magnitud de la diferencia de medias (ΔM)	Estadístico	P Valor ajustado	Significancia
<i>Nivel académico</i>				
Nivelación - 1 - 2 Nivel = 0	0.22	-2.31	0.04898	**
Nivelación - 5 - 6 Nivel = 0	0.33	-2.59	0.03994	**
Nivelación - 9 - 10 Nivel = 0	0.37	-3.26	0.0236	**
Nivelación - Egresados = 0	-0.73	2.45	0.0429	**
1 - 2 Nivel - Egresados = 0	-0.95	2.65	0.03994	**
3 - 4 Nivel - Egresados = 0	-0.87	2.49	0.0429	**
5 - 6 Nivel - Egresados = 0	-1.05	2.91	0.031	**
7 - 8 Nivel - Egresados = 0	-0.88	2.35	0.04898	**
9 - 10 Nivel - Egresados = 0	-1.10	2.85	0.031	**
<i>Edad</i>				
[15 - 20] - [26 - 30] = 0	-0.26	2.13	0.0492	**
[15 - 20] - [31 - más] = 0	-0.30	2.15	0.0492	**
[21 - 25] - [26 - 30] = 0	-0.22	2.88	0.01304	**
[21 - 25] - [31 - más] = 0	-0.26	2.85	0.01304	**

Nota: Códigos de significancia: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘.’

Fuente: elaboración propia.

Sin duda, los resultados obtenidos son interesantes, en especial el de los Egresados, quienes han experimentado dos etapas cruciales de la educación en Ecuador: 1) Pre-COVID-19 y 2) Post-COVID-19. Su percepción positiva podría deberse a que vivieron la evolución del diseño de las clases en línea, especialmente lo concerniente a la “evaluación y retroalimentación”, así como podría estar asociado con factores de madurez.

Componente 4: “diseño del curso o la asignatura”

Por último, en el componente 4 se encontraron diferencias significativas en la percepción según el sexo y la edad, como se muestra en la Tabla 4 del análisis de *PERMANOVA*. Al comparar por pares mediante permutaciones de la variable sexo, se encontró que la diferencia de medias entre masculino y femenino es pequeña

($\Delta M = -0.08$), pero significativa (Tabla 8). Al ser la diferencia promedio ajustada negativa, esto indica que el valor promedio de la variable dependiente es menor para el sexo femenino en comparación con el sexo masculino.

En contraste, las variaciones en función de la edad son más evidentes y se alinean con los hallazgos obtenidos en los otros componentes (CP1, CP2 y CP3). Como se evidencia en la Tabla 8, los segmentos demográficos más jóvenes (15-20 y 21-25 años) suelen otorgar una valoración menos favorable al “diseño del curso o la asignatura” en comparación con los segmentos de mayor edad (26-30 y 31 o más años).

Se observan diferencias de medias más notables entre el grupo de 15-20 años y sus similares de 26-30 ($\Delta M = -0.49$), y de 31 o más ($\Delta M = -0.46$). Esto indica que los estudiantes más jóvenes podrían notar más falencias o retos en el “diseño del curso o la asignatura”, en contraste con sus pares de mayor edad, quienes probablemente poseen experiencia o estrategias de adaptación más efectivas al contexto educativo.

Tabla 8

Resultados significativos al 0.05 de significancia en la percepción en el “diseño del curso o la asignatura”

Comparación	Magnitud de la diferencia de medias (ΔM)	Estadístico	P Valor ajustado	Significancia
<i>Sexo</i>				
Femenino - Masculino = 0	-0.08	2.35	0.01888	**
<i>Edad</i>				
[15 - 20] - [26 - 30] = 0	-0.49	4.03	0.00034	***
[15 - 20] - [31 - más] = 0	-0.46	3.48	0.00151	***
[21 - 25] - [26 - 30] = 0	-0.34	3.02	0.00504	***
[21 - 25] - [31 - más] = 0	-0.31	2.49	0.01926	**

Nota: Códigos de significancia: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘.’

Fuente: elaboración propia.

DISCUSIÓN

El objetivo de la presente investigación fue determinar el efecto de las variables sociodemográficas nivel académico, sexo, área de residencia y edad en la percepción del alumnado de carreras en línea de la Universidad Técnica de Manabí sobre los componentes “organización pedagógica, tecnológica y social del proceso de enseñanza aprendizaje en línea”, “logros de aprendizaje”, “evaluación y retroalimentación”, y el

“diseño del curso o la asignatura” en el contexto Post-COVID-19. A continuación, se presenta la discusión examinando la influencia que tiene cada variable sociodemográfica significativa sobre la percepción de la educación en línea, medida sobre cada componente analizado.

El nivel académico mostró tener efecto sobre la percepción de la educación en línea del alumnado, con tendencias similares en los componentes “organización pedagógica, tecnológica y social del proceso de enseñanza-aprendizaje en línea” y “evaluación y retroalimentación”. En ambos casos, los estudiantes en nivelación valoran más positivamente la educación en línea en comparación con aquellos en niveles intermedios o avanzados. Esto coincide con los hallazgos de Remesal y Villarroel (2023), quienes sugieren que los estudiantes en etapas iniciales tienen menos referencias para comparar modalidades educativas y, por ende, muestran mayor aceptación. Además, su entusiasmo y expectativas iniciales pueden influir en esta valoración.

Por otro lado, en el CP3, los egresados también presentan percepciones significativamente más favorables que los estudiantes en niveles intermedios, lo que puede deberse a su experiencia tanto en educación Pre-COVID-19 como Post-COVID-19, permitiéndoles valorar objetivamente las mejoras implementadas en la enseñanza en línea (Dinu et al., 2022; Vaca-Cárdenas et al., 2024). La implementación de estrategias como personalización del aprendizaje, trabajo colaborativo, evaluaciones más auténticas y retroalimentación efectiva podrían explicar esta tendencia (Amaning, 2024; Estrada-Araoz et al., 2023; Rivera Tejada et al., 2023; Wang et al., 2022; Yan et al., 2022).

Estos hallazgos resaltan la importancia de adaptar continuamente enfoques pedagógicos y estrategias tecnológicas según el nivel académico del alumnado, garantizando una educación en línea eficaz, equitativa y adaptada a sus necesidades individuales y colectivas.

El sexo evidenció un efecto sobre la percepción de la educación en línea, diferenciado para los componentes “logros de aprendizaje” y “diseño del curso o la asignatura”.

En el CP2, no se encontraron diferencias significativas entre el sexo masculino y femenino ($p = 0.0735$), lo que sugiere que la educación en línea impacta de manera similar en ambos sexos en cuanto a “logros de aprendizaje”, en concordancia con los hallazgos de Bazán-Ramírez et al. (2023). Sin embargo, la proximidad al umbral de significancia indica que factores como la experiencia con tecnologías o el entorno familiar podrían influir en la percepción, como lo han sugerido estudios previos con resultados divergentes (Estrada y Paz Delgado, 2022; Estrada-Araoz et al., 2023; Remesal y Villarroel, 2023).

En el caso del CP4, sí se evidenció una diferencia significativa ($p = 0.01888$), con el sexo masculino mostrando una percepción más positiva del “diseño del curso o la asignatura” en comparación con el femenino, resultado consistente con los obtenidos por Estrada-Araoz et al. (2023), pero en contraste con estudios que

reportan una percepción más favorable por parte del sexo femenino (Estrada y Paz Delgado, 2022; Remesal y Villarroel, 2023). Esta disparidad, aunque pequeña en magnitud, sugiere que ambos sexos pueden interactuar de manera diferente con el diseño pedagógico y los recursos educativos, destacando la necesidad de estudios adicionales para comprender mejor estos patrones.

Estos hallazgos destacan la importancia de adoptar enfoques pedagógicos inclusivos que consideren las diferencias en la experiencia de aprendizaje según el sexo, promoviendo un diseño instruccional de cursos que atiendan la diversidad de necesidades y expectativas del alumnado.

Uno de los hallazgos más relevantes del estudio se centra en las diferencias significativas percibidas por lo estudiantes según la edad, en relación con los cuatro componentes principales identificados para evaluar la calidad de la educación en línea. Específicamente, los estudiantes de menor edad (15-20 años) poseen percepciones menos favorables en comparación con sus pares de edad avanzada (de 26 a 30 años y de 31 a más años). Estas divergencias podrían ser atribuibles a la ausencia de experiencia y competencias de autorregulación en los estudiantes de menor edad, lo que obstaculiza su adaptación a las exigencias de dicha modalidad de estudio. Investigaciones anteriores han indicado que los alumnos de menor edad pueden experimentar obstáculos en la administración del tiempo, la autodirección y la autorregulación del aprendizaje, lo cual puede influir en su percepción y desempeño en contextos virtuales (Colvin et al., 2024; Cramarenco et al., 2023; Garzón Umerenkova y Gil Flores, 2017).

En contraposición, los estudiantes de edad avanzada pueden exhibir una madurez y experiencia académica superiores, lo que les facilita una adaptación más efectiva a los desafíos de la educación en línea. Adicionalmente, es plausible que estos alumnos hayan desarrollado estrategias eficaces de aprendizaje autónomo y competencias de administración del tiempo, lo que pudo haber contribuido a un desempeño más positivo en contextos virtuales, similar al resultado obtenido en su estudio por Ferrer-Cascales et al. (2011).

Las evidencias encontradas en este estudio no solo brindan información clave sobre la percepción de calidad de la educación en línea en las universidades públicas del contexto ecuatoriano, sino que también tienen implicaciones prácticas relevantes para los gestores educativos en el diseño de estrategias educativas y políticas públicas. Una contribución particular se enfoca en la utilización del cuestionario, cuyas características psicométricas aseguran su validez y fiabilidad, con el objetivo de obtener de los estudiantes la percepción de la calidad de la educación en línea a nivel nacional. Consecuentemente, estos hallazgos pueden servir como fundamento para el desarrollo de modelos adaptativos de enseñanza aprendizaje en línea que integren inteligencia artificial y analíticas de aprendizaje para personalizar la experiencia del estudiante en función del nivel académico, sexo y edad. Por ejemplo, la implementación de sistemas de recomendación fundamentados en datos podría contribuir a la optimización de la organización pedagógica y la retroalimentación,

adaptando los contenidos y metodologías a las necesidades individuales y estilos de aprendizaje de los estudiantes.

CONCLUSIONES

Esta investigación, fundamentada en los hallazgos de validez y fiabilidad del cuestionario, enfatizan la relevancia de tener en cuenta las variables sociodemográficas, tales como el nivel académico, el sexo y la edad, como factores predictores en la toma de decisiones relativas a la administración académica, así como la planificación y ejecución de políticas públicas, con el objetivo de fomentar una educación en línea inclusiva, equitativa y adaptada a las circunstancias del alumnado universitario. Tomando en cuenta estos elementos, las conclusiones derivadas son las siguientes:

1. *Mejora en la planificación educativa*: los resultados sobre la influencia del nivel académico en la percepción del alumnado sugieren la necesidad de diseñar estrategias pedagógicas diferenciadas para estudiantes en distintos niveles de su carrera. Los estudiantes de nivelación y egresados tienen percepciones significativamente más positivas sobre la “organización pedagógica, tecnológica y social del proceso de enseñanza aprendizaje en línea” y la “evaluación y retroalimentación” en la educación en línea, lo que indica que las políticas universitarias deberían enfocarse en ajustar las metodologías y herramientas educativas a lo largo de la trayectoria académica del estudiante, atendiendo a sus cambios en expectativas y necesidades.
2. *Impulso a la inclusión y equidad de género*: el hallazgo de diferencias en la percepción del “diseño del curso o la asignatura” entre el sexo masculino y femenino destaca la necesidad de adoptar políticas que promuevan un entorno de aprendizaje más equitativo. Las instituciones públicas y las universidades pueden utilizar estos datos para diseñar programas de mejora que equilibren las percepciones y experiencias de ambos sexos, asegurando que la estructura de los cursos y asignaturas en línea sea inclusiva y que responda a la diversidad de formas en que interactúan con el entorno digital.
3. *Evolución de la educación en línea y su impacto*: los resultados obtenidos sobre los egresados muestran que estos estudiantes, que vivieron tanto la fase Pre-COVID-19 como la Post-COVID-19, tienen una percepción positiva sobre la evolución de la educación en línea, especialmente en términos de “evaluación y retroalimentación”. Este hallazgo subraya la necesidad de seguir perfeccionando los entornos virtuales de aprendizaje, considerando las lecciones aprendidas durante la pandemia. Los gestores educativos y las universidades deben asegurar que las mejoras implementadas se mantengan y se adapten continuamente para satisfacer las nuevas demandas del contexto educativo digital.
4. *Contribución a las políticas públicas*: este estudio puede servir como una herramienta clave para la gestión pública, ya que los resultados obtenidos pueden

ser utilizados por los gestores educativos para formular políticas que fortalezcan la infraestructura y calidad de la educación en línea en las universidades públicas de Ecuador. Dado el contexto actual de transformación digital en la educación, los hallazgos permiten a los responsables de la gestión pública tomar decisiones informadas que busquen cerrar las brechas digitales y mejorar la accesibilidad y efectividad del aprendizaje en línea.

Sin embargo, las aportaciones significativas de este estudio también presentan ciertas limitaciones. Aunque la muestra seleccionada para este estudio es representativa y el contexto de las universidades públicas en Ecuador establece una estructura mínima que asegura una educación en línea de alta calidad de forma homogénea, esto no obstaculiza que cada entidad invierta recursos adicionales para optimizar dichas condiciones. Esta heterogeneidad podría limitar la generalización de los descubrimientos a escala nacional, incluso cuando la muestra comprende estudiantes de todas las provincias del país. La característica inherente del cuestionario de autoevaluación en línea puede haber inducido un sesgo de deseabilidad social en las respuestas.

Finalmente, es imperativo que las investigaciones futuras contemplen, desde una perspectiva transversal, comparaciones entre carreras en función de sus respectivos campos del conocimiento, así como entre universidades en función de su región geográfica y su naturaleza (públicas o privadas). Además, resultaría provechoso extender el estudio a un enfoque de investigación longitudinal. También los hallazgos podrían ser incrementados mediante una investigación cualitativa que utilice técnicas como entrevistas y grupos focales a estudiantes, docentes y gestores educativos, lo que facilitaría una comprensión más detallada de las percepciones. Otra línea futura de investigación podría enfocarse en la valoración del efecto de la personalización adaptativa del aprendizaje, considerándola como un nuevo componente esencial de estudio a incorporar en la evaluación de la calidad de la educación en línea.

REFERENCIAS

- Amaning, N. (2024). Qualitative analysis on technology acceptance model of accounting faculty perceptions of virtual accounting coursework. *Cogent Education*, 11(1), 2331345. <https://doi.org/10.1080/2331186X.2024.2331345>
- Anderson, M. J. (2001). A new method for non-parametric multivariate analysis of variance. *Austral Ecology*, 26(1), 32-46. <https://doi.org/10.1111/j.1442-9993.2001.01070.pp.x>
- Arslan, M., Edirne, T., Emre, N. y Ozsahin, A. (2023). Medical Students' Perceptions of Online Education during the COVID-19 Pandemic. *Eurasian Journal of Family Medicine*, 12(1), 1-10. <https://doi.org/10.33880/ejfm.2023120101>
- Avendaño, W. R., Luna, H. O. y Rueda, G. (2021). Educación virtual en tiempos de COVID-19: percepciones de estudiantes universitarios. *Formación Universitaria*,

- 14(5), 119-128. <https://doi.org/10.4067/S0718-50062021000500119>
- Barrientos Oradini, N., Yáñez Jara, V., Pennanen-Arias, C. y Aparicio Puentes, C. (2022). Análisis sobre la educación virtual, impactos en el proceso formativo y principales tendencias. *Revista de Ciencias Sociales*, 28(4), 496-511. <https://doi.org/10.31876/rcs.v28i4.39144>
- Bazán-Ramírez, A., Capa-Luque, W., Ango-Aguilar, H., Anaya-González, R. y Cárdenas-López, V. (2023). Perception of Peruvian Students Studying in Biological Sciences about the Advantages of Virtual Classes during the COVID-19 Pandemic. *Education Sciences*, 13(6), 626. <https://doi.org/10.3390/educsci13060626>
- Benjamini, Y. y Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 57(1), 289-300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>
- Burnham, K. P. y Anderson, D. R. (2004). Multimodel Inference: Understanding AIC and BIC in Model Selection. *Sociological Methods & Research*, 33(2), 261-304. <https://doi.org/10.1177/0049124104268644>
- Cabero-Almenara, J., Valencia-Ortiz, R., Llorente-Cejudo, C. y Palacios-Rodríguez, A. de P. (2023). Nativos e inmigrantes digitales en el contexto de la COVID-19: las contradicciones de una diversidad de mitos. *Texto Livre*, 16, e42233. <https://doi.org/10.1590/1983-3652.2023.42233>
- Carbonell García, C. E., Rodríguez Román, R., Sosa Aparicio, L. A. y Alva Olivos, M. A. (2021). De la educación a distancia en pandemia a la modalidad híbrida en pospandemia. *Revista Venezolana de Gerencia*, 26(96), 1154-1171. <https://doi.org/10.52080/rvgluz.26.96.10>
- Colvin, R. G., Delcours, N., Krueger, T. M. y Singh, H. (2024). The Impact of University Student Services and Student Life Characteristics on Students' Perceptions of Online Education. *Journal of Higher Education Theory and Practice*, 24(1), 186-199. <https://doi.org/10.33423/jhetp.v24i1.6770>
- Consejo de Educación Superior (2023). *Reglamento de Régimen Académico*. Quito, Ecuador.
- Cramarenco, R. E., Burcă-Voicu, M. I. y Dabija, D. C. (2023). Student Perceptions of Online Education and Digital Technologies during the COVID-19 Pandemic: A Systematic Review. *Electronics*, 12(2), 319. <https://doi.org/10.3390/electronics12020319>
- Crisol-Moya, E., Herrera-Nieves, L. y Montes-Soldado, R. (2020). Educación virtual para todos: una revisión sistemática. *Education in the Knowledge Society (EKS)*, 21, 13. <https://doi.org/10.14201/eks.23448>
- Cruz, I. R. (2023). Análisis bibliométrico de la educación virtual y rendimiento académico universitario periodo 1996-2021, utilizando datos Scopus. *RISTI - Revista Ibérica de Sistemas e Tecnologías de Informação*, (52), 126-141. <https://doi.org/10.17013/risti.52.126-141>
- Dinu, L. M., Baykoca, A., Dommett, E. J., Mehta, K. J., Everett, S., Foster, J. L. H. y Byrom, N. C. (2022). Student Perceptions of Online Education during COVID-19 Lockdowns: Direct and Indirect Effects on Learning. *Education Sciences*, 12(11), 813. <https://doi.org/10.3390/educsci12110813>
- Ernst, M. D. (2004). Permutation Methods: A Basis for Exact Inference. *Statistical Science*, 19(4), 676-685. <https://doi.org/10.1214/088342304000000396>
- Estrada, L. y Paz Delgado, C. L. (2022). Enseñanza remota de emergencia: Estudio de satisfacción en estudiantes universitarios durante la pandemia por COVID-19. *Perfiles Educativos*, 44(178),

- 46-62. <https://doi.org/10.22201/iisue.24486167e.2022.178.60503>
- Estrada-Araoz, E. G., Gallegos Ramos, N. A., Paredes Valverde, Y., Quispe Herrera, R. y Mori Bazán, J. (2023). Perception of Peruvian university students about virtual education in the context of the COVID-19 pandemic. *Archivos Venezolanos de Farmacología y Terapéutica*, 41(9). http://saber.ucv.ve/ojs/index.php/rev_aavft/article/view/25908
- Ferrer-Cascales, R., Walker, S. L., Reig-Ferrer, A., Fernández-Pascual, M. D. y Albaladejo-Blázquez, N. (2011). Evaluation of hybrid and distance education learning environments in Spain. *Australasian Journal of Educational Technology*, 27(7). <https://doi.org/10.14742/ajet.906>
- García, S. y Herrera, F. (2008). An extension on Statistical comparisons of classifiers over multiple data sets for all pairwise comparisons. *Journal of Machine Learning Research*, 9(12). <https://www.jmlr.org/papers/volume9/garcia08a/garcia08a.pdf>
- García Aretio, L. (2004). Viejos y nuevos modelos de educación a distancia. *Bordón. Revista de Pedagogía*, 56(3-4), 409-429
- García Aretio, L. (2020). Bosque semántico: ¿educación/enseñanza/aprendizaje a distancia, virtual, en línea, digital, eLearning...? *RIED. Revista Iberoamericana de Educación a Distancia*, 23(1), 9-28. <https://doi.org/10.5944/ried.23.1.25495>
- Garzón Umerenkova, A. y Gil Flores, J. (2017). Gestión del tiempo en alumnado universitario con diferentes niveles de rendimiento académico. *Educação e Pesquisa*, 44, e157900. <https://doi.org/10.1590/s1678-4634201708157900>
- Gómez-Arteta, I., Escobar-Mamani, F. y Bonifaz Valdez, B. (2024). Contexto geográfico versus tecnología: una pugna que determina la satisfacción estudiantil con la educación virtual universitaria en Puno-Perú. *RELIEVE-Revista Electrónica de Investigación y Evaluación Educativa*, 30(1). <https://doi.org/10.30827/relieve.v30i1.25133>
- Good, P. (2000). *Permutation Tests: A Practical Guide to Resampling Methods for Testing Hypotheses*. Springer. <https://doi.org/10.1007/978-1-4757-3235-1>
- Harpe, S. E. (2015). How to analyze Likert and other rating scale data. *Currents in Pharmacy Teaching and Learning*, 7(6), 836-850. <https://doi.org/10.1016/j.cptl.2015.08.001>
- Hernández, R. y Mendoza, C. (2018). *Metodología de la investigación. Las rutas cuantitativa, cualitativa y mixta*. Mc Graw Hill Education.
- Higgins, J. J. (2004). *Introduction to Modern Nonparametric Statistics*. Brooks/Cole.
- Lobos, K., Bustos, C., Cobo-Rendón, R. C. y Cisternas, N. V. (2022). Características psicométricas de un cuestionario de expectativas hacia la educación virtual en estudiantes universitarios durante la pandemia COVID-19. *Formación Universitaria*, 15(1), 3-18. <https://doi.org/10.4067/S0718-50062022000100003>
- Martín-Párraga, L., Llorente-Cejudo, C. y Cabero-Almenara, J. (2024). Las TIC como espacio de progreso hacia el Objetivo de Desarrollo Sostenible 4 (ODS4). *Revista Lusófona de Educação*, 61(61), 75-88. <https://doi.org/10.24140/issn.1645-7250.rle61.05>
- 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>
- Pelletier, K., Robert, J., Muscanell, N., McCormack, M., Reeves, J., Arbino, N., Grajek, S., Birdwell, T., Liu, D., Mandernach, J., Moore, A., Porcaro, A., Rutledge, R. y Zimmern, J. (2023). *EDUCAUSE Horizon Report, Teaching and Learning Edition*. EDUCAUSE. <https://>

- library.educase.edu/-/media/files/library/2023/4/2023hrteachinglearning.pdf
- Rahmatpour, P., Nazari, R., Sharif-Nia, H., Mohammad Ibrahim, M., Muhammad Ibrahim, F. y Tayebi Myaneh, Z. (2024). The roles of online courses' design and evaluation in nursing student satisfaction with online learning. *E-Learning and Digital Media*. <https://doi.org/10.1177/20427530241239436>
- Ramírez García, A. G., Espejel García, A., Pirela Hernández, A. A. y Castillo Escalante, I. C. (2021). Educación virtual: alternativa en un sistema educativo globalizado. *Revista Venezolana de Gerencia*, 26(6), 376-389. <https://doi.org/10.52080/rvgluz.26.e6.23>
- Remesal, A. y Villarreal, V. (2023). Challenges for Post-Pandemic Virtual Education in Latin America: A Comparative Analysis of the Emergency Remote Higher Education Process in Chile, Mexico, and Ecuador. *Sustainability*, 15(19), 14199. <https://doi.org/10.3390/su151914199>
- Rivera Tejada, H. S., Otiniano García, N. M. y Goicochea Ríos, E. del S. (2023). Estrategias didácticas de la educación virtual universitaria: Revisión sistemática. *EduTec. Revista Electrónica de Tecnología Educativa*, (83), 120-134. <https://doi.org/10.21556/edutec.2023.83.2683>
- Roco-Videla, Á., Aguilera-Eguía, R. y Olguin-Barraza, M. (2024). Ventajas del uso del coeficiente de omega de McDonald frente al alfa de Cronbach. *Nutrición Hospitalaria*, 41(1), 262-263. <https://doi.org/10.20960/nh.04879>
- Senescyt. (2024). *Buscador de oferta académica*. <https://siau.senescyt.gob.ec/buscador-oferta/>
- Sullivan, G. M. y Artino, A. R. (2013). Analyzing and Interpreting Data From Likert-Type Scales. *Journal of Graduate Medical Education*, 5(4), 541-542. <https://doi.org/10.4300/JGME-5-4-18>
- Vaca-Cárdenas, M. E., Ordoñez-Ávila, E. R. Vaca-Cárdenas, L. A. y Vaca-Cárdenas, A. N. (2024). Students' acceptance toward Asynchronous Virtual Education during COVID-19 pandemic. *Kasetsart Journal of Social Sciences*, 45(2), 483-492. <https://doi.org/10.34044/j.kjss.2024.45.2.14>
- Vásquez Peñafiel, M. S., Nuñez, P. y Cuestas Caza, J. (2023). Competencias digitales docentes en el contexto de COVID-19. Un enfoque cuantitativo. *Pixel-Bit, Revista de Medios y Educación*, 67(67), 155-185. <https://doi.org/10.12795/pixelbit.98129>
- Wang, W., Peslak, A., Kovacs, P. y Kovalchick, L. (2022). Have students' perceptions of online education changed with the easing of COVID-19 restrictions? *Issues in Information Systems*, 23(3), 199-208. https://doi.org/10.48009/3_iis_2022_116
- Yan, Y., Vyas, L., Wu, A. M. y Rawat, S. (2022). Effective online education under COVID-19: Perspectives from teachers and students. *Journal of Public Affairs Education*, 28(4), 422-439. <https://doi.org/10.1080/15236803.2022.2110749>

Fecha de recepción del artículo: 1 de diciembre de 2024

Fecha de aceptación del artículo: 4 de marzo de 2025

Fecha de aprobación para maquetación: 17 de marzo de 2025

Fecha de publicación en OnlineFirst: 3 de abril de 2025

Fecha de publicación: 1 de julio de 2025

Adopción del aprendizaje en línea en la educación superior de China: la perspectiva de los docentes

Adoption of e-learning technology in China's higher education: insights from teachers' perspectives



- ✉ Chongya Ma - *Jilin Normal University, JLNU (China)*
✉ Farrukh Ijaz - *Punjab Municipal Development Fund Company, PMDFC (Pakistán)*
✉ Muhammad Fareed - *School of Business, VIZJA University in Warsaw (Polonia)*
✉ Hood Laeeq - *University of Management and Technology, UMT (Pakistán)*

RESUMEN

El sistema de educación superior ha experimentado profundos cambios a nivel global en la era post-COVID-19. Los avances en ciencia y tecnología, especialmente en las TIC, han ejercido una influencia considerable en este ámbito. La tecnología educativa ha cobrado relevancia, y la investigación sobre el aprendizaje en línea es esencial para explorar esta nueva frontera. Este estudio analiza los factores que influyen en la adopción del aprendizaje en línea en la educación superior de China. Se utilizó una encuesta estructurada para recopilar datos de 280 docentes de diversas instituciones públicas, con el fin de evaluar el impacto de estos factores en la integración de tecnologías de aprendizaje en línea. Se consideraron los docentes de instituciones públicas, ya que presentan mayores niveles de angustia financiera, ansiedad y depresión en comparación con los de instituciones privadas, lo que limita el acceso a recursos en línea, genera problemas de compatibilidad y afecta la autoeficacia. Los hallazgos revelaron que la falta de infraestructura TIC y apoyo técnico, junto con una formación limitada en aprendizaje en línea, representan barreras clave para su implementación. Además, la influencia social, la utilidad percibida y la facilidad de uso desempeñan un papel crucial en la intención de adopción. Este estudio ofrece información valiosa para los responsables de políticas y administradores educativos que buscan fortalecer los entornos de aprendizaje en línea y superar los desafíos del sector de educación superior en China.

Palabras clave: aprendizaje en línea; instituciones de educación superior; UTAUT; TAM; China.

ABSTRACT

The higher education system has been subject to great change in the post-COVID-19 era globally. Developments in science and technology, especially in the field of information, communication and technology (ICT) have a massive influence on the higher education system. Technology in education has garnered importance and research on e-learning has become necessary to explore this new cutting edge of education. This study examines the factors affecting the adoption of e-learning within the higher education system of China. The study applied the structured survey approach to collect data from 280 faculty members of various public institutions to assess how these factors impact the integration of e-learning technologies in China. Public institute teachers are considered as they have high financial distress, anxiety and depression compared to private teachers leading to a lack of online facilities, compatibility issues and self-efficacy. The study findings revealed that a lack of ICT infrastructure and technical support, combined with a lack of e-learning knowledge among educators, are significant barriers to the effective implementation of e-learning in the higher education system of China. Social influence, perceived usefulness and ease of use also play crucial roles in shaping intent towards e-learning adoption. This study provides insights for policymakers and educational administrators seeking to enhance e-learning environments and overcome the challenges faced in the higher education sector in China.

Keywords: e-learning; higher education institutions; UTAUT; TAM; China.

INTRODUCCIÓN

En la era digital, el formato de educación a distancia ha surgido como un componente crucial de las instituciones de educación superior, permitiendo a las instituciones ampliar sus ofertas y llegar a un grupo demográfico más amplio de estudiantes (Gupta y Gupta, 2020). La mayoría de las instituciones de educación superior han insistido recientemente en el uso adecuado de la Tecnología de la Información y la Comunicación (TIC) en el proceso de enseñanza para compensar la escasez de recursos y satisfacer una demanda de educación (Khaldi et al., 2023). El *desafío* enfrenan en su desarrollo profesional y conocimientos es que las instituciones de educación superior adopten un sistema de educación a distancia (Milićević et al., 2021).

China, siendo líder en avances tecnológicos, ha visto un aumento constante en el uso de sistemas de educación a distancia en sus instituciones públicas de educación superior (McConnell, 2018). La educación a distancia está planificada para facilitar y mejorar los procesos de enseñanza al menor costo posible (Osadcha et al., 2021). El gobierno chino ha realizado inversiones estratégicas en tecnología educativa digital, y las instituciones públicas están adoptando gradualmente estos sistemas para aumentar la calidad de los procesos de aprendizaje y enseñanza. La educación a distancia no solo mejora la eficiencia en la entrega de contenido educativo, sino que también garantiza la accesibilidad, flexibilidad y participación de los estudiantes (Chang y Fang, 2020).

Muchos investigadores (Ouadoud et al., 2021; Janelli, 2018; Ennouamani y Mahani, 2017; Lee et al., 2009) presentan definiciones de educación a distancia desde diversos ángulos. Shee y Wang (2008) afirman que la educación a distancia es la combinación de diferentes dispositivos electrónicos (como computadoras portátiles, computadoras, correo electrónico, Internet, intranet y web) para proporcionar educación o capacitación a los estudiantes en el ámbito de la educación a distancia. Turnbull et al. (2021) afirman que la educación a distancia es un modo de tecnología digital para entregar contenido educativo, involucrar a los estudiantes y gestionar el proceso de aprendizaje. Es la provisión en tiempo real de educación y capacitación a través de Internet a un dispositivo de usuario final (Lee et al., 2009). Con la educación a distancia, los estudiantes pueden recibir instrucción o capacitación en cualquier momento y desde cualquier lugar. Es aplicable tanto dentro como fuera del aula y puede ser dirigida por el instructor o por el estudiante (Baig et al., 2022).

Este estudio considera el uso y/o utilización de herramientas TIC, como software orientado a la educación, computadoras e Internet en instituciones de educación superior. Numerosos estudios han examinado los beneficios de la educación a distancia en la educación superior, incluidos la flexibilidad en el aprendizaje (Al Rawashdeh et al., 2021; Salamat et al., 2018), la rentabilidad (Sandars, 2021) y el aumento de la participación de los estudiantes (Serrano et al., 2019). Los sistemas de educación a distancia permiten a las instituciones superar las barreras geográficas,

haciendo que la educación sea más accesible para una diversa gama de estudiantes (Al-Fraihat et al., 2020). En el contexto de la educación superior china, la adopción de sistemas de educación a distancia ha sido impulsada por la necesidad de acomodar a una creciente población estudiantil y el impulso hacia la innovación digital en la educación (Wang et al., 2018).

Anteriormente, el sistema educativo estaba centrado en el instructor, pero ahora se ha centrado en el estudiante. A pesar de estos esfuerzos, existen varios desafíos que impiden la adaptación completa de la tecnología de educación a distancia en el sistema de educación superior (Wang et al., 2018). Estos incluyen infraestructura tecnológica insuficiente (Turnbull et al., 2021; Birch y Burnett, 2009; Surry y Ensminger, 2006), capacitación docente (Islam et al., 2015), apoyo técnico y gerencial (Fan y Yi, 2012), conocimientos informáticos y de educación a distancia entre los facilitadores (Chatti y Hadoussa, 2021), actitudes culturales (Kanuka, 2008), influencia social (Chang y Fang, 2020) y la preparación de maestros y estudiantes para adoptar entornos de aprendizaje digital (Fan y Yi, 2012). Este estudio tiene como objetivo investigar los factores que afectan la adopción de sistemas de educación a distancia en instituciones públicas de educación superior en China.

A pesar de la gran inversión realizada en la infraestructura de educación a distancia, junto con el creciente interés en la educación digital, China aún no ha alcanzado su máxima capacidad para adoptar sistemas de educación a distancia en instituciones de educación superior (Wu, 2024). Como señalan Madni et al. (2022) y Yakubu y Dasuki (2019), el éxito de los sistemas de aprendizaje en línea no depende solo de la infraestructura tecnológica, sino también de factores como la preparación y receptividad de docentes y estudiantes, la provisión de capacitación adecuada, la utilidad y facilidad de uso percibidas, y el apoyo institucional general para la transformación digital. Este estudio tiene como objetivo llenar esta brecha comprendiendo los factores que obstaculizan o facilitan la adopción de sistemas de educación a distancia por parte de las instituciones públicas de educación superior en China. Este estudio tiene como objetivo cerrar esta brecha mediante la comprensión de los factores que dificultan o facilitan la adopción de sistemas de aprendizaje en línea por parte de las instituciones públicas de educación superior en China. El estudio intenta explorar todos los factores que afectan la implementación del aprendizaje en línea en los institutos de educación superior del país.

El estudio tiene como propósito investigar los factores que influyen en la adopción de sistemas de aprendizaje en línea en las instituciones públicas de educación superior en China. Los objetivos específicos son:

1. Examinar el efecto de la falta de infraestructura TIC en la adopción de la educación a distancia en la educación superior.
2. Examinar el efecto de la falta de apoyo técnico y gerencial en la adopción de la educación a distancia en la educación superior.

3. Examinar el efecto de la infraestructura informática y los conocimientos de educación a distancia entre los facilitadores en la adopción de la educación a distancia en la educación superior.
4. Examinar el efecto de la influencia social en la adopción de la educación a distancia en la educación superior.
5. Examinar el efecto de la utilidad percibida en la adopción de la educación a distancia en la educación superior.
6. Examinar el efecto de la facilidad de uso percibida en la adopción de la educación a distancia en la educación superior.

Se discuten los factores que influyen en la adopción de la educación a distancia y se hacen recomendaciones para maximizar la adopción y el uso y utilizar todo el potencial de la tecnología de educación a distancia. El estudio es significativo de varias maneras. En primer lugar, este estudio proporciona información valiosa a los responsables de la formulación de políticas y los líderes educativos para mejorar la infraestructura y las estrategias de aprendizaje digital en instituciones públicas basadas en los factores que influyen en la adopción de sistemas de educación a distancia. El estudio también destaca los desafíos tecnológicos enfrentados por las instituciones y proporciona recomendaciones para mejorar la infraestructura tecnológica necesaria para una adopción exitosa de la educación a distancia. Además, los hallazgos del estudio ayudan a los organismos gubernamentales y a los responsables de la formulación de políticas educativas a desarrollar políticas de educación a distancia más efectivas, asegurando su alineación con las necesidades específicas de las instituciones públicas de educación superior en China. Para las instituciones públicas de educación superior, esta investigación ofrece soluciones prácticas para superar las barreras a la adopción de la educación a distancia, mejorando así la calidad de la educación y los resultados de aprendizaje de los estudiantes.

Revisión de literatura

El aprendizaje en línea es una modalidad de enseñanza que proporciona un entorno educativo y facilita a los estudiantes estudiar cuando y donde que elijan, a través de paneles interactivos que permiten la comunicación por video y audio, así como chats en vivo para el aula virtual y la interacción con el instructor (Andrews y Haythomthwaite, 2007). Bsharat et al. (2023) señalaron que el aprendizaje en línea es el medio más reciente de aprendizaje a través de Internet, que ofrece un entorno de aula virtual para la interacción entre estudiantes e instructores. También puede definirse como un medio de aprendizaje completamente basado en la web, que utiliza redes de área local y tecnologías digitales. (Becker et al., 2012). Sin embargo, el éxito de los sistemas de e-learning depende de la infraestructura tecnológica, la disposición y las actitudes del profesorado y los estudiantes, la utilidad percibida, la influencia

social, la facilidad de uso y el apoyo institucional general para la transformación digital (Yakubu y Dasuki, 2019).

La infraestructura inadecuada de TIC, incluida la conectividad a Internet poco fiable, el hardware y el software insuficientes y el acceso limitado a las herramientas tecnológicas, se ha identificado como una barrera importante para la adopción del e-learning (Nchunge et al., 2013). Huang y Liu (2024) argumentaron que una infraestructura tecnológica adecuada es el elemento secreto para el éxito de los sistemas de e-learning. La conectividad inadecuada a Internet, la ausencia de dispositivos digitales y el software antiguo resultan en una utilización inútil de la tecnología de e-learning (Semlambo et al., 2022). Según Moore y Fodrey (2018), la falta de infraestructura digital necesaria hace que las instituciones no puedan proporcionar un entorno de aprendizaje adecuado a los estudiantes que desean participar en el e-learning. Esto es especialmente notable en China, ya que China tiene disparidades tanto en entornos urbanos como rurales en lo que respecta a sus instituciones educativas (Wang et al., 2018).

Alhabeeb y Rowley (2018) destacaron que la falta de compromiso directivo y de experiencia técnica puede obstaculizar gravemente la eficacia de los sistemas de e-learning. Jiang et al. (2022) descubrieron que la falta de apoyo directivo, junto con la falta de formación en habilidades técnicas tanto para educadores como para el personal administrativo, impide el proceso de implementación del e-learning en las universidades chinas. Además, los desafíos financieros para las instituciones también obstaculizan la adopción de instalaciones de e-learning (Levine y Sun, 2002). Estos incluyen el costo de los cursos de capacitación de apoyo, fondos para el personal de soporte técnico y para proporcionar recursos suficientes para apoyar el desarrollo y mantenimiento del equipo (Fan y Yi, 2012). Muchas instituciones subestiman la adopción de las tecnologías más recientes, ya que calculan erróneamente los costos necesarios para realizar la transición hacia nuevas tecnologías (Mackeogh y Fox, 2009). Sin la guía adecuada, los miembros de la facultad pueden encontrar desconcertante integrar el e-learning en sus metodologías de enseñanza, lo que resulta en resultados subóptimos (Alqahtani y Rajkhan, 2020).

Assarehy Bidokht (2011) aseguran que la falta de familiaridad con las metodologías de enseñanza digital y la limitada alfabetización informática entre los miembros de la facultad pueden ralentizar significativamente la adopción del e-learning. Nawaz (2013) afirmó que el rápido crecimiento y desarrollo en el campo del e-learning ha creado una amplia gama de cursos de e-learning en diversas instituciones académicas, lo que supone un gran desafío para estudiantes y tutores en diferentes niveles de experiencia en la academia. También se requiere que los instructores y estudiantes sean alfabetizados en informática y deben poseer un conocimiento básico de la web, para poder utilizar el software y las herramientas de e-learning. En el contexto de la educación superior de China, Jin et al. (2021) encontraron que los miembros de la facultad que carecen de formación en tecnologías de e-learning a menudo son reacios a su adopción, ya que lo perciben como más difícil y que consume más

tiempo en comparación con los métodos de enseñanza tradicionales cara a cara. En consecuencia, el conocimiento insuficiente entre los facilitadores no solo afecta la implementación, sino también la calidad de los programas de e-learning (Callinan, 2020).

La influencia social, o el impacto que tienen los compañeros, los superiores y la comunidad académica en la decisión de una persona de adoptar una nueva tecnología, es un factor significativo en la implementación del e-learning (Fishbein y Ajzen, 1977). En China, la naturaleza jerárquica de las instituciones de educación superior significa que las decisiones sobre la adopción de tecnología a menudo están influenciadas por la alta dirección y los colegas (Duan et al., 2010). Según Al-Ammari y Hamad (2008), los miembros de la facultad adoptan la tecnología si sus colegas y los líderes de la institución adoptan el e-learning. Por lo tanto, por influencia, existe la capacidad de tener una cultura innovadora en la que el respaldo colectivo del e-learning puede provocar su adopción. Yakubu y Dasuki (2019) confirmaron que las actitudes y percepciones de los estudiantes hacia los sistemas de e-learning han contribuido extensamente a su disposición para participar en la interfaz de aprendizaje electrónico. Por ejemplo, la sustitución del contacto físico en las interacciones puede ir en contra de la aceptación del sistema de aprendizaje en línea debido a la cultura china.

La utilidad percibida del e-learning es el grado en que maestros y alumnos creen que el uso de sistemas de e-learning mejorará su rendimiento académico (Siron et al., 2020). Este concepto, arraigado en el Modelo de Aceptación de la Tecnología (TAM), ha sido ampliamente estudiado en la literatura sobre tecnología educativa (Rahmi et al., 2018). En la educación superior china, Chatti y Hadoussa (2021) encontraron que la utilidad percibida es uno de los principales predictores de la adopción del e-learning. El éxito de los sistemas de e-learning depende en gran medida de la disposición de los miembros de la facultad para integrar la tecnología en sus prácticas docentes (Daneji et al., 2019). Un estudio de Zhong et al. (2019) indicaron que los miembros de la facultad y los estudiantes deben adoptar e internalizar el e-learning solo si creen que el e-learning puede ofrecer flexibilidad, mejorar los resultados de aprendizaje y proporcionar un acceso adecuado a los recursos educativos.

Es el grado hasta el que los individuos creen que el uso de un sistema de aprendizaje en línea no requerirá esfuerzo (Venkatesh y Davis, 1996). En el contexto de la adopción del e-learning en China, la facilidad de uso percibida es crucial, ya que los educadores pueden ser reacios a implementar sistemas que perciben como complejos o difíciles de navegar. Aguti et al. (2013) destacaron que el instructor debe considerar todas las deficiencias que un estudiante podría enfrentar al preparar y esbozar el curso en línea. Además, el idioma y el contenido del curso deben ser fáciles de entender para los estudiantes, para que puedan comprender la conferencia fácilmente. Mensah et al. (2022) sugirieron que en la educación superior china, si los miembros de la facultad y los estudiantes encuentran que la tecnología de e-learning es fácil de usar, con desafíos técnicos mínimos, su disposición para adoptar dichos

sistemas aumenta significativamente. Por el contrario, las interfaces complicadas y la falta de instrucciones claras pueden desalentar el uso, inhibiendo así la adopción generalizada (Rahmi et al., 2018).

Hipótesis de Investigación

Basándose en la literatura anterior, se formulan las siguientes hipótesis de investigación:

- H1: Existe una influencia negativa de la falta de infraestructura de TIC en la adopción del e-learning.
- H2: Existe una influencia negativa de la falta de apoyo directivo y técnico en la adopción del e-learning.
- H3: Existe una influencia negativa de la falta de conocimientos de e-learning y computación entre los facilitadores en la adopción del e-learning.
- H4: Existe una influencia positiva de la influencia social en la adopción del e-learning.
- H5: Existe una influencia positiva de la utilidad percibida en la adopción del e-learning.
- H6: Existe una influencia positiva de la facilidad de uso percibida en la adopción del e-learning.

MARCO TEÓRICO

Modelo de aceptación de tecnología (MAT)

El modelo TAM, desarrollado por Davis (1989), afirma que la facilidad de uso percibida y la utilidad percibida son determinantes clave de la aceptación del usuario de la tecnología. En el contexto del e-learning, si los educadores y los estudiantes perciben que el e-learning mejorará el rendimiento académico (utilidad) y que es fácil de usar sin un esfuerzo excesivo (facilidad de uso), es más probable que implementen y adopten estas tecnologías.

Teoría unificada de aceptación y uso de la tecnología (UTAUT)

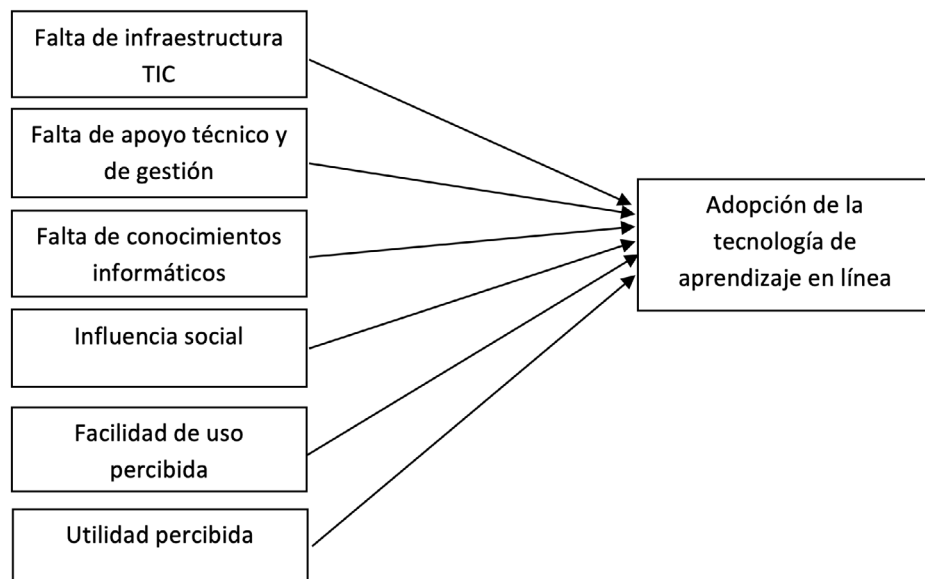
De manera similar, el modelo UTAUT de Venkatesh et al. (2003) amplía el marco TAM al incluir factores adicionales como condiciones facilitadoras, influencia social e intenciones del usuario. El modelo es altamente relevante para comprender cómo la influencia social (el efecto de compañeros y supervisores), la presencia de apoyo directivo y técnico (condiciones facilitadoras) y la percepción de facilidad de uso (expectativa de esfuerzo) contribuyen a la adopción del e-learning.

Teoría de la difusión de innovaciones (DOI)

La DOI explica cómo las nuevas tecnologías se difunden a través de culturas y sistemas sociales (Wejnert, 2002). En el contexto del e-learning, la DOI sugiere que la falta de infraestructura de TIC y el conocimiento de los facilitadores (complejidad) pueden ralentizar la adopción, mientras que la influencia social y el apoyo directivo (observabilidad y posibilidad de prueba) pueden acelerarla. Estos modelos proporcionan una base sólida para examinar cómo las percepciones individuales moldean la adopción del e-learning en la educación superior china.

Basándonos en estos marcos teóricos, desarrollamos el modelo de investigación del presente estudio, que se presenta a continuación en la Figura 1.

Figura 1
Modelo de investigación del estudio



METODOLOGÍA

Diseño de investigación

Este estudio adopta un enfoque de investigación cuantitativo. Se utiliza un diseño de encuesta transversal para recopilar datos de miembros de la facultad de diversas universidades públicas en China.

Población y muestreo

La población de estudio está compuesta por miembros de la facultad que enseñan en instituciones públicas de educación superior en China¹. Hay más de 1500 universidades públicas en China. Se utiliza una técnica de muestreo por conveniencia para recopilar datos de diferentes universidades públicas, asegurando la representación de diversas regiones y disciplinas. En este estudio, se utiliza el muestreo por conveniencia debido a su practicidad, eficiencia y capacidad para recopilar datos de una muestra diversa pero accesible. Dado que los miembros de la facultad en las instituciones públicas de educación superior chinas pueden ser difíciles de alcanzar (Li y Xue, 2021) a través de un muestreo aleatorio debido a la amplia dispersión geográfica y el gran número de instituciones, el muestreo por conveniencia permite a los investigadores recopilar datos de participantes que son fácilmente accesibles, como aquellos dentro de sus redes profesionales o aquellos dispuestos a participar. El tamaño de la muestra de 280 se determina utilizando un análisis estadístico de potencia para asegurar que los resultados sean generalizables a la población más amplia y basados en las muestras de encuestas de estudios similares. Etikan et al. (2016) aseguran que una potencia suficiente para detectar efectos significativos ayuda a compensar algunas de las limitaciones del muestreo por conveniencia al asegurar que el tamaño de la muestra sea adecuado para el análisis. Los maestros de institutos públicos son considerados para el estudio, ya que tienen un alto estrés financiero, ansiedad y depresión en comparación con los maestros del sector privado, lo que lleva a una falta de facilidades en línea, problemas de compatibilidad y autoeficacia.

Recopilación de datos

El cuestionario se utiliza para recopilar datos de la muestra objetivo. Se diseña en función de los factores identificados en la revisión de la literatura, que incluyen la falta de infraestructura de TIC, el apoyo gerencial y técnico, el conocimiento de e-learning y computación entre los facilitadores, la influencia social, la utilidad y facilidad de uso percibidas (véase Tabla 1). Los cuestionarios se distribuyen electrónicamente a través de correo electrónico y portales universitarios. Se pidió a los participantes que respondieran preguntas sobre sus experiencias con los sistemas de e-learning, sus percepciones sobre la tecnología y los desafíos que enfrentan al adoptar la tecnología de e-learning. Las variables del estudio son las siguientes:

Ma, C., Ijaz, F., Fareed, M., & Laeeq, H. (2025). Adoption of e-learning technology in China's higher education: insights from teachers' perspectives [Adopción del aprendizaje en línea en la educación superior de China: la perspectiva de los docentes]. *RIED-Revista Iberoamericana de Educación a Distancia*, 28(2), pp. 325-352.
<https://doi.org/10.5944/ried.28.2.43514>

Tabla 1
Variables de investigación

Número de serie	Variables	Fuentes
1	Falta de infraestructura de TIC	Huang y Liu (2024); Fan y Yi (2012)
2	Falta de apoyo gerencial y técnico	Semlambo et al. (2022)
3	Falta de conocimientos informáticos	Semlambo et al. (2022); Fan y Yi (2012)
4	Influencia social	Mashhour y Saleh (2010)
5	Utilidad percibida	Chatti y Hadoussa (2021)
6	Facilidad de uso percibida	Mensah et al. (2022)
7	Adopción del aprendizaje electrónico	Chatti y Hadoussa (2021)

Análisis de datos

El estudio empleó el software Smart PLS para el análisis de datos. Se utilizaron estadísticas descriptivas, análisis de correlación y PLS-SEM para analizar los datos. Además, se realizó un estudio piloto con una pequeña muestra de participantes de la población objetivo para asegurar la validez y fiabilidad del instrumento de encuesta. La retroalimentación del estudio piloto se utilizó para perfeccionar el cuestionario y mejorar su claridad y coherencia.

Consideraciones Éticas. El estudio se adhiere a las directrices éticas para la realización de investigaciones, incluyendo la obtención del consentimiento informado de los participantes y garantizando la confidencialidad y anonimato de sus respuestas.

RESULTADOS

Estadísticas descriptivas

La Tabla 2 proporciona las estadísticas descriptivas de las variables. Los valores medios de todas las variables superan el 4 en la escala, siendo la utilidad percibida la que tiene la media más alta (4.76), lo que sugiere que los encuestados consideran que el e-learning es altamente útil. Las desviaciones estándar de todas las variables son relativamente bajas, oscilando entre -3.0 y +3.0, lo que indica que las respuestas están concentradas alrededor de la media con una variabilidad limitada (Sandler, 2021). Además, la asimetría negativa observada en todas las variables sugiere que

334 Ma, C., Ijaz, F., Fareed, M., & Laeeq, H. (2025). Adoption of e-learning technology in China's higher education: insights from teachers' perspectives [Adopción del aprendizaje en línea en la educación superior de China: la perspectiva de los docentes]. *RIED-Revista Iberoamericana de Educación a Distancia*, 28(2), pp. 325-352.
<https://doi.org/10.5944/ried.28.2.43514>

las distribuciones están sesgadas a la izquierda, lo que significa que los encuestados tienden a calificar estos aspectos por encima de la media, reflejando actitudes favorables hacia los factores examinados. Los valores de curtosis, que oscilan entre 1.02 y 2.68, indican que algunas variables presentan una distribución leptocúrtica, con mayor agrupamiento alrededor de la media y menos valores extremos, lo que refuerza aún más la consistencia en las respuestas.

Tabla 2

Análisis descriptivo (n = 280)

Variables	Media	Desviación estándar	Oblicuidad	Curtosis
Falta de infraestructura de las TIC	4.27	1.12	-1.19	2.41
Falta de apoyo gerencial y técnico	4.16	1.08	-1.37	2.68
Falta de conocimientos informáticos	4.23	1.03	-0.82	1.02
Influencia social	4.14	1.11	-1.09	1.42
Utilidad percibida	4.76	1.02	-1.25	2.59
Facilidad de uso percibida	4.37	1.01	-0.782	1.11

Prueba de normalidad

Las pruebas de Kolmogorov-Smirnov y Shapiro-Wilk se utilizan para evaluar la normalidad de los datos. En ambas pruebas, un valor p (Sig.) menor de 0.05 indica que los datos se desvían significativamente de una distribución normal. Según la Tabla 3, todas las variables tienen valores p de 0.000 tanto en las pruebas de Kolmogorov-Smirnov como en las de Shapiro-Wilk. Esto significa que, para todas las variables, los resultados de las pruebas indican una desviación significativa de la normalidad. Por lo tanto, se concluye que los datos de todas las variables no siguen una distribución normal, lo que puede tener implicaciones para la selección de métodos estadísticos; por ello, se utilizan pruebas no paramétricas, es decir, PLS-SEM.

Tabla 3*Prueba de normalidad (n = 280)*

Variables	Kolmogorov-Smirnovaa			Shapiro-Wilk		
	Estadística	df	Significado	Estadística	df	Significado
Falta de infraestructura de las TIC	.145	273	.000	.876	273	.000
Soporte gerencial y técnico	.134	273	.000	.903	273	.000
Falta de conocimientos informáticos	.177	273	.000	.901	273	.000
Influencia social	.241	273	.000	.895	273	.000
Utilidad percibida	.149	273	.000	.921	273	.000
Facilidad de uso percibida	.133	273	.000	.923	273	.000
Adopción del aprendizaje electrónico	.208	273	.000	.897	273	.000

Nota: a. Corrección de significancia de Lilliefors

Análisis de confiabilidad

El Alfa de Cronbach mide la consistencia interna, y los valores superiores a 0.7 se consideran generalmente aceptables (Kline, 2023). Raykov y Grayson (2003) argumentaron que la fiabilidad compuesta (CR) proporciona una evaluación más matizada de la fiabilidad, buscando también valores superiores a 0.7. La Tabla 4 presenta las estadísticas de fiabilidad de todos los constructos o escalas. Todas las escalas tienen valores de CR superiores a 0.7, lo que demuestra una fiabilidad aceptable a excelente, respaldando así la robustez de los instrumentos de medición utilizados en este estudio.

Tabla 4*Análisis de confiabilidad*

Número de serie	Escala	Número de artículos	Alfa de Cronbach	CR
1	Falta de infraestructura de las TIC	5	0.782	0.848
2	Soporte gerencial y técnico	4	0.859	0.898
3	Falta de conocimientos informáticos	6	0.923	0.939
4	Influencia social	2	0.635	0.844
5	Utilidad percibida	4	0.703	0.817
6	Facilidad de uso percibida	4	0.870	0.911
7	Adopción del aprendizaje electrónico	5	0.830	0.881

Validez discriminante

La validez discriminante proporciona información sobre la distintividad de cada constructo en relación con los demás, evaluada utilizando el criterio de Fornell-Larcker, donde la raíz cuadrada de la Varianza Media Extraída (AVE) se compara con las correlaciones entre constructos. En la Tabla 5, los valores diagonales (en negrita, aunque no se muestran aquí) representan “la raíz cuadrada de la AVE para cada constructo, que debe ser mayor que las correlaciones inter-constructos en las filas y columnas correspondientes para confirmar la validez discriminante”. La validez discriminante se confirma para todos los constructos, ya que la raíz cuadrada de la AVE para cada constructo es mayor que sus correlaciones inter-constructos correspondientes, lo que indica que cada constructo se mide de manera única y no se superpone excesivamente con otros en el modelo.

Tabla 5
Validez discriminante

	ECK	ICT	AEL	MTS	PUF	PEU	INF	AVE
Falta de conocimientos de aprendizaje electrónico y de informática (ECK)	0.849							0.721
Falta de infraestructura de las TIC (ICT)	0.698	0.726						0.528
Adopción del aprendizaje electrónico (AEL)	0.573	0.548	0.773					0.597
Falta de apoyo gerencial y técnico (MTS)	-0.079	0.305	-0.262	0.830				0.688
Utilidad percibida (PUF)	0.425	0.581	0.605	0.467	0.727			0.528
Facilidad de uso percibida (PEU)	0.442	0.200	0.427	-0.296	0.112	0.848		0.719
Influencia social (INF)	0.166	0.441	0.437	0.284	0.411	0.167	0.854	0.730

Cargas factoriales

Las cargas factoriales representan la correlación entre cada variable observada (indicador) y su constructo latente correspondiente (factor), mostrando cuán bien las variables observadas explican el constructo subyacente. Generalmente, se consideran aceptables las cargas superiores a 0.6, ya que indican que la variable observada se correlaciona fuertemente con su factor latente. La Tabla 6 y la Figura 2 indican que todas las variables/elementos de los constructos tienen cargas factoriales superiores a 0.6, lo que respalda la validez del modelo de medición y confirma que los indicadores son confiables para capturar los constructos previstos.

Tabla 6
Cargas factoriales

Escala	Factores						
	1	2	3	4	5	6	7
Aprendizaje electrónico y conocimientos informáticos	0.852						
	0.843						
	0.899						
	0.834						
	0.848						
	0.818						
Infraestructura TIC		0.733					
		0.639					
		0.734					
		0.779					
		0.740					
Soporte gerencial y técnico			0.891				
			0.829				
			0.881				
			0.703				
Utilidad percibida				0.713			
				0.719			
				0.778			
				0.694			
Facilidad de uso percibida					0.826		
					0.778		
					0.878		
					0.904		
Influencia social						0.815	
						0.892	
Adopción del aprendizaje electrónico							0.741
							0.744
							0.838
							0.793
							0.742

Prueba de hipótesis

Los hallazgos del estudio, tal como se presentan en la Tabla 7 y se muestran en la Figura 3, revelan que varios factores influyen significativamente en el AEL, con la influencia social (INF), la utilidad percibida (PUF) y la facilidad de uso percibida (PEU) demostrando relaciones positivas, como lo indican los coeficientes de 0.456, 0.264 y 0.331, respectivamente. Estos hallazgos sugieren que la INF, junto con una mayor PUF y PEU, son determinantes críticos que facilitan la adopción de iniciativas de e-learning. Por otro lado, el análisis indica que la falta de infraestructura de TIC (ICT) y la falta de conocimiento del contenido de e-learning (ECK) tienen impactos negativos en el AEL, con coeficientes de -0.142 y -0.312, respectivamente. Aunque se encontró que la ICT influye significativamente en el AEL ($p < 0.01$), la relación negativa sugiere que la dependencia de las TIC puede obstaculizar inadvertidamente la adopción debido a una posible sobrecomplejidad o un soporte inadecuado para el usuario. Además, la falta de apoyo gerencial y técnico (MTS) no afectó significativamente al AEL ($p = 0.226$), lo que indica la necesidad de una mayor exploración sobre cómo la capacitación gerencial podría alinearse mejor con las estrategias de e-learning.

Tabla 7
Prueba de hipótesis

Núm.	Variables	Coefficiente	Error Estándar	estadísticas t	valores p	Decisión
H1	ICT > AEL	-0.142	0.069	-2.058	0.000***	Aceptada
H2	MTS > AEL	-0.168	0.139	-1.212	0.226	No aceptada
H3	ECK > AEL	-0.312	0.122	-2.548	0.011**	Aceptada
H4	INF > AEL	0.456	0.121	3.768	0.001***	Aceptada
H5	PUF > AEL	0.264	0.122	2.164	0.006***	Aceptada
H6	PEU > AEL	0.331	0.131	2.523	0.012**	Aceptada

Nota: *** $p < 0.01$ ** $p < 0.05$ * $p < 0.10$

Figura 2
Modelo de medida

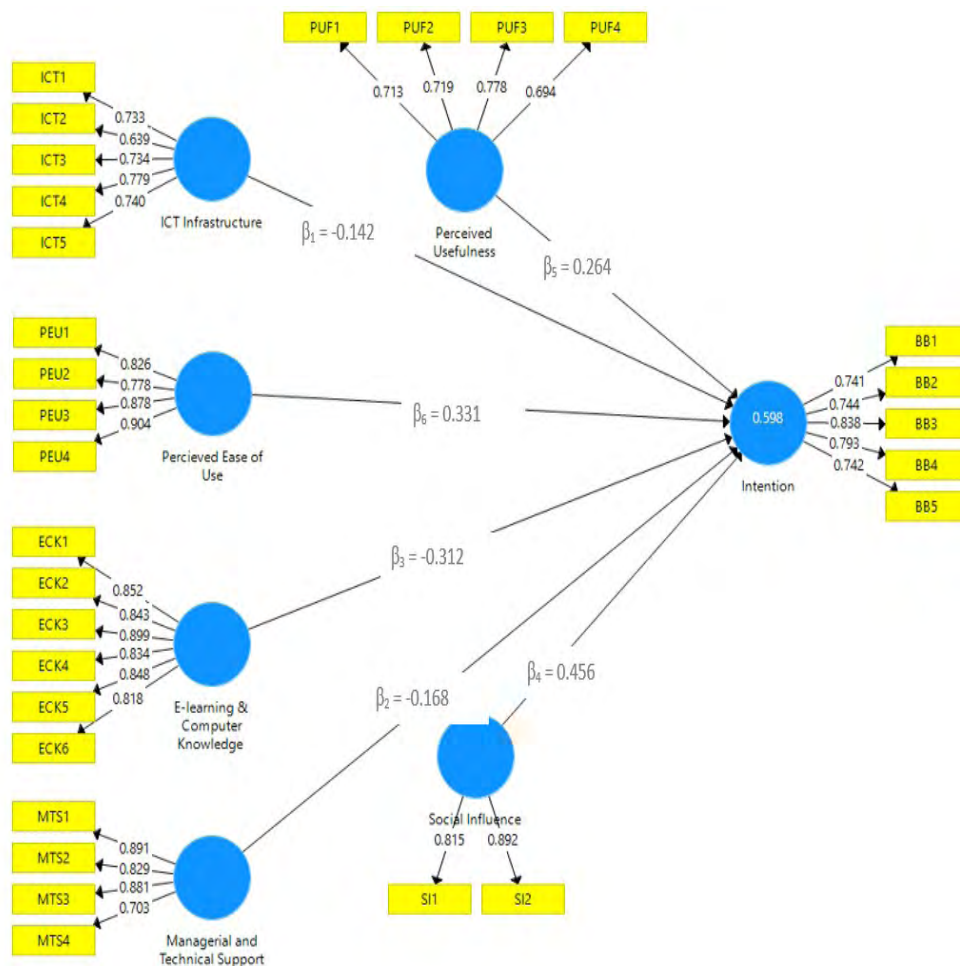
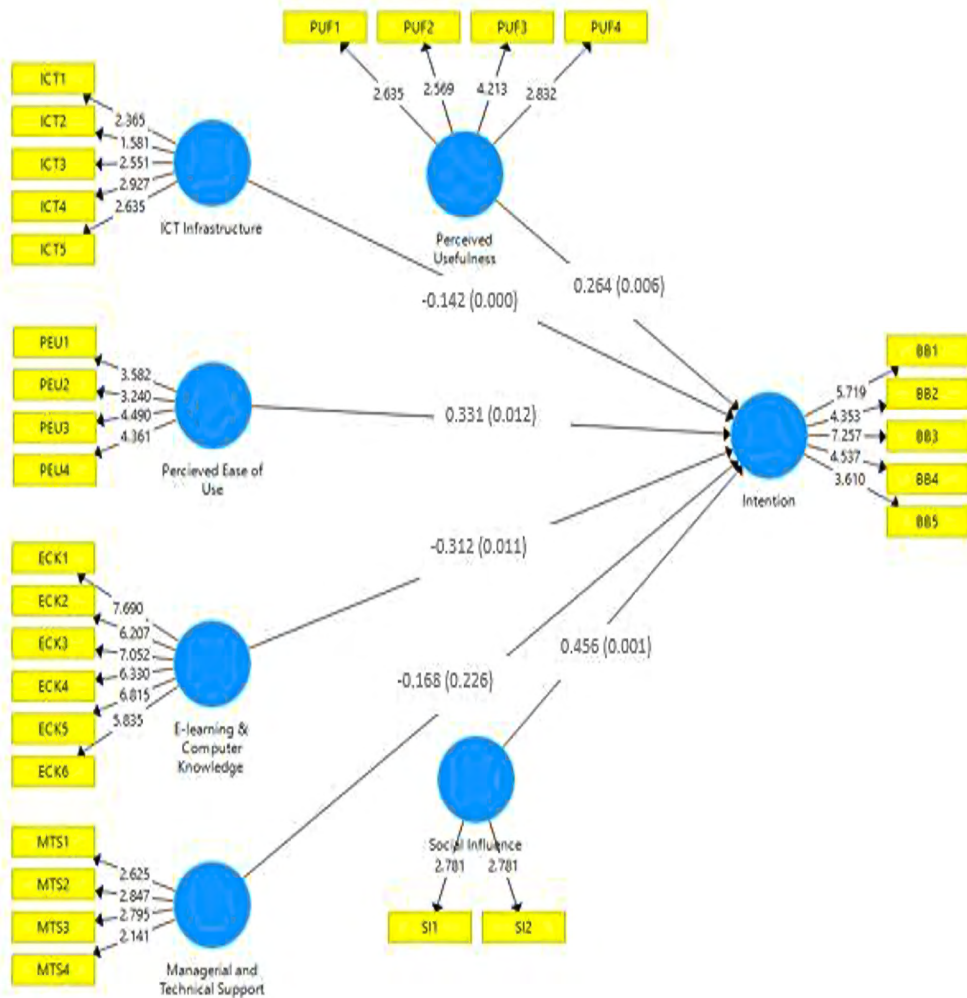


Figura 3
Modelo estructural



DISCUSIÓN

Hipótesis 1

El coeficiente negativo y el valor p significativo sugieren que la falta de infraestructura de TIC tiene una influencia estadísticamente significativa en la adopción de la tecnología de e-learning. Esto significa que mejorar la infraestructura de TIC aumenta positivamente la intención de implementar sistemas de e-learning. Este hallazgo es coherente con investigaciones anteriores que afirman que los entornos tecnológicos excesivamente complejos o mal soportados pueden frustrar a los usuarios, actuando como impedimentos para la adopción de nuevas tecnologías (Alharbi y Drew, 2014). Chuang et al. (2016) sugirieron que una infraestructura inadecuada en términos de TIC afecta la experiencia de aprendizaje y el compromiso de los estudiantes. El desafío que se presenta durante el proceso de adaptación a nuevas tecnologías puede convertirse fácilmente en un obstáculo. Por lo tanto, hay un fuerte énfasis en los buenos sistemas de apoyo y en las instalaciones de capacitación que deberían aliviar los efectos adversos.

Hipótesis 2

La hipótesis de la falta de apoyo gerencial y técnico (MTS) no fue respaldada ($p = 0.226$), lo que sugiere que el apoyo gerencial y técnico puede no ser adecuadamente requerido para la implementación de la tecnología de e-learning en el sistema de educación superior de China. Esto es inconsistente con los hallazgos de Hwang y Chang (2011), quienes observan que el apoyo institucional puede marcar una diferencia efectiva en el compromiso del usuario en un entorno de e-learning. Sin embargo, esto se alinea con estudios que indican que el apoyo gerencial a menudo carece de un enfoque en la integración de la tecnología en las prácticas de gestión (Culnan, 2016). Se necesita más investigación para explorar cómo la capacitación gerencial puede adaptarse para apoyar eficazmente las iniciativas de e-learning.

Hipótesis 3

El coeficiente negativo significativo de -0.312 ($p = 0.011$) para la falta de conocimiento sobre contenido de aprendizaje en línea (ECK) indica que un mayor nivel de falta de ECK puede afectar negativamente la adopción del aprendizaje en línea (AEL). Este hallazgo contrasta con la literatura existente que generalmente enfatiza la importancia del conocimiento del contenido en la promoción de resultados exitosos de e-learning (Ertmer y Ottenbreit-Leftwich, 2010). Al-Azawei et al. (2016) también hipotetizan que cuanto más conozca un usuario sobre su conocimiento del sistema, más probable será que se comprometa con intenciones positivas. La relación

negativa observada aquí puede sugerir que un énfasis excesivo en el conocimiento del contenido, sin abordar las estrategias pedagógicas, puede restar valor a la implementación efectiva del e-learning.

Hipótesis 4

La influencia social (INF) demostró una relación positiva y significativa con el AEL (coeficiente = 0.456, $p < 0.01$), corroborando los hallazgos de la Teoría del Comportamiento Planificado de Ajzen (1991) y estudios más recientes de Rogers (2019), que enfatizan el papel de los factores sociales en la adopción de tecnología. Esto respalda numerosos estudios que destacan el papel crítico de la influencia social o las normas subjetivas en la facilitación del e-learning (Bakia et al., 2012). Una infraestructura efectiva es vital para proporcionar los recursos y herramientas necesarios que mejoren la experiencia de aprendizaje en general. La relación positiva sugiere que el estímulo de compañeros y redes sociales puede mejorar significativamente las intenciones de los usuarios de participar en el e-learning.

Hipótesis 5

El coeficiente positivo de 0.264 ($p = 0.006$) para la utilidad percibida (PUF) confirma la literatura existente que subraya la importancia de la utilidad percibida en los modelos de adopción de tecnología, como el TAM. Cuando los usuarios creen que la tecnología de e-learning proporciona recursos valiosos y mejora su experiencia de aprendizaje, es más probable que se comprometan con ella. Varios estudios, incluidos los de Venkatesh y Davis (2000), destacaron la utilidad percibida en la adopción y la intención conductual de los usuarios hacia el uso de tecnologías de la información. Los hallazgos de este estudio confirmaron que los usuarios tienen más probabilidades de adoptar el e-learning si creen que les ayudará a lograr mejores resultados de aprendizaje.

Hipótesis 6

La facilidad de uso percibida (PEU) mostró una relación positiva significativa con el AEL (coeficiente = 0.331, $p = 0.012$). Este hallazgo se alinea con la literatura del TAM, sugiriendo que la facilidad de uso es crucial para promover el compromiso y la satisfacción del usuario con los sistemas de e-learning (Venkatesh y Davis, 2000). Sugiere que cuando los usuarios encuentran un sistema fácil de usar, es más probable que lo adopten. La relación positiva entre la facilidad de uso percibida y la intención en este estudio subraya la importancia de diseñar un modo de entrega de educación de e-learning fácil de usar para fomentar el compromiso del usuario, lo que lleva a tasas de adopción significativamente mejoradas.

CONCLUSIÓN

Este estudio discute la influencia multifactorial de la adopción y cómo estos factores interactúan entre sí, basándose en el marco conceptual del TAM y el UTAUT. Aunque las percepciones de los usuarios, en particular la utilidad y la facilidad de uso percibidas, fueron los factores determinantes positivos más significativos que están en consonancia con la literatura desarrollada, los efectos adversos de la dependencia de las TIC y el conocimiento del contenido de e-learning sugieren que se necesita un enfoque adecuado para la introducción del e-learning. Además, la falta de sistemas de formación directiva refuerza una vez más la necesidad de que los directivos participen en programas de capacitación adecuados y personalizados que integren competencias tecnológicas en las prácticas de gestión. Estos resultados explican por qué los elementos tecnológicos, gerenciales y pedagógicos deben integrarse para que la adopción del e-learning sea exitosa, ya que mejora las experiencias y resultados educativos. En este sentido, el enfoque en la investigación sobre el diseño centrado en el usuario y los contactos sociales dentro de un entorno de aprendizaje será útil para las partes interesadas para mejorar las intenciones del usuario, lo que también puede mejorar la efectividad y la adopción de los procesos de e-learning. La investigación continuará en estas áreas a medida que el panorama de la tecnología educativa siga evolucionando, por lo que se asegura que los hallazgos sean relevantes y accionables.

Implicaciones teóricas

Esta investigación contribuye a una creciente bibliografía sobre la adopción del e-learning al proporcionar un análisis exhaustivo de los factores que afectan la adopción de tecnologías de e-learning en el sistema de educación superior de China. El estudio amplía la comprensión contemporánea de los modelos de aceptación de tecnología, especialmente en el contexto de la educación superior, que destaca la importancia de la infraestructura de TIC, el apoyo técnico y el conocimiento del educador en la configuración de la adopción del e-learning. Los hallazgos respaldan la importancia de la influencia social, la utilidad percibida y la facilidad de uso, realizando nuevas contribuciones a estos constructos y cómo pueden afectar potencialmente las actitudes y la intención conductual del profesorado hacia la adopción de herramientas de e-learning. El trabajo contribuye a llenar el vacío al traer teorías y aplicaciones de adopción de tecnología hacia adelante dentro de entornos educativos y, por lo tanto, proporciona un marco para estudios futuros de contextos similares.

Implicaciones prácticas

Desde una perspectiva práctica, los hallazgos ofrecen información valiosa para los formuladores de políticas, los administradores educativos y las instituciones que buscan mejorar la experiencia de e-learning en el sistema de educación superior de China. Abordar las barreras identificadas, como la falta de infraestructura de TIC, el apoyo técnico insuficiente y el conocimiento inadecuado de e-learning entre los miembros de la facultad, puede mejorar la integración y la efectividad de las plataformas de e-learning. Las instituciones educativas también pueden centrarse en programas de desarrollo profesional para aumentar la confianza de los educadores en el uso de tecnologías de e-learning. Además, fomentar una influencia social positiva y enfatizar la utilidad y la facilidad de uso percibidas de los sistemas de e-learning puede ayudar a aumentar las tasas de adopción.

El estudio proporciona recomendaciones viables, como priorizar la asignación de recursos, mejorar la capacitación de las partes interesadas y aprovechar la tecnología para abordar los desafíos de implementación. Estos conocimientos tienen como objetivo reducir la brecha entre los hallazgos teóricos y las aplicaciones del mundo real, añadiendo valor al estudio.

Limitaciones y futuras direcciones

Este estudio está limitado por su dependencia de datos autoinformados de miembros de la facultad de instituciones públicas en China, lo que puede limitar la generalización de los hallazgos a instituciones privadas o a otros países con diferentes estructuras educativas. Los estudios futuros pueden abordar tales limitaciones teniendo una muestra más diversificada y empleando métodos longitudinales para la observación del cambio en la adopción del e-learning a lo largo del tiempo. Los estudios futuros pueden aplicar muestreos estratificados o aleatorios que aumentarían su representatividad y generalización. Los datos cualitativos a través de entrevistas o grupos focales también podrían contribuir a una comprensión mucho más profunda sobre las experiencias de los maestros al superar los desafíos que enfrenta el e-learning. Estas estrategias mejoran el rigor en el estudio, pero también proporcionan claridad sobre cómo dicha investigación puede mejorar el proceso de adopción del e-learning en instituciones de educación superior.

NOTAS

1. https://www.cucas.cn/studyinchina/attribute/Public_Universities_23.html

REFERENCIAS

- Aguti, B., Walters, R. y Wills, G. (2013). A framework for evaluating the effectiveness of blended e-learning within universities. En *Society for Information Technology & Teacher Education International Conference* (pp. 1982-1987). Association for the Advancement of Computing in Education (AACE).
- Al Rawashdeh, A. Z., Mohammed, E. Y., Al Arab, A. R., Alara, M. y Al-Rawashdeh, B. (2021). Advantages and disadvantages of using e-learning in university education: Analyzing students' perspectives. *Electronic Journal of E-learning*, 19(3), 107-117. <https://doi.org/10.34190/ejel.19.3.2168>
- Al-Ammari, J. y Hamad, S. (2008, April). Factors influencing the adoption of e-learning at UOB. En *2nd International Conference and Exhibition for Zain E-learning Center* (pp. 28-30). University of Bahrain.
- Al-Azawei, A., Parslow, P. y Lundqvist, K. (2016). Barriers and opportunities of e-learning implementation in Iraq: A case of public universities. *The International Review of Research in Open and Distributed Learning*, 17(5). <https://doi.org/10.19173/irrodl.v17i5.2501>
- Al-Fraihat, D., Joy, M. y Sinclair, J. (2020). Evaluating E-learning systems success: An empirical study. *Computers in Human Behavior*, 102, 67-86. <https://doi.org/10.1016/j.chb.2019.08.004>
- Alhabeeb, A. y Rowley, J. (2018). E-learning critical success factors: Comparing perspectives from academic staff and students. *Computers & Education*, 127, 1-12. <https://doi.org/10.1016/j.compedu.2018.08.007>
- Alharbi, A. y Drew, S. (2014). Factors influencing the adoption of e-learning systems in Saudi Arabia. *International Journal of Education and Information Technologies*, 8(3), 135-141.
- Alqahtani, A. Y. y Rajkhan, A. A. (2020). E-learning critical success factors during the covid-19 pandemic: A comprehensive analysis of e-learning managerial perspectives. *Education Sciences*, 10(9), 216. <https://doi.org/10.3390/educsci10090216>
- Andrews, R. y Haythornthwaite, C. (2007). Introduction to e-learning research. *The SAGE handbook of e-learning research*, 10(9781848607859), n1. <https://doi.org/10.4135/9781848607859.n1>
- Assareh, A. y Bidokht, M. H. (2011). Barriers to e-teaching and e-learning. *Procedia Computer Science*, 3, 791-795. <https://doi.org/10.1016/j.procs.2010.12.129>
- Baig, M. I., Shuib, L. y Yadegaridehkordi, E. (2022). E-learning adoption in higher education: A review. *Information Development*, 38(4), 570-588. <https://doi.org/10.1177/026666669211008224>
- Bakia, M., Shear, L., Toyama, Y. y Lasseter, A. (2012). *Understanding the impacts of technology on education*. U.S. Department of Education, Office of Educational Technology.
- Becker, K., Fleming, J. y Keijsers, W. (2012). E-learning: ageing workforce versus technology-savvy generation. *Education + Training*, 54(5), 385-400. <https://doi.org/10.1108/00400911211244687>
- Birch, D. y Burnett, B. (2009). Bringing academics on board: Encouraging institution-wide diffusion of e-learning environments. *Australasian Journal of Educational Technology*, 25(1). <https://doi.org/10.14742/ajet.1184>
- Bsharat, T. R., Behak, F. y Ismail, I. A. (2023). A guide to e-learning: When the entire world was caught in the center of a storm, technology played a critical role. En *Implementing Rapid E-Learning Through Interactive Materials Development* (pp. 1-10). IGI Global.

- <https://doi.org/10.4018/978-1-6684-4940-0.ch001>
- Callinan, J. (2020). Barriers and facilitators to e-learning in palliative care. *International Journal of Palliative Nursing*, 26(8), 394-402. <https://doi.org/10.12968/ijpn.2020.26.8.394>
- Chang, C. L. y Fang, M. (2020, June). E-Learning and online instructions of higher education during the 2019 novel coronavirus diseases (COVID-19) epidemic. *Journal of Physics: Conference Series*, 1574(1), p. 012166. IOP Publishing. <https://doi.org/10.1088/1742-6596/1574/1/012166>
- Chatti, H. y Hadoussa, S. (2021). Factors affecting the adoption of e-learning technology by students during the COVID-19 quarantine period: The application of the UTAUT model. *Engineering, Technology & Applied Science Research*, 11(2), 6993-7000. <https://doi.org/10.48084/etasr.3985>
- Chuang, M. Y., Chen, C. J. y Lin, M. J. J. (2016). The impact of social capital on competitive advantage: The mediating effects of collective learning and absorptive capacity. *Management Decision*, 54(6), 1443-1463. <https://doi.org/10.1108/MD-11-2015-0485>
- Culnan, M. J. (2016). Information Systems Research: Current Status and Future Directions. *Journal of Management Information Systems*, 33(4), 928-949.
- Daneji, A. A., Ayub, A. F. M. y Khambari, M. N. M. (2019). The effects of perceived usefulness, confirmation and satisfaction on continuance intention in using massive open online course (MOOC). *Knowledge Management & E-Learning*, 11(2), 201-214. <https://doi.org/10.34105/j.kmel.2019.11.010>
- Davis, F. D. (1989). Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology. *MIS Quarterly*, 13(3), 319-340. <https://doi.org/10.2307/249008>
- Duan, Y., He, Q., Feng, W., Li, D. y Fu, Z. (2010). A study on e-learning take-up intention from an innovation adoption perspective: A case in China. *Computers & Education*, 55(1), 237-246. <https://doi.org/10.1016/j.compedu.2010.01.009>
- Ennouamani, S. y Mahani, Z. (2017, December). An overview of adaptive e-learning systems. En *2017 Eighth International Conference on Intelligent Computing and Information Systems (ICICIS)* (pp. 342-347). IEEE. <https://doi.org/10.1109/INTELCIS.2017.8260060>
- Ertmer, P. A. y Ottenbreit-Leftwich, A. T. (2010). Teachers' beliefs and technology integration practices: A developmental approach. *Technology, Pedagogy and Education*, 19(3), 319-336.
- Etikan, I., Musa, S. A. y Alkassim, R. S. (2016). Comparison of convenience sampling and purposive sampling. *American Journal of Theoretical and Applied Statistics*, 5(1), 1-4. <https://doi.org/10.11648/j.ajtas.20160501.11>
- Fan, W. y Yi, S. (2012). Barriers to E-learning adoption in China's traditional higher education institutions: An exploratory study at the institutional level. En *Proceedings of the IADIS International Conference e-Learning*. IADIS Press.
- Fishbein, M. y Ajzen, I. (1977). *Belief, attitude, intention, and behaviour: An introduction to theory and research*. Addison-Wesley
- Gupta, S. B. y Gupta, M. (2020). Technology and E-learning in higher education. *Technology*, 29(4), 1320-1325.
- Huang, H. y Liu, G. (2024). Evaluating students' behavioral intention and system usability of augmented reality-aided distance design learning during the COVID-19 pandemic. *Universal Access in the Information Society*, 23(3), 1217-1231. <https://doi.org/10.1007/s10209-022-00920-9>
- Hwang, G. J. y Chang, H. F. (2011). A formative assessment-based mobile

- learning approach to improving the learning attitudes and achievements of students. *Computers & Education*, 56(4), 1023-1031. <https://doi.org/10.1016/j.compedu.2010.12.002>
- Islam, N., Beer, M. y Slack, F. (2015). E-learning challenges faced by academics in higher education. *Journal of Education and Training Studies*, 3(5), 102-112. <https://doi.org/10.1114/jets.v3i5.947>
- Janelli, M. (2018). E-learning in theory, practice, and research. *Education Issues*, 4(eng), 81-98. <https://doi.org/10.17323/1814-9545-2018-4-81-98>
- Jiang, H., Islam, A. A., Gu, X., Spector, J. M. y Chen, S. (2022). Technology-enabled e-learning platforms in Chinese higher education during the pandemic age of COVID-19. *Sage Open*, 12(2), 21582440221095085. <https://doi.org/10.1177/21582440221095085>
- Jin, Y. Q., Lin, C. L., Zhao, Q., Yu, S. W. y Su, Y. S. (2021). A study on traditional teaching method transferring to E-learning under the COVID-19 pandemic: From Chinese students' perspectives. *Frontiers in Psychology*, 12, 632787. <https://doi.org/10.3389/fpsyg.2021.632787>
- Kanuka, H. (2008). Understanding e-learning technologies-in-practice. En T. Anderson (Ed.), *The theory and practice of online learning* (2^a ed., pp. 91-118). Athabasca University Press. <https://doi.org/10.15215/aupress/9781897425084.006>
- Khaldi, A., Bouzidi, R. y Nader, F. (2023). Gamification of e-learning in higher education: a systematic literature review. *Smart Learning Environments*, 10(1), 10. <https://doi.org/10.1186/s40561-023-00227-z>
- Kline, R. B. (2023). *Principles and practice of structural equation modelling*. Guilford Publications.
- Lee, B. C., Yoon, J. O. y Lee, I. (2009). Learners' acceptance of e-learning in South Korea: Theories and results. *Computers & Education*, 53(4), 1320-1329. <https://doi.org/10.1016/j.compedu.2009.06.014>
- Levine, A. y Sun, J. C. (2002). *Barriers to Distance Education*. American Council on Education, EDUCAUSE.
- Li, J. y Xue, E. (2021). Mapping the education policy of foreign faculty for creating world-class universities in China: Advantage, conflict, and ambiguity. *Educational Philosophy and Theory*, 53(14), 1454-1463. <https://doi.org/10.1080/00131857.2020.1863213>
- MacKeogh, K. y Fox, S. (2009). Strategies for Embedding e Learning in Traditional Universities: Drivers and Barriers. *Electronic Journal of E-learning*, 7(2), pp147-154.
- Madni, S. H. H., Ali, J., Husnain, H. A., Masum, M. H., Mustafa, S., Shuja, J., Maray, M. y Hosseini, S. (2022). Factors influencing the adoption of IoT for E-learning in higher educational institutes in developing countries. *Frontiers in Psychology*, 13, 915596. <https://doi.org/10.3389/fpsyg.2022.915596>
- Mashhour, A., & Saleh, Z. (2010). Evaluating E-learning in Jordanian institutions: why is it lagging? *Quarterly Review of Distance Education*, 11(4), 269.
- McConnell, D. (2018). E-learning in Chinese higher education: the view from inside. *Higher Education*, 75(6), 1031-1045. <https://doi.org/10.1007/s10734-017-0183-4>
- Mensah, I. K., Zeng, G., Luo, C., Lu, M. y Xiao, Z. W. (2022). Exploring the E-Learning Adoption Intentions of College Students Amidst the COVID-19 Epidemic Outbreak in China. *Sage Open*, 12(2), 21582440221086629. <https://doi.org/10.1177/21582440221086629>
- Milićević, V., Denić, N., Milićević, Z., Arsić, L., Spasić-Stojković, M., Petković, D., Stojanović, J., Krkic, M., Sokolov Milovančević, N. y Jovanović, A. (2021). E-learning perspectives in higher

- education institutions. *Technological Forecasting and Social Change*, 166, 120618. <https://doi.org/10.1016/j.techfore.2021.120618>
- Moore, R. L. y Fodrey, B. P. (2018). Distance education and technology infrastructure: Strategies and opportunities. *Leading and managing e-learning: What the e-learning leader needs to know*, 87-100. https://doi.org/10.1007/978-3-319-61780-0_7
- Nawaz, A. (2013). Using e-learning as a tool for 'education for all' in developing states. *International Journal of Science and Technology Education Research*, 4(3), 38-46.
- Nchunge, D. M., Sakwa, M. y Mwangi, W. (2013). Assessment of ICT infrastructure on ICT adoption in educational institutions: a descriptive survey of secondary schools in Kiambu county Kenya. *Journal of Computer Science & Information Technology*, 1(1), 32-45.
- Osadcha, K. P., Osadchyi, V. V. y Spirin, O. (2021). Current state and development trends of e-learning in China. *Information Technologies and Learning Tools*, 85(5), 208-227. <https://doi.org/10.33407/itlt.v85i5.4399>
- Ouadoud, M., Rida, N. y Chafiq, T. (2021). Overview of E-learning Platforms for Teaching and Learning. *Int. J. Recent Contributions Eng. Sci. IT*, 9(1), 50-70. <https://doi.org/10.3991/ijes.v9i1.21111>
- Rahmi, B. A. K. I., Birgoren, B. y Aktepe, A. (2018). A meta-analysis of factors affecting perceived usefulness and perceived ease of use in the adoption of e-learning systems. *Turkish Online Journal of Distance Education*, 19(4), 4-42. <https://doi.org/10.17718/tojde.471649>
- Raykov, T. y Grayson, D. (2003). A test for change of composite reliability in scale development. *Multivariate Behavioral Research*, 38(2), 143-159. https://doi.org/10.1207/S15327906MBR3802_1
- Rogers, M. S. (2019, November). *Creating an Effective e-Learning Culture: The Pedagogical Variation Model (PVM) for Online Learning and Teaching*. En eLearn: World Conference on EdTech (pp. 337-347). Association for the Advancement of Computing in Education (AACE).
- Salamat, L., Ahmad, G., Bakht, M. I. y Saifi, I. L. (2018). Effects of E-Learning on Students' Academic learning at the university Level. *Asian Innovative Journal of Social Sciences and Humanities*, 2(2).
- Sandars, J. (2021). Cost-effective e-learning in medical education. En *Cost-effectiveness in medical education* (pp. 40-47). CRC Press. <https://doi.org/10.1201/9780429091285-6>
- Sandler, A. (2021). Against Standard Deviation as a Quality Control Maxim in Anthropometry. *Econ Journal Watch*, 18(1).
- Semlambo, A. A., Sengati, F. y Angalia, B. (2022). Factors affecting the adoption of e-learning systems in public higher learning institutions in Tanzania: A Case of Institute of Accountancy Arusha (IAA). *Journal of Computer and Communications*, 10(9), 113-126. <https://doi.org/10.4236/jcc.2022.109008>
- Serrano, D. R., Dea Ayuela, M. A., Gonzalez Burgos, E., Serrano Gil, A. y Lalatsa, A. (2019). Technology enhanced learning in higher education: How to enhance student engagement through blended learning. *European Journal of Education*, 54(2), 273-286. <https://doi.org/10.1111/ejed.12330>
- Shee, D. Y. y Wang, Y. S. (2008). Multi-criteria evaluation of the web-based e-learning system: A methodology based on learner satisfaction and its applications. *Computers & education*, 50(3), 894-905. <https://doi.org/10.1016/j.compedu.2006.09.005>
- Siron, Y., Wibowo, A. y Narmaditya, B. S. (2020). Factors affecting the adoption

- of e-learning in Indonesia: Lesson from Covid-19. *JOTSE: Journal of Technology and Science Education*, 10(2), 282-295. <https://doi.org/10.3926/jotse.1025>
- Surry, D. W. y Ensminger, D. C. (2006). *Facilitating the Use of Web-Based Learning by Higher Education Faculty*. Online Submission.
- Turnbull, D., Chugh, R. y Luck, J. (2021). Transitioning to E-Learning during the COVID-19 pandemic: How have Higher Education Institutions responded to the challenge? *Education and Information Technologies*, 26(5), 6401-6419. <https://doi.org/10.1007/s10639-021-10633-w>
- Venkatesh, V. y Davis, F. D. (1996). A model of the antecedents of perceived ease of use: Development and test. *Decision Sciences*, 27(3), 451-481. <https://doi.org/10.1111/j.1540-5915.1996.tb01822.x>
- Venkatesh, V. y Davis, F. D. (2000). A theoretical extension of the Technology Acceptance Model: Four longitudinal field studies. *Management Science*, 46(2), 186-204. <https://doi.org/10.1287/mnsc.46.2.186.11926>
- Venkatesh, V., Morris, M. G., Davis, G. B. y Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 425-478. <https://doi.org/10.2307/30036540>
- Wang, Y., Liu, X. y Zhang, Z. (2018). An overview of e-learning in China: History, challenges and opportunities. *Research in Comparative and International Education*, 13(1), 195-210. <https://doi.org/10.1177/1745499918763421>
- Wejnert, B. (2002). Integrating models of diffusion of innovations: A conceptual framework. *Annual Review of Sociology*, 28(1), 297-326. <https://doi.org/10.1146/annurev.soc.28.110601.141051>
- Wu, J. (2024). E-learning management systems in higher education: Features of the application at a Chinese vs. European university. *Journal of the Knowledge Economy*, 1-31. <https://doi.org/10.1007/s13132-024-02159-6>
- Yakubu, M. N. y Dasuki, S. I. (2019). Factors affecting the adoption of e-learning technologies among higher education students in Nigeria: A structural equation modelling approach. *Information Development*, 35(3), 492-502. <https://doi.org/10.1177/02666666918765907>
- Zhong, J., Xie, H. y Wang, F.L. (2019). The research trends in recommender systems for e-learning: A systematic review of SSCI journal articles from 2014 to 2018. *Asian Association of Open Universities Journal*, 14(1), pp.12-27. <https://doi.org/10.1108/AAOUJ-03-2019-0015>

Fecha de recepción del artículo: 1 de diciembre de 2024

Fecha de aceptación del artículo: 18 de febrero de 2025

Fecha de aprobación para maquetación: 28 de marzo de 2025

Fecha de publicación en OnlineFirst: 11 de abril de 2025

Fecha de publicación: 1 de julio de 2025

ANEXO

Elementos variables

Variable	Elementos
ICT	Tu instituto cuenta con infraestructuras TIC confiables para facilitar cualquier sistema de e-learning.
	Tu instituto dispone de un respaldo de energía confiable (generadores eléctricos, inversores y UPS) para respaldar la adopción de cualquier sistema de e-learning.
	Tu instituto actualiza regularmente sus infraestructuras TIC para adaptarse a la adopción de sistemas de e-learning actualizados.
	En tu instituto, hay acceso a internet confiable con alta velocidad y buen ancho de banda, ya que el costo de internet para uso institucional en el país es asequible.
	Tu instituto tiene una cultura de mantenimiento para reparar y actualizar las instalaciones TIC, incluyendo la garantía de que todas las computadoras en los laboratorios y oficinas estén conectadas a la LAN y funcionen sin problemas.
MTS	Tu instituto tiene políticas de seguridad de sistemas de información/TIC que respaldan adecuadamente la adopción de sistemas de e-learning.
	Cada departamento de tu instituto tiene plena propiedad del sistema de e-learning que utiliza.
	Tu instituto proporciona soporte técnico confiable y oportuno tanto para los facilitadores como para los estudiantes mientras usan los sistemas de e-learning.
	La administración de tu instituto apoya la investigación y el desarrollo en TIC y ofrece programas regulares de concienciación y capacitación.
ECK	Tu instituto ofrece programas regulares de concienciación y capacitación sobre e-learning.
	Todos los facilitadores de tu instituto tienen las habilidades necesarias para utilizar el sistema de e-learning proporcionado por la institución.
	Tu instituto dispone de suficientes computadoras para que todos los estudiantes puedan utilizar el sistema de e-learning en todos los módulos.
	Mi computadora/portátil oficial está actualizada y es confiable para utilizar cualquier sistema de e-learning.
	Prefiero un sistema de e-learning para mi módulo en lugar de un sistema de enseñanza tradicional (basado en clases presenciales).
	Tu instituto lleva a cabo suficientes programas de investigación y desarrollo para mejorar las capacidades de TIC en el apoyo al e-learning.

Variable	Elementos
PUF	El uso de los sistemas de e-learning me ayuda a realizar mis tareas de aprendizaje de una mejor manera.
	El uso de los sistemas de aprendizaje en línea me ayuda a seguir mejor mis cursos.
	El uso de los sistemas de e-learning puede mejorar mi nivel de interacción con mis profesores y compañeros.
	El uso de los sistemas de e-learning puede mejorar la efectividad de mi aprendizaje.
PEU	Me sería fácil aprender a usar el sistema de e-learning recomendado por la universidad.
	Encuentro fácil de usar el sistema de e-learning recomendado por la universidad.
	Podría tener dificultades para usar el sistema de e-learning recomendado por la universidad.
INF	Me sería fácil volverme competente en el uso del sistema de aprendizaje.
	Mis compañeros de facultad me ayudan a usar el sistema de e-learning.
	Interactúo con mis compañeros de facultad sobre el uso de los sistemas de e-learning.
AEL	El e-learning es importante para el futuro de la educación superior.
	Estoy dispuesto/a a adoptar nuevas tecnologías de e-learning en mis prácticas de enseñanza.
	Estoy feliz de integrar herramientas de e-learning en el currículo de mi curso.
	Estoy satisfecho/a con el uso actual de las tecnologías de e-learning en la enseñanza.
	Seguiré utilizando/adaptando herramientas y sistemas de e-learning en mis prácticas de enseñanza.

Propiedades psicométricas de una escala para medir estilos de enseñanza en línea en docentes universitarios

Psychometric properties of a scale to measure online teaching styles in university teachers



- ID Estanislao Casanova Sánchez - *Instituto Tecnológico de Sonora, ITSON (México)*
- ID Armando Lozano Rodríguez - *Instituto Tecnológico de Sonora, ITSON (México)*
- ID Beatriz Yesenia Ahumada Flores - *Instituto Tecnológico de Sonora, ITSON (México)*
- ID Fernanda Inéz García Vázquez - *Instituto Tecnológico de Sonora, ITSON (México)*

RESUMEN

Los estilos de enseñanza se relacionan con el compromiso académico y la motivación de los estudiantes, e influyen de forma significativa en su satisfacción y su percepción sobre el proceso de aprendizaje en línea. Esto destaca la necesidad de contar con instrumentos psicométricamente robustos, que reflejen las particularidades de los entornos de educación en línea. La revisión de literatura mostró que existen múltiples escalas sobre estilos de enseñanza diseñadas para contextos presenciales; sin embargo, son limitadas las propuestas específicas para ambientes virtuales. El objetivo del estudio fue analizar las propiedades psicométricas de una escala diseñada para medir los estilos de enseñanza en línea (EEL) en docentes universitarios. Participaron 240 docentes, quienes completaron una escala de 28 ítems organizada en cuatro dimensiones (diseñador, corrector, mediador y facilitador) y evaluada con una escala Likert de cinco puntos. La validez de contenido fue establecida mediante un juicio de tres expertos, obteniendo un índice de validez de contenido (S-CVI) de .99. El Análisis Factorial Confirmatorio (AFC) mostró un ajuste satisfactorio para una escala de 16 ítems (CFI=.97, TLI=.96, RMSEA=.03), confirmando las cuatro dimensiones teóricas. Además, se encontró evidencia de invarianza factorial entre sexos. Los coeficientes alfa de Cronbach y omega de McDonald oscilaron entre .67 y .71 (.84 global), reflejando consistencia interna aceptable. En conclusión, se presenta una escala con evidencias sólidas de validez y fiabilidad, lo que la posiciona como una herramienta útil para evaluar los estilos de enseñanza en línea y contribuir al mejoramiento de las prácticas pedagógicas en entornos digitales.

Palabras clave: educación en línea; estilos de enseñanza; docentes universitarios; medición; validez.

ABSTRACT

Teaching styles are related to students' academic engagement and motivation. They significantly influence their satisfaction and perception of the online learning process. This highlights the need for psychometrically robust instruments that reflect the particularities of online education environments. The literature review showed that there are multiple scales on teaching styles designed for face-to-face contexts; however, there are limited specific proposals for virtual environments. The study aimed to analyze the psychometric properties of a scale designed to measure online teaching styles (OTS) in university teachers. A total of 240 teachers participated, completing a 28-item scale organized into four dimensions (designer, corrector, mediator, and facilitator) and evaluated with a five-point Likert scale. Content validity was established by a judgment of three experts, obtaining a content validity index (S-CVI) of .99. Confirmatory Factor Analysis (CFA) showed a satisfactory fit for a 16-item scale (CFI=.97, TLI=.96, RMSEA=.03), confirming the four theoretical dimensions. In addition, evidence of factorial invariance between sexes was found. Cronbach's alpha and McDonald's omega coefficients ranged between .67 and .71 (.84 overall), reflecting acceptable internal consistency. In conclusion, a scale with solid evidence of validity and reliability is presented, which positions it as a useful tool to evaluate online teaching styles and contribute to the improvement of pedagogical practices in digital environments.

Keywords: online education; teaching styles; university teachers; measurement; validity.

INTRODUCCIÓN

Los estilos de enseñanza se refieren al conjunto de creencias, valores, actitudes y habilidades que los docentes emplean en el proceso educativo. Estos combinan ideas, técnicas y métodos de enseñanza que brindan coherencia y consistencia a la labor docente, e influyen en cómo los maestros presentan los contenidos, interactúan con los estudiantes, asignan tareas, supervisan el progreso y orientan el aprendizaje (González-Peiteado et al., 2012; Grasha, 1994; Laudadio, 2012). En el ámbito de la educación en línea, Lozano-Rodríguez y Flores-Fahara (2010), describen a los estilos de enseñanza en línea (EEL) como los patrones de comportamiento que los docentes exhiben durante el desarrollo de sus actividades educativas, que involucran aspectos relacionados con el diseño instruccional, la corrección de las tareas, la mediación pedagógica y la facilitación del aprendizaje, todo ello enmarcado en interacciones que se llevan a cabo a través de plataformas digitales.

El estudio de los EEL cobra especial relevancia en un contexto donde la educación virtual ha experimentado un crecimiento acelerado (Leino et al., 2024), impulsado tanto por el avance de las tecnologías digitales, como por la necesidad de adaptarse a nuevas formas de impartir clases; especialmente tras la pandemia por COVID-19, que obligó a las instituciones educativas a trasladarse a entornos virtuales (Arenas-Martínez y Ramírez-Martínez, 2024). Además, es importante analizar cómo la tecnología se ha incorporado al desarrollo profesional de los docentes (Soussi, 2020).

Diversas investigaciones han relacionado los estilos de enseñanza con variables como el logro y el compromiso académico, la motivación, la satisfacción con la enseñanza en línea, el aprendizaje percibido y el tipo de habilidades desarrolladas por los estudiantes (Bartholomew et al., 2018; Franker, 2016; Hidalgo-Cabrillana y Lopez-Mayan, 2018; Shaari et al., 2014; Zante y Klasen, 2021). También han vinculado los estilos de enseñanza con las creencias epistemológicas del docente (Soleimani, 2020), sus competencias, su sentido de autoeficacia y su compromiso (González et al., 2018).

Además, se ha encontrado que los estilos de enseñanza no solo influyen en el rendimiento académico de los estudiantes, sino que afectan de forma significativa su percepción sobre el proceso de aprendizaje en línea (Vikas y Mathur, 2022). Asimismo, autores como Kebritchi et al. (2017), señalan que parte de los problemas asociados con la educación en línea están relacionados con los estilos de enseñanza adoptados por los docentes. En este sentido, entender los EEL es importante para contribuir en la mejora de la calidad de la enseñanza y el aprendizaje en los entornos educativos digitales (Noguez Ortiz, 2024; Wang et al., 2024), en el entendido de que el cambio de modalidad (presencial-virtual) impacta en las prácticas pedagógicas de los docentes (El-Soussi, 2022).

Existen diferencias significativas entre los modelos de estilos de enseñanza presenciales y en línea. Las más relevantes están relacionadas con la interacción docente-estudiante, ya que en la enseñanza presencial el docente puede confrontar

directamente a los estudiantes, gestionando el tiempo y el espacio de manera distinta a como ocurre en entornos virtuales (Silva, 2017). La enseñanza se permea sin intermediarios. Por otro lado, en la virtualidad, existe un equipo tecnológico de por medio. Los esfuerzos del docente para instruir a sus estudiantes no siempre son sincrónicos, lo que implica que la enseñanza puede ocurrir en distintos momentos. Esto requiere que el estudiante desarrolle un mayor autocontrol y gestión del tiempo para participar y mantenerse activo en un curso virtual (Silva-Vera et al., 2023).

Medición de los estilos de enseñanza en línea

La evaluación de los estilos de enseñanza en la educación en línea ha demandado la creación y adaptación de instrumentos específicos que reflejen las particularidades de este entorno. En este sentido, el Cuestionario de Estilos de Enseñanza de Grasha-Riechmann (GRTSI) propuesto por Grasha (1996) y el Inventario de Estilos de Enseñanza en Educación Superior (TSIHE) desarrollado por Abello et al. (2020), son los instrumentos que han sido más ajustados para evaluar los estilos de enseñanza en entornos digitales (Abell et al., 2016; Malay et al., 2024; Malay et al., 2022; Sanje y Varnali, 2014).

Estas escalas permiten identificar la prevalencia de estilos como el facilitador, delegador o el enfoque centrado en el estudiante, y su impacto en variables como el compromiso y éxito académico de los estudiantes en contextos de educación en línea. No obstante, no capturan con precisión las dinámicas específicas que emergen en la interacción educativa mediada por tecnologías digitales, puesto que no fueron desarrolladas con ese propósito.

La revisión de literatura muestra una cantidad importante de estudios sobre estilos de enseñanza para contextos presenciales, que reportan las propiedades psicométricas de diversas escalas y que revelan una gran cantidad de enfoques (Tabla 1). Dichas investigaciones emplean tanto el análisis factorial exploratorio (AFE) como el confirmatorio (AFC), además del análisis de consistencia interna para analizar las escalas; lo que proporciona evidencias robustas sobre su validez y fiabilidad (Abello et al., 2020; Arbabisarjou et al., 2020; Espada et al., 2019; Hurriyetoglu y Kilicoglu, 2020; Merino-Barrero et al., 2017). Sin embargo, esto deja ver una carencia importante de estudios centrados en el análisis de las propiedades psicométricas de escalas diseñadas para entornos digitales.

Tabla 1

Estudios que analizan las propiedades psicométricas de las escalas de estilos de enseñanza para ambientes presenciales

Autor	Método de análisis	Evidencias de validez	Dimensiones
Arbabisarjou et al. (2020)	Análisis factorial confirmatorio y consistencia interna.	AFC: RMSEA=.075, IC 90%, GFI=.74; Fiabilidad: dimensiones α =.68 – .75, test completo α =.72	Autoridad formal, modelo personal, facilitador y delegador.
Abello et al. (2020)	Análisis factorial confirmatorio y consistencia interna.	AFC: $\chi^2=1284.13$, $df=346$, $p<0.001$, $\chi^2/df=3.7$, TLI=0.96, CFI=0.96, RMSEA=0.079; Fiabilidad: $\omega=0.94$ - 0.96	Interacción profesor-alumno, negociación de decisiones, estructuración de la enseñanza, control del comportamiento.
Hurriyetoglu y Kilicoglu (2020)	Análisis factorial confirmatorio y consistencia interna.	AFC: RMSEA=0.06, CFI=0.95, NNFI=0.93, NFI=0.97, GFI=0.95; Fiabilidad: $\alpha=0.88$ (general), subfactores: 0.82 (visual), 0.76 (auditivo), 0.65 (kinestésico)	Visual, auditivo, kinestésico.
Espada et al. (2019)	Análisis factorial exploratorio, análisis factorial confirmatorio y consistencia interna.	AFC: $\chi^2=3743.96$; $gl=1375$; $p=.000$; $\chi^2/gl=2.72$; RMSEA=0.07; Fiabilidad: $\alpha=.78$	Mando directo, práctica, enseñanza recíproca, autoevaluación, inclusión, descubrimiento guiado, resolución de problemas, individualizado, alumnos iniciados, libre exploración, autoenseñanza.

Autor	Método de análisis	Evidencias de validez	Dimensiones
Merino-Barrero et al. (2017)	Análisis factorial exploratorio, análisis factorial confirmatorio y consistencia interna.	AFC: $\chi^2=200.41$; $gl=158$; $p=0.013$; $\chi^2/gl=1.26$; IFI=0.94; CFI=0.93; TLI=0.92; RMSEA=0.04. Fiabilidad: $\alpha=.79$	Tradicional, individualizadores, participativos y socializadores, cognitivos y creativos.
Batista et al. (2015)	Análisis factorial exploratorio y de consistencia interna.	Fiabilidad: $\alpha=.70$	Dinámico, analítico, sistemático y práctico.
González-Peiteado et al. (2012)	Análisis factorial exploratorio y consistencia interna.	Fiabilidad: $\alpha=.89$	Reflexivo, cooperador, académico, individualizador, innovador, indagador.
Laudadío (2012)	Análisis factorial exploratorio y consistencia interna.	Fiabilidad: $\alpha=.81$	Centrado en el docente y centrado en el alumno.
Leung et al. (2003)	Análisis factorial exploratorio y consistencia interna.	Fiabilidad: $\alpha=.73-.83$	Asertivo, sugestivo, colaborativo, facilitador.

Nota: α = Alfa de Cronbach, ω = Omega de McDonald

Uno de los vacíos detectados en la literatura especializada, fue precisamente que los estilos de enseñanza no han sido estudiados a profundidad desde la perspectiva virtual o en línea. Aunque Vieira Barros et al. (2008) catalogaron e identificaron cuatro estilos de uso de espacios virtuales, no especificaban si se trataba desde una perspectiva del aprendizaje o de la enseñanza, dejando abierta la posibilidad. En ese respecto, García-Aranda et al. (2017) retomaron los estilos de uso de espacios virtuales, para enfocarlos a docentes de bachillerato a fin de mejorar los procesos de enseñanza aprendizaje en ambientes virtuales. En esta misma línea, se encontraron también estudios que utilizaban instrumentos de estilos de enseñanza en la presencialidad aplicados a ambientes virtuales o en línea (Deseo et al., 2024; Rakha, 2023).

Ahora bien, en la investigación documental se identificaron únicamente tres instrumentos desarrollados específicamente para analizar los estilos de enseñanza en línea. El primero, desarrollado por Arslan et al. (2014), se diseñó a partir de la literatura sobre estilos de enseñanza y se compone de cuatro dimensiones: diseñador, evaluador, innovador y facilitador. El segundo y el tercero, propuestos por Mori et al. (2023) y Romero-Félix (2023), se basaron en la propuesta teórica de

estilos de enseñanza en línea de Lozano-Rodríguez y Flores-Fahara (2010), quienes establecieron un constructo con cuatro dimensiones: diseñador, corrector, mediador y facilitador.

No obstante, la información revisada permite plantear algunas limitaciones: a) dos de los tres instrumentos identificados para analizar los EEL lo hacen solo a nivel exploratorio (Arslan et al., 2014; Mori et al., 2023) y uno a nivel confirmatorio (Romero-Félix, 2023), por lo que es necesario realizar más estudios que contribuyan al análisis de las estructuras propuestas, puesto que el análisis de validez y fiabilidad de estos instrumentos es esencial para asegurar que reflejen con precisión lo que están midiendo (Valdés-Cuervo et al., 2019); b) el estudio de Arslan et al. (2014), analiza las propiedades psicométricas de su escala a partir de una prueba piloto con 39 docentes, mientras que el de Mori et al. (2023), realiza el estudio desde la perspectiva de los estudiantes. Por lo que, el único instrumento que aporta más evidencias de validez dirigido a docentes universitarios es la propuesta de Romero-Félix (2023). Sin embargo, esta última plantea una escala con base en preferencias y no en acciones observables; c) no se encontraron estudios que examinen la invarianza de medida por sexo. Sin embargo, las investigaciones han señalado que hombres y mujeres pueden tener diferentes enfoques pedagógicos en la práctica docente (Arabit García et al., 2021; Navarro-Patón et al., 2020), por lo que el análisis de invarianza es clave para asegurar la comparabilidad de las mediciones entre géneros (Krosnick y Presser, 2010); d) no se hallaron investigaciones que analicen la validez concurrente, un aspecto importante para evaluar en qué medida los puntajes del instrumento se correlacionan con otras escalas que valoran constructos similares o teóricamente relacionados (American Educational Research Association et al., 2014) y; e) no se encontraron estudios robustos sobre estilos de enseñanza en línea en México, que informen sobre sus propiedades psicométricas.

Validez concurrente

Para garantizar la validez concurrente, es fundamental analizar la relación entre la escala EEL y otros constructos teóricamente relacionados. En este caso, se seleccionó la felicidad subjetiva como variable de referencia, definida como la experiencia habitual de emociones positivas (Lyubomirsky et al., 2005), y considerada un indicador de la calidad de vida del docente que influye en su disposición para enseñar e interactuar de manera positiva con los estudiantes (Chan, 2009). Estudios previos han demostrado que las emociones de los docentes impactan su comportamiento (Frenzel, 2014), sus estilos de enseñanza (Burel et al., 2021) y las funciones que desempeñan en el ámbito educativo (Yirci et al., 2018). Estas relaciones permiten analizar la validez concurrente al explorar la conexión entre la felicidad subjetiva y la escala EEL.

Modelo teórico

Se tomó como referencia la propuesta teórica de Lozano-Rodríguez y Flores-Fahara (2010), quienes propusieron un modelo que examina la relación entre las experiencias docentes vividas durante sus clases en línea y cinco filosofías fundamentadas en la teoría de Zinn (2004): conductual, liberal, progresista, humanista y radical. Dicha propuesta identificó cuatro estilos de enseñanza basados en las percepciones sobre las estrategias de enseñanza en línea de docentes universitarios y, parte del supuesto de que los estilos no son fijos, sino que se adaptan a las necesidades, personalidades y preferencias de los instructores (Sternberg, 1997). A continuación, se describen los cuatro estilos:

Diseñador: es un docente innovador y adaptable, enfocado en el desarrollo ético y pragmático de los estudiantes. Prioriza actividades ajustadas a las necesidades, se comunica por foros y emplea métodos educativos humanistas y progresistas. Ajusta el diseño de su proceso de enseñanza según la retroalimentación de los estudiantes.

Corrector: este profesor enfatiza la excelencia y la estructura en las tareas, promueve la interacción mediante foros y comunicación sincrónica. Su enfoque se alinea con el conductismo y las artes liberales, fomenta la comunicación clara y empática y la adaptabilidad estructural.

Mediador: este educador es un humanista que enfatiza la confianza en los estudiantes y la resolución de conflictos mediante la negociación y el bien común. Promueve la justicia social y utiliza foros en línea para interactuar con los estudiantes. Prioriza una enseñanza colaborativa y adaptativa con retroalimentación constructiva.

Facilitador: este docente reconoce a cada estudiante como único en desarrollo. Ofrece orientación y asesoramiento personalizado. Equilibra el apoyo con distancia, valorando tanto el aspecto humano como el logro de las tareas. Utiliza herramientas sincrónicas y foros interactivos. Guía con principios progresistas y artes liberales, complementados con diversos métodos pedagógicos y un proceso de evaluación formativa.

Esta propuesta teórica, permite orientar el análisis de los estilos de enseñanza a un contexto de educación en línea, considerando: a) el uso de la tecnología como base del proceso educativo; b) los diferentes tipos de interacción y comunicación (síncrona o asíncrona) que suceden en un ambiente virtual de aprendizaje y; c) que el modelo se desarrolla a partir de la perspectiva de docentes universitarios.

Con base en este modelo, Romero-Félix (2023) diseñó una propuesta inicial de 32 ítems que redujo a una escala de 12 ítems tras el proceso de análisis de validez y confiabilidad. No obstante, en la revisión del instrumento se identificó que la redacción de los ítems se orientó hacia las preferencias docentes y que la escala Likert utilizada fue de cuatro puntos. Por ello, se revisó y modificó el instrumento con el objetivo de transformar los ítems de un enfoque basado en preferencias a uno centrado en acciones observables, para obtener datos más objetivos y precisos sobre

el comportamiento de los docentes. Además, se cambió la escala Likert a una de cinco puntos.

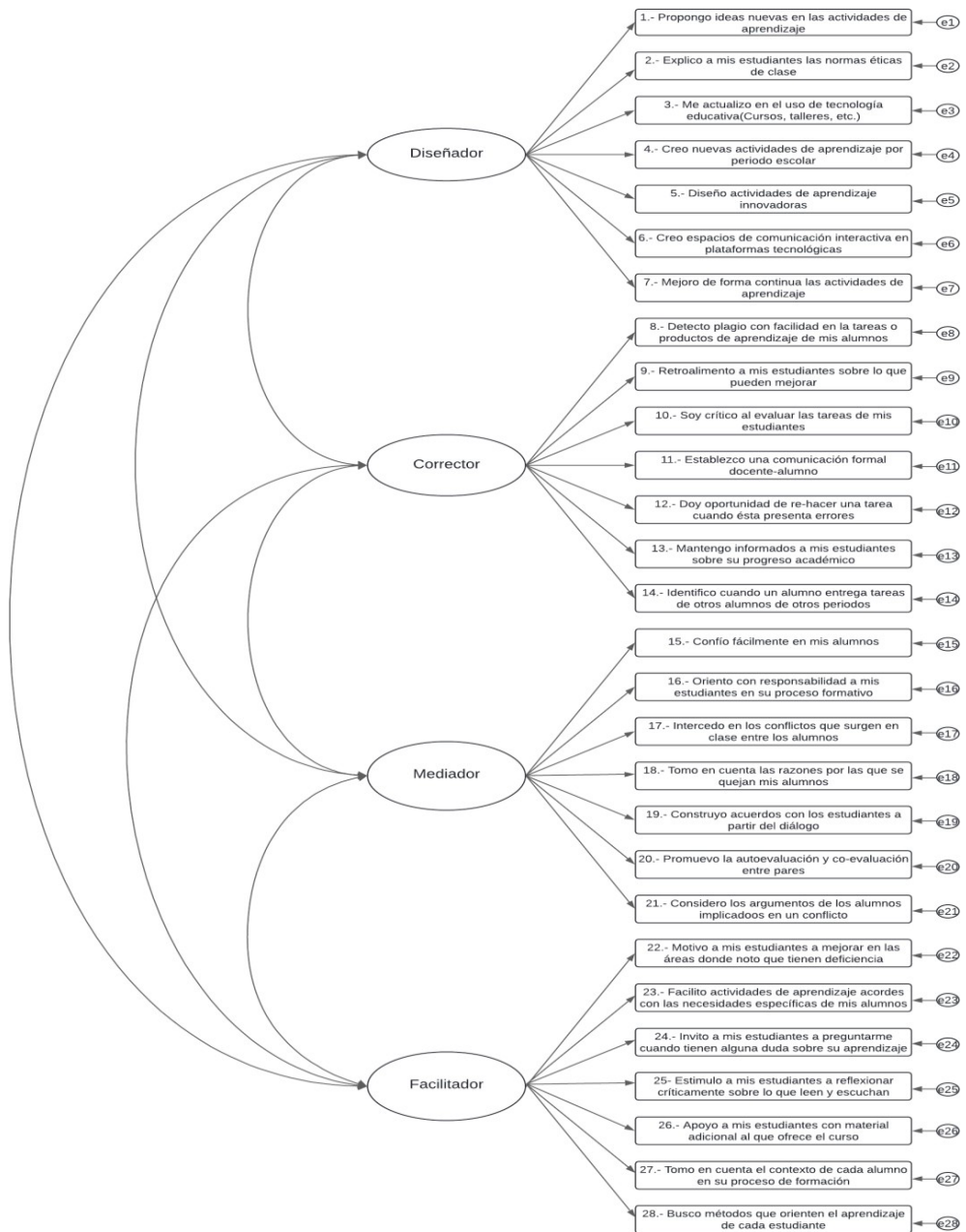
Lo anterior, considerando que los ítems de preferencia que tradicionalmente se formulan en términos de gusto o acuerdo con afirmaciones, son útiles para medir actitudes, creencias y disposiciones internas (Ajzen, 2005); sin embargo, estos constructos pueden no correlacionarse directamente con los comportamientos reales (Krosnick y Presser, 2010). Para lograr una medición más directa y fiable del desempeño o de las prácticas cotidianas, es preferible emplear ítems que capten la frecuencia con la que los sujetos realizan acciones observables y concretas (Podsakoff et al., 2003).

En este sentido, de acuerdo con Bandura (1997), los reportes de acciones ofrecen una vía más precisa para evaluar la conducta en contextos específicos, lo que permite vincular los resultados con intervenciones más efectivas en el ámbito educativo. Así, esta transformación fortalece la validez del cuestionario, al reflejar mejor el comportamiento en la práctica real de los docentes (Furr, 2011). Por otro lado, la utilización de una escala Likert de cinco o más puntos es recomendada en la literatura psicométrica, ya que incrementa la sensibilidad del instrumento, permitiendo captar variaciones más finas en las respuestas de los participantes y mejorando la consistencia interna de la escala (Lozano et al., 2008).

Por lo anterior, el presente estudio propone analizar las propiedades psicométricas de una escala diseñada para medir los estilos de enseñanza en la educación en línea (EEL) en docentes de educación superior. Para lograr este objetivo, se llevó a cabo un análisis que incluye la validez de contenido de la escala mediante el juicio de expertos, la evaluación de la validez de la estructura interna del modelo a través de un Análisis Factorial Confirmatorio (AFC) (Figura 1), el análisis de la invarianza de medida para ambos sexos, el análisis de validez concurrente y el análisis de las evidencias de fiabilidad de la escala.

Figura 1

Modelo teórico para medir estilos de enseñanza en línea



MÉTODO

Participantes

En el estudio participaron 240 docentes de educación superior de un municipio del sur de Sonora, México. El muestreo se realizó de forma no probabilística por conveniencia y solo participaron aquellos docentes que impartieron al menos un curso en línea en los años 2022 y 2023. El 48.8 % (n = 117) fueron mujeres, el 49.5 % (n = 119) fueron hombres y el 1.7 % (n = 4) prefirió no especificar. Sus edades oscilaron entre los 25 y los 75 años (M = 44.5 años, DE = 10.14). La media respecto a la experiencia docente fue de 15.8 años (DE = 9.4) e impartiendo clases en línea fue de cinco años (DE = 4.78).

Instrumentos

Estilos de enseñanza en línea

El instrumento EEL se compone de 28 ítems con una escala de respuesta tipo Likert de cinco puntos (0 = Nunca, 1 = Casi nunca, 2 = A veces, 3 = Casi siempre, 4 = Siempre), distribuidos en cuatro dimensiones de siete ítems cada una: estilo diseñador (ej., Propongo ideas nuevas en las actividades de aprendizaje); estilo corrector (ej., Retroalimentación a mis estudiantes sobre lo que pueden mejorar); estilo mediador (ej., Oriento con responsabilidad a mis alumnos en su proceso formativo) y; estilo facilitador (ej., Apoyo a mis estudiantes con material adicional al que ofrece el curso).

Felicidad subjetiva

Se empleó la versión adaptada por Valdez (2022) de la Escala de Felicidad Subjetiva, propuesta por Lyubomirsky y Lepper (1999). Esta escala mide la percepción global de felicidad a través de declaraciones que los participantes valoran de forma personal o en comparación con otras personas (ej., En general, me considero muy feliz). El instrumento consta de seis ítems, evaluados mediante una escala Likert de cinco puntos (0 = En total desacuerdo, 1 = En desacuerdo, 2 = Ni de acuerdo ni en desacuerdo, 3 = De acuerdo, 4 = Totalmente de acuerdo).

Los resultados del AFC indican que el modelo teórico presenta un buen ajuste con los datos ($X^2 = 12.44$, $gl = 7$, $p = .087$; SRMR = .03; AGFI = .95; TLI = .98; CFI = .99; RMSEA = .06, IC 90 % [.00, .11]) (Blunch, 2013; Byrne, 2010; Valdés-Cuervo et al., 2019). Además, la escala mostró una fiabilidad adecuada, con un coeficiente omega de McDonald de .78 (Katz, 2006).

Procedimiento

Recolección de datos

Se solicitó la autorización del comité de ética institucional para llevar a cabo la investigación, así como el permiso de las autoridades educativas para la recolección de datos. La administración del instrumento se realizó de la siguiente manera: a) de forma presencial mediante un cuadernillo que se aplicó en los cubículos y zonas de descanso de los docentes en dos campus universitarios; b) a través de un cuestionario en línea que se distribuyó en los grupos académicos de WhatsApp y, c) distribuyendo el cuestionario en línea mediante el correo electrónico institucional. En todos los casos, se extendió una invitación para participar en la investigación de forma voluntaria, donde se describió el objetivo y el consentimiento informado.

Validez de expertos

Para evaluar la validez de expertos, el instrumento se envió por correo electrónico a un grupo de nueve profesores investigadores. De ellos, solo tres respondieron: dos hombres y una mujer, todos con estudios de doctorado y un promedio de 20 años de experiencia en el tema de “Estilos de enseñanza”. La evaluación se realizó en cuatro categorías a través de doce indicadores: claridad (ej. El ítem es claro), coherencia (ej. El ítem se encuentra completamente relacionado con la dimensión que está midiendo), relevancia (ej. El ítem es relevante y debe ser incluido) y suficiencia (evaluada por dimensión, ej. Los ítems son aptos y suficientes).

Los datos recopilados se analizaron utilizando el índice de validez de contenido (CVI), el cual es la proporción de expertos que califican los ítems como 3 o 4 en una escala de Likert de 4 puntos (1 = No cumple con el criterio; 2 = Bajo nivel; 3 = Moderado nivel; 4 = Alto nivel), conforme a la metodología de Almanasreh et al. (2019).

Este enfoque establece criterios específicos en función del número de expertos involucrados y, evalúa el grado de acuerdo entre ellos respecto a cada ítem a través del índice de validez de contenido individual (I-CVI) (Castillo Salcido et al., 2024). Además, se evaluó la calidad global de la escala utilizando el índice de validez de contenido escalar promedio (S-CVI/ave). Para determinar la concordancia entre los jueces, se tomaron como referencia los siguientes valores: I-CVI, igual o superior a 0.78 en cada ítem y para la escala completa, S-CVI/ave igual o superior a 0.90 (Polit et al., 2007).

El análisis reveló un alto nivel de acuerdo sobre la validez de los ítems del instrumento, con un I-CVI = 1 en los 28 ítems en las categorías claridad y suficiencia, 27 en relevancia y 26 en coherencia. Sin embargo, dos ítems obtuvieron un I-CVI de .67 en la categoría de coherencia y uno en relevancia, por lo que se modificaron siguiendo las sugerencias de los evaluadores para mantenerlos en la escala.

Respecto al S-CVI/ave, los resultados mostraron valores superiores a .90 en todas las categorías evaluadas: .99 en relevancia, 1 en claridad, .97 en coherencia y 1 en suficiencia. Esto confirmó la validez de contenido de la escala EEL en sus cuatro dimensiones (conformadas por siete ítems cada una), con un S-CVI promedio de .99.

Análisis de datos

Se evaluó la normalidad de los datos calculando los estadísticos descriptivos (media, desviación estándar, asimetría y curtosis) con el software SPSS. Se consideraron valores de asimetría entre -2 y 2, y de curtosis entre -7 y 7, como indicadores de una distribución normal (Bandalos y Finney, 2019; Muthén y Kaplan, 1985). Con el propósito de evaluar el grado de ajuste de los datos al modelo teórico propuesto, se llevó a cabo un Análisis Factorial Confirmatorio (AFC) basado en un modelo previamente establecido, lo que permitió comprobar la estructura interna de la escala. Además, este análisis posibilitó el cálculo de la invarianza factorial, la cual permite comprobar si la escala mantiene su estructura factorial en diferentes grupos poblacionales, asegurando la equivalencia en la medición y la posibilidad de comparaciones válidas entre grupos (Furr, 2022).

El Análisis Factorial Confirmatorio se realizó con el software AMOS, empleando el método de Máxima Verosimilitud y el método Bootstrap (500 repeticiones, IC 95 %) (Byrne, 2010; Hancock y Liu, 2012). Se consideraron los siguientes índices de bondad de ajuste para evaluar el ajuste del modelo a los datos: Chi cuadrado con probabilidad asociada ($X^2, p > .001$), grados de libertad ($gl > 0$), raíz cuadrada de la media de residuos cuadrados ($SRMR \leq .05$), índice de bondad de ajuste ajustado ($AGFI \geq 0.90$), índice de Tucker-Lewis ($TLI \geq 0.90$), índice de ajuste comparativo ($CFI \geq 0.95$) y error cuadrático de la media de aproximación ($RMSEA \leq 0.08$) (Blunch, 2013; Byrne, 2010; Valdés-Cuervo et al., 2019). Se empleó el software estadístico JASP para calcular los coeficientes de fiabilidad omega de McDonald ($\omega \geq 0.65$) y alfa de Cronbach ($\alpha \geq 0.65$), con el propósito de evaluar la fiabilidad de las medidas (Katz, 2006).

El análisis de la validez concurrente se llevó a cabo mediante una correlación de Pearson entre las puntuaciones de la escala EEL y la medida de felicidad subjetiva, las cuales se recolectaron de manera simultánea. Este enfoque se utilizó para evaluar si las dimensiones de la escala EEL y su puntaje global se relacionaban con la felicidad subjetiva, lo que permitió comprobar si la escala mide constructos coherentes con otros teóricamente relacionados (Furr, 2022). Se consideraron como correlaciones significativas aquellas con un valor de $p < .01$ (Funder y Ozer, 2019). El análisis se realizó utilizando el software SPSS. Se interpretaron las correlaciones para determinar la magnitud y dirección de la relación, tomando en cuenta que correlaciones positivas y significativas proporcionarían evidencia de validez concurrente para la escala EEL.

Para examinar la invarianza de medida de la escala para los sexos masculino y femenino, se desarrollaron varios modelos anidados. Cada modelo fue comparado

progresivamente: el primer modelo sin restricciones (invarianza configural), el segundo con restricciones en los coeficientes de regresión (invarianza métrica), el tercero con restricciones en las cargas factoriales e interceptos de los factores (invarianza escalar), y el cuarto con restricciones en cargas factoriales, interceptos, covarianzas y varianzas residuales (invarianza residual). Se emplearon como indicadores de invarianza, las diferencias ΔX^2 con $p > .001$, ΔCFI menores a .01 y $\Delta RMSEA$ menores a .015 (Blunch, 2013; Byrne, 2010).

RESULTADOS

Análisis descriptivo

El análisis descriptivo de los datos de la escala nos permite observar que las respuestas de los docentes ($n=240$) muestran una tendencia a “casi siempre”, con un rango de medias que va de 2.9 puntos (ítem 3, $DE=0.99$) a 3.72 puntos (Ítem 7, $DE=0.53$) (Tabla 2). Además, el análisis de asimetría y curtosis sugiere que los datos tienen un comportamiento que asemeja normalidad (asimetría $+/-2$, curtosis $+/-7$) (Bandalos y Finney, 2019).

Tabla 2

Media, desviación estándar, asimetría y curtosis de los ítems de la escala EEL

Ítems	M	DE	Asimetría	Curtosis
1. Propongo ideas nuevas en las actividades de aprendizaje.	3.18	0.76	-0.42	-0.77
3. Me actualizo en el uso de la tecnología educativa (cursos, talleres, etc.).	3.43	0.73	-1.05	0.38
6. Creo espacios de comunicación interactiva en plataformas tecnológicas.	2.90	0.99	-0.52	-0.45
7. Mejoro de forma continua las actividades de aprendizaje.	3.29	0.69	-0.61	-0.16
9. Retroalimentación a mis estudiantes sobre lo que pueden mejorar.	3.56	0.56	-0.81	-0.38
10. Soy crítico al evaluar las tareas de mis estudiantes.	3.51	0.63	-1.13	1.16
11. Establezco una comunicación formal docente-alumno.	3.72	0.53	-1.95	3.83

Ítems	M	DE	Asimetría	Curtosis
13. Mantengo informados a mis estudiantes sobre su progreso académico.	3.51	0.63	-1.23	1.83
16. Oriento con responsabilidad a mis alumnos en su proceso formativo.	3.70	0.52	-1.49	1.32
17. Intercedo en los conflictos que surgen en clase entre los alumnos.	3.06	1.00	-0.94	0.40
18. Tomo en cuenta las razones por las que se quejan mis alumnos.	3.62	0.57	-1.17	0.38
21. Considero los argumentos de los alumnos implicados en un conflicto.	3.45	0.72	-1.39	2.44
23. Facilito actividades de aprendizaje acordes con las necesidades específicas de mis alumnos.	3.45	0.70	-1.12	0.76
26. Apoyo a mis estudiantes con material adicional al que ofrece el curso.	3.43	0.71	-1.05	0.54
27. Tomo en cuenta el contexto de cada alumno en su proceso de formación.	3.13	0.78	-0.44	-0.61
28. Busco métodos que orienten el aprendizaje de cada estudiante.	3.12	0.82	-0.73	0.28

Nota: M = Media; DE = desviación estándar; la numeración de los ítems corresponde al modelo teórico inicial.

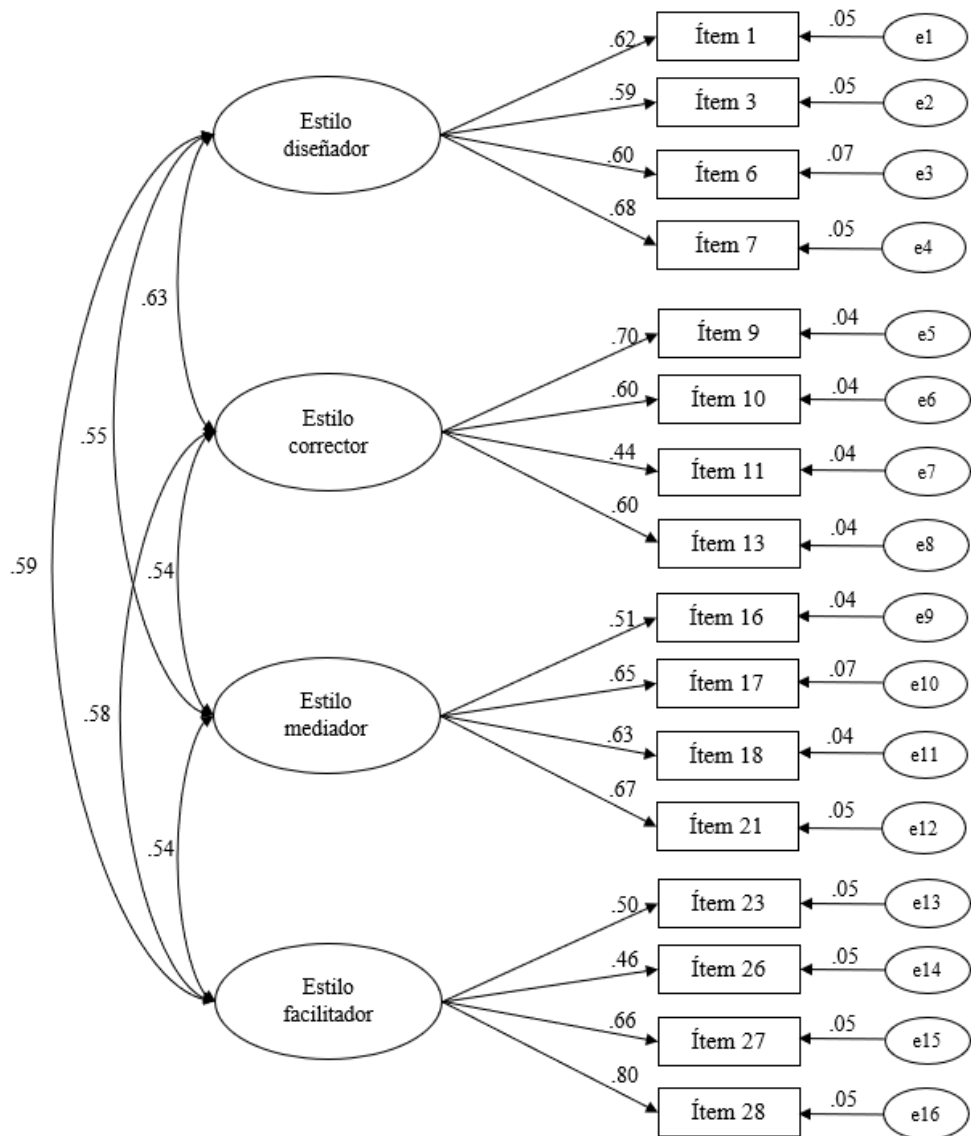
Análisis factorial confirmatorio

La literatura sugiere que la carga factorial de los ítems debe ser superior a .40, por lo que durante el proceso de análisis se eliminaron aquellos que no cumplían con este criterio, resultando en una escala final de cuatro dimensiones con cuatro ítems cada una (Brown, 2015; Hair et al., 2010; Kline, 2016).

Al realizar el AFC con la muestra del estudio (n = 240), los resultados evidenciaron un buen ajuste del modelo de medición, confirmando la estructura teórica ($X^2 = 126.48$ gl = 98, $p = .028$; SRMR = .05; AGFI = .92; TLI = .96; CFI = .97; RMSEA = .03 IC 90 % [.01, .05]). Las cargas factoriales para cada dimensión oscilaron entre: estilo diseñador, .59 - .68; estilo corrector, .44 - .70; estilo mediador, .51 - .67 y; estilo facilitador, .46 - .80, en todos los casos significativas ($p < 0.001$). Las correlaciones entre las dimensiones se encontraron entre .54 y .63 (Figura 2).

Figura 2

Resultado del AFC del modelo para medir estilos de enseñanza en línea



Invarianza media por sexo

Los hallazgos del análisis multigrupo indican que el modelo factorial mantiene su invarianza para los grupos comparados, considerando que los índices de ajuste muestran que el modelo base mantiene una estructura factorial similar para ambos grupos ($X^2 = 279.85$, $gl = 196$, $p = .000$; $SRMR = .08$; $CFI = .91$; $TLI = .88$; $RMSEA = .04$, 90 % CI [.03, .05]). Al igualarse las cargas factoriales de los grupos (invarianza métrica), se observó que las diferencias entre los valores de CFI y RMSEA fueron menores a .01 y .015 respectivamente, sugiriendo la presencia de invarianza métrica. Al añadir constricciones en los interceptos de los grupos (invarianza escalar), tampoco se encontraron diferencias en el CFI y RMSEA. Al comparar el modelo base con el modelo de invarianza residual, los valores también cumplieron con los criterios de igualdad (Tabla 3).

Tabla 3
Índices de bondad de ajuste de los modelos de invarianza

Medida	Modelos	X^2	gl	ΔX^2	Δgl	p	ΔCFI	$\Delta RMSEA$
Estilos de enseñanza en línea	M ₁ Invarianza configuracional	279.85	196					
	M ₂ Invarianza métrica	288.35	212	8.50	16	.933	.008	-.004
	M ₃ Invarianza escalar	309.53	228	21.18	16	.172	-.006	<.000
	M ₄ Invarianza residual	339.15	250	29.62	22	.128	-.008	<.000

Validez concurrente

El análisis de validez concurrente se realizó mediante una correlación de Pearson entre las dimensiones de la escala EEL y la felicidad subjetiva, medidas simultáneamente. Los resultados muestran una relación positiva y significativa ($p < .01$) entre el puntaje global de la escala EEL, sus dimensiones y la felicidad subjetiva. En particular, la dimensión de corrector presentó la correlación más alta ($r = .27$), seguida de diseñador ($r = .22$), facilitador ($r = .18$) y mediador ($r = .13$). La relación positiva entre ambas variables respalda la validez concurrente de la escala, indicando que los estilos de enseñanza en línea están asociados con la felicidad de los docentes (Tabla 4).

Tabla 4*Correlaciones entre la escala EEL, sus dimensiones y la felicidad subjetiva*

Variables	M	DE	1	2	3	4	5	6
1 Felicidad Subjetiva	3.38	0.59	-					
2 EEL Diseñador	3.20	0.58	.22**	-				
3 EEL Corrector	3.58	0.42	.27**	.45**	-			
4 EEL Mediador	3.46	0.42	.13**	.38**	.40**	-		
5 EEL Facilitador	3.28	0.55	.18**	.45**	.43**	.38**	-	
6 EEL Global	3.38	0.39	.26**	.78**	.72**	.72**	.76**	-

Nota: **= $p < .01$, EEL= Estilos de enseñanza en línea; M = Media; DE = desviación estándar.

Análisis de fiabilidad

La fiabilidad de la escala EEL se analizó por medio de los coeficientes omega de McDonald (ω) y alfa de Cronbach (α). Los resultados muestran que los valores de los dos estadísticos son aceptables (Katz, 2006) (Tabla 5).

Tabla 5*Coeficientes de fiabilidad de la escala EEL y sus dimensiones*

Escala estilos de enseñanza en línea	ω	α
EEL Global	.84	.84
EEL Diseñador	.71	.70
EEL Corrector	.68	.67
EEL Mediador	.71	.68
EEL Facilitador	.71	.68

Nota: EEL= Estilos de enseñanza en línea, α = Alfa de Cronbach, ω = Omega de McDonald

DISCUSIÓN

La escala EEL presenta propiedades psicométricas sólidas, lo que respalda su validez y fiabilidad como herramienta para evaluar los estilos de enseñanza en entornos digitales. Los resultados del análisis factorial confirmatorio mostraron un ajuste satisfactorio del modelo (CFI=.97, TLI=.96, RMSEA=.03), lo que indica una buena correspondencia entre la estructura teórica propuesta de cuatro dimensiones

(diseñador, corrector, mediador y facilitador) y los datos empíricos. Estos valores superan los umbrales mínimos recomendados para modelos de medida ($CFI \geq .95$, $TLI \geq .95$, $RMSEA \leq .06$), lo cual refuerza la validez de constructo de la escala (Hu y Bentler, 1999; Kline, 2016).

En comparación con investigaciones previas, los hallazgos coinciden con estudios que han validado instrumentos de evaluación en entornos digitales, como el Cuestionario de Estilos de Enseñanza de Grasha-Riechmann adaptado a contextos en línea (Abell et al., 2016) y el Inventario de Estilos de Enseñanza en Educación Superior (TSIHE) (Malay et al., 2024). Estos estudios también reportaron estructuras multidimensionales y ajustes satisfactorios, aunque con variaciones en las dimensiones específicas. Sin embargo, la inclusión de las dimensiones diseñador, corrector, mediador y facilitador en la escala de EEL amplía el marco teórico de referencia (Grasha, 1996; Mori et al., 2023; Romero-Félix, 2023; Sanje y Varnali, 2014), adaptándolo a las demandas pedagógicas de la educación en línea.

Por otro lado, los coeficientes de consistencia interna (alfa de Cronbach y omega de McDonald) oscilaron entre .67 y .71 (.84 global). Aunque estos valores se encuentran en el límite inferior de aceptabilidad, son considerados adecuados en instrumentos en etapas iniciales de validación (DeVellis, 2017). Si bien los coeficientes obtenidos son inferiores a los reportados en otras escalas de estilos de enseñanza (Abello et al., 2020; González-Peiteado et al., 2012), son comparables a los hallados en escalas como las de Arbabisarjou et al. (2020), Hurriyetoglu y Kilicoglu (2020) y Merino-Barrero et al. (2017). Sin embargo, el valor de consistencia interna global de la escala analizada fue superior al de estas últimas, lo que sugiere una mayor estabilidad en la puntuación total. Asimismo, los resultados son coherentes con los obtenidos en instrumentos diseñados para contextos digitales, donde la variabilidad en la interacción docente-estudiante puede afectar la homogeneidad de las respuestas (Scherer et al., 2019).

El análisis de validez concurrente aporta evidencia empírica sobre la relación entre los estilos de enseñanza en línea y la felicidad subjetiva de los docentes, lo que respalda la utilidad de la escala EEL para evaluar estos estilos en el contexto digital. Según Furr (2022), la validez concurrente se establece cuando un instrumento muestra asociaciones significativas con variables teóricamente relacionadas. En este sentido, la correlación positiva y significativa entre el puntaje global de la escala EEL y la felicidad subjetiva ($p < .01$) respalda su validez, sugiriendo que la adopción de determinados estilos de enseñanza en línea podría estar vinculada con el bienestar emocional de los docentes. Este hallazgo es consistente con investigaciones previas que han demostrado que las emociones de los docentes influyen en su comportamiento, sus estilos de enseñanza y sus funciones dentro del proceso formativo (Burel et al., 2021; Frenzel, 2014; Yirci et al., 2018).

En cuanto a la invarianza factorial por sexo, la escala demostró equivalencia configural, métrica, escalar y residual, lo que permite comparaciones significativas entre hombres y mujeres. Esto coincide con estudios que sugieren que las diferencias en los estilos de enseñanza no se deben tanto al género, sino a las expectativas del rol

docente (Eagly y Wood, 2012) o a sus distintos enfoques pedagógicos (Arabit García et al., 2021; Navarro-Patón et al., 2020). En este sentido, estudios anteriores señalan que existen diferencias de género en el desarrollo de las competencias digitales entre los docentes (Palacios-Rodríguez et al., 2025), posiblemente influenciadas por la percepción del rol de género en relación con la tecnología (Usart et al., 2021). Es por ello, que explorar estas diferencias con la escala EEL podría proporcionar información valiosa sobre la influencia del género en las prácticas pedagógicas en línea.

Implicaciones teóricas y prácticas

Los resultados de este estudio presentan importantes implicaciones tanto teóricas como prácticas. En términos aplicados, esta investigación amplía el estudio de los estilos de enseñanza más allá de los entornos presenciales, incorporando un análisis específico de su desarrollo en contextos virtuales. La pandemia de COVID-19 evidenció la necesidad de adaptar la enseñanza a modalidades en línea, resaltando la importancia de garantizar la calidad educativa a distancia. Factores como la retroalimentación, el acompañamiento y la supervisión de los estudiantes cobran un papel crucial cuando la interacción presencial es sustituida por plataformas tecnológicas. En este sentido, la evaluación del docente no solo debe considerar su dominio de la materia, sino también sus estrategias pedagógicas y preferencias en entornos digitales (Deseo et al., 2024).

Además, se ha encontrado que los estilos de enseñanza pueden estar influenciados por la preparación docente y su estado emocional. Durante la pandemia, la transición abrupta a la enseñanza en línea generó diversas afectaciones psicológicas tanto en docentes como en estudiantes (Cuadra-Martínez et al., 2020; Delgado, 2023), lo que refuerza la importancia de analizar las prácticas pedagógicas en línea.

Desde una perspectiva teórica, este estudio reafirma que la enseñanza en ambientes virtuales no es equivalente a la enseñanza presencial, por lo que los principios pedagógicos tradicionales no pueden aplicarse sin adaptaciones. Investigaciones previas han señalado que la docencia en línea requiere ajustes metodológicos específicos para garantizar la efectividad del aprendizaje y la interacción con los estudiantes (Garrison y Cleveland-Innes, 2010). Así, la escala EEL proporciona un marco adecuado para evaluar cómo los docentes ajustan sus prácticas pedagógicas a las exigencias de los entornos digitales, lo que puede servir de base para futuras investigaciones y programas de formación docente.

CONCLUSIONES

El presente estudio tuvo como objetivo analizar las propiedades psicométricas de una escala diseñada para medir los estilos de enseñanza en la educación en línea en docentes de educación superior. Los resultados de la investigación confirman

que la escala Estilos de Enseñanza en Línea es un instrumento psicométricamente sólido, con evidencias robustas de validez y fiabilidad para evaluar los estilos de enseñanza en entornos digitales. El análisis factorial confirmatorio reveló un ajuste adecuado del modelo de cuatro dimensiones ($CFI = .97$, $TLI = .96$, $RMSEA = .03$), lo que respalda la estructura teórica propuesta. La validez de contenido, establecida a través del juicio de expertos, mostró índices excepcionales ($S-CVI = .99$), mientras que los coeficientes de fiabilidad (α y ω entre $.67$ y $.71$) reflejan consistencia interna aceptable, considerando que se trata de una escala en etapa inicial de validación.

El análisis de invarianza factorial indicó que la escala es equivalente para los sexos masculino y femenino, permitiendo comparaciones válidas entre grupos. Asimismo, la validez concurrente mostró correlaciones positivas significativas con la felicidad subjetiva de los docentes, sugiriendo que los estilos de enseñanza en línea podrían estar vinculados al bienestar emocional del profesorado. Estos hallazgos refuerzan la relevancia de contar con instrumentos específicos para contextos virtuales, que permitan mejorar las prácticas pedagógicas y promover la calidad educativa en entornos digitales.

A pesar de la contribución de este estudio a la medición de los estilos de enseñanza en línea mediante una escala útil para la investigación en este campo, se deben tener en cuenta algunas limitaciones. En primer lugar, las respuestas del instrumento se obtuvieron mediante el autorreporte de los profesores, por lo que pudieron ser afectadas por la deseabilidad social. En segundo lugar, los participantes fueron docentes de un municipio del sur de Sonora, por lo que se requiere de una muestra con mayor tamaño y diversidad geográfica para poder generalizar los resultados. Además, este estudio consideró los estilos de enseñanza únicamente desde la perspectiva de los docentes, por lo que futuras investigaciones podrían indagar en la perspectiva de los estudiantes acerca del estilo de enseñanza de sus profesores para identificar posibles discrepancias y coincidencias.

Por tanto, los resultados de este estudio dan evidencia de una escala confiable y válida para evaluar los estilos de enseñanza en la educación en línea. El modelo teórico de cuatro dimensiones, desarrollado específicamente para el contexto digital, ofrece una medida precisa del constructo de estilos de enseñanza, adaptada a las particularidades de la educación en línea.

Finalmente, se propone un instrumento breve con evidencias de validez aceptables, que contribuye significativamente a la medición del constructo en docentes universitarios, lo que amplía la aplicabilidad y utilidad de la herramienta en la evaluación y mejora de la enseñanza en entornos digitales. Además, se abren nuevas vías para investigaciones futuras que busquen explorar más a fondo la influencia de los estilos de enseñanza en el desempeño académico y la experiencia de los estudiantes en plataformas en línea.

REFERENCIAS

- Abell, N., Cain, M. y Cheng-Yuan, C. L. (2016). Essential attributes for online success: Student learning preferences and faculty teaching styles. *International Journal on E-Learning*, 15(4), 401-422. <http://bit.ly/411UKAL>
- Abello, D. M., Tapia, J. A. y Panadero-Calderón, E. (2020). Development and validation of the Teaching Styles Inventory for Higher Education (TSIH). *Anales de Psicología/Annals of Psychology*, 36(1), 143-154. <https://doi.org/10.6018/analesps.370661>
- Ajzen, I. (2005). *Attitudes, personality, and behavior*. Open University Press.
- Almanasreh, E., Moles, R. y Chen, T. F. (2019). Evaluation of methods used for estimating content validity. *Research in Social and Administrative Pharmacy*, 15(2), 214-221. <https://doi.org/10.1016/j.sapharm.2018.03.066>
- American Educational Research Association, American Psychological Association y National Council on Measurement in Education (Eds.). (2014). *Estándares para pruebas educativas y psicológicas*. American Educational Research Association.
- Arabit García, J., Prendes Espinosa, M. P. y Serrano Sánchez, J. L. (2021). La enseñanza de STEM en Educación Primaria desde una perspectiva de género. *Revista Fuentes*, 23(1), 64-76. <https://doi.org/10.12795/revistafuentes.2021.v23.i1.12266>
- Arbabisarjou, A., Akbarilakeh, M., Soroush, F. y Payandeh, A. (2020). Validation and normalization of Grasha-Riechmann Teaching Style Inventory in faculty members of Zahedan University of Medical Sciences. *Advances in Medical Education and Practice*, 11, 305-312. <https://doi.org/10.2147/AMEP.S244313>
- Arenas-Martínez, R. y Ramírez-Martínez, I. (2024). Certezas e incertidumbres sobre el abandono estudiantil en la Universidad Autónoma Juan Misael Saracho durante la pandemia del COVID-19. *593 Digital Publisher CEIT*, 9(5), 37-46. <https://doi.org/10.33386/593dp.2024.5.2574>
- Arslan, O., Gülbahar, Y., Rapp, C. y Deniz, K. Z. (2014). Development and pilot implementation of online teaching style survey (OTSS). En J. Viteli y M. Leikomaa (Eds.), *Proceedings of EdMedia 2014-World Conference on Educational Media and Technology* (pp. 1530-1535). Association for the Advancement of Computing in Education (AACE). <https://www.learntechlib.org/primary/p/147682/>
- Bandalos, D. L. y Finney, S. J. (2019). Factor analysis: Exploratory and confirmatory. En G. R. Hancock y R. O. Mueller (Eds.), *The reviewer's guide to quantitative methods in the social sciences* (pp. 98-122). Routledge. <https://doi.org/10.4324/9781315755649-8>
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. W.H. Freeman.
- Bartholomew, K. J., Ntoumanis, N., Mouratidis, A., Katartzis, E., Thøgersen-Ntoumani, C. y Vlachopoulos, S. (2018). Beware of your teaching style: A school-year long investigation of controlling teaching and student motivational experiences. *Learning and Instruction*, 53, 50-63. <https://doi.org/10.1016/j.learninstruc.2017.07.006>
- Batista, G. D. P., Labatut Portilho, E. M. y Rufini, S. É. (2015). Evidence of validity for the Portilho/Banas teaching style questionnaire. *Paidéia (Ribeirão Preto)*, 25(62), 317-324. <https://doi.org/10.1590/1982-43272562201505>
- Blunch, N. J. (2013). *Introduction to structural equation modeling using IBM SPSS Statistics and AMOS*. SAGE. <https://doi.org/10.4135/9781526402257>

- Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). The Guilford Press.
- Burel, N., Tessier, D. y Langdon, J. (2021). Are teachers' subjective feelings linked with need-supportive and need-thwarting motivating styles? A cross-lagged pilot study in physical education. *European Journal of Psychology of Education*, 36(4), 1221-1241. <https://doi.org/10.1007/s10212-020-00517-x>
- Byrne, B. M. (2010). *Structural equation modeling with AMOS: Basic concepts, applications, and programming* (2nd ed.). Routledge/Taylor & Francis Group.
- Castillo Salcido, P. F., García Martínez, R., Hernández Gómez, J. A., Poblano Ojinaga, E. R. y Noriega Morales, S. A. (2024). Validez de contenido de un instrumento de medición para medir el éxito en la implementación del TPM. *Revista Iberoamericana para la Investigación y el Desarrollo Educativo*, 14(28), Artículo e674. <https://doi.org/10.23913/ride.v14i28.1928>
- Chan, D. W. (2009). Orientations to happiness and subjective well-being among Chinese prospective and in-service teachers in Hong Kong. *Educational Psychology*, 29(2), 139-151. <https://doi.org/10.1080/01443410802570907>
- Cuadra-Martínez, D., Castro-Carrasco, P. J., Sandoval-Díaz, J., Pérez-Zapata, D. y Mora Dabancens, D. (2020). COVID-19 y comportamiento psicológico: Revisión sistemática de los efectos psicológicos de las pandemias del siglo XXI. *Revista Médica de Chile*, 148(8), 1139-1154. <https://doi.org/10.4067/S0034-98872020000801139>
- Delgado, G. (2023). Efectos Psicológicos causados por la COVID-19 en los estudiantes con discapacidad y sus progenitores. *RETOS XXI*, 7(1), 1-13. <https://doi.org/10.30827/retosxxi.7.2023.24319>
- Deseo, R. S., Teaño, M. E. J., Steven, J. y Hayg, S. C. (2024). Mental health amidst pandemic: Predictor of online teaching style and readiness among educators. *International Journal of Research in Education Humanities and Commerce*, 5(5), 238-250. <https://doi.org/10.37602/IJREHC.2024.5517>
- DeVellis, R. F. (2017). *Scale development: Theory and applications* (4th ed.). Sage.
- Eagly, A. H. y Wood, W. (2012). Social role theory. En P. A. M. Van Lange, A. W. Kruglanski, y E. T. Higgins (Eds.), *Handbook of theories of social psychology* (Vol. 2, pp. 458-476). Sage. <https://doi.org/10.4135/9781446249222.n49>
- El-Soussi, A. (2022). The shift from face-to-face to online teaching due to COVID-19: Its impact on higher education faculty's professional identity. *International Journal of Educational Research Open*, 3, Article 100139. <https://doi.org/10.1016/j.ijedro.2022.100139>
- Espada, M., Fernández, M. y Calero, J. C. (2019). Validación de la versión española del cuestionario de uso y percepción del espectro de estilos de enseñanza en educación física. *Revista Mexicana de Investigación Educativa*, 24(80), 271-285. <https://bit.ly/3AXAsiK>
- Franker, K. A. (2016). *The influence of teaching style on instructor satisfaction with teaching fully online courses* (Publication No. 10124138) [Doctoral dissertation, Edgewood College]. ProQuest Dissertations and Theses Global. <https://bit.ly/3T4BlM4>
- Frenzel, A. C. (2014). Teacher emotions. En R. Pekrun y L. Linnenbrink-Garcia (Eds.), *International handbook of emotions in education* (pp. 494-519). Routledge.
- Funder, D. C. y Ozer, D. J. (2019). Evaluating effect size in psychological research: Sense and nonsense. *Advances in Methods and Practices in Psychological Science*, 2(2), 156-168. <https://doi.org/10.1177/2515245919847202>

- Furr, R. M. (2011). *Scale construction and psychometrics for social and personality psychology*. SAGE. <https://doi.org/10.4135/9781446287866>
- Furr, R. M. (2022). *Psychometrics: An introduction* (4th ed.). SAGE Publications.
- García-Aranda, P. C., Pérez-Polanco, P. O. y Canto-Herrera, P. J. (2017). Los profesores de bachillerato y sus estilos de uso del espacio virtual. *Revista de Estilos de Aprendizaje*, 10(19), 112-134. <https://doi.org/10.55777/rea.v10i19.1072>
- Garrison, D. R., Cleveland-Innes, M. y Fung, T. S. (2010). Exploring causal relationships among teaching, cognitive and social presence: Student perceptions of the community of inquiry framework. *The Internet and Higher Education*, 13(1-2), 31-36. <https://doi.org/10.1016/j.iheduc.2009.10.002>
- González, A., Conde, Á., Díaz, P., García, M. y Ricoy, C. (2018). Instructors' teaching styles: Relation with competences, self-efficacy, and commitment in pre-service teachers. *Higher Education*, 75(4), 625-642. <https://doi.org/10.1007/s10734-017-0160-y>
- González-Peiteado, M., López-Castedo, A. y Pino-Juste, M. (2012). Análisis psicométrico de una escala sobre estilos de enseñanza (ESEE). *Enseñanza & Teaching*, 31(1), 181-198. <https://bit.ly/3VVxaSD>
- Grasha, A. F. (1994). A matter of style: The teacher as expert, formal authority, personal model, facilitator, and delegator. *College Teaching*, 42(4), 142--149. <https://doi.org/10.1080/87567555.1994.9926845>
- Grasha, A. F. (1996). *Teaching with style: A guide to enhancing learning by understanding teaching and learning styles*. Alliance Publishers.
- Hair, J. F., Black, W. C., Babin, B. J. y Anderson, R. E. (2010). *Multivariate data analysis* (7th ed.). Pearson Prentice Hall.
- Hancock, G. R. y Liu, M. (2012). Bootstrapping standard errors and data-model fit statistics in structural equation modeling. En R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 296-306). The Guilford Press.
- Hidalgo-Cabrillana, A. y Lopez-Mayan, C. (2018). Teaching styles and achievement: Student and teacher perspectives. *Economics of Education Review*, 67, 184-206. <https://doi.org/10.1016/j.econedurev.2018.10.009>
- Hu, L. y Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, 6(1), 1-55. <https://doi.org/10.1080/10705519909540118>
- Hurriyetoglu, N. y Kilicoglu, E. (2020). Teaching styles scale: Validity and reliability study. *Educational Policy Analysis and Strategic Research*, 15(2), 222-237. <https://doi.org/10.29329/epasr.2020.251.12>
- Katz, M. H. (2006). *Multivariable analysis* (2nd ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9780511811692>
- Kebritchi, M., Lipschuetz, A. y Santiago, L. (2017). Issues and challenges for teaching successful online courses in higher education: A literature review. *Journal of Educational Technology Systems*, 46(1), 4-29. <https://doi.org/10.1177/0047239516661713>
- Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). Guilford Press.
- Krosnick, J. A. y Presser, S. (2010). Question and questionnaire design. En P. V. Marsden y J. D. Wright (Eds.), *Handbook of survey research* (2nd ed., pp. 264-313). Emerald.
- Laudadio, J. (2012). Evaluación de estilos de enseñanza en la universidad: Estudio preliminar de las propiedades

- psicométricas del Cuestionario sobre la Orientación Docente del Profesor Universitario (CODPU). *Interdisciplinaria*, 29(1), 79-93. <https://doi.org/10.16888/interd.2012.29.1.5>
- Leino, R. K., Gardner, M. R., Cartwright, T. y Döring, A. K. (2024). Engagement in a virtual learning environment predicts academic achievement in research methods modules: A longitudinal study combining behavioral and self-reported data. *Scholarship of Teaching and Learning in Psychology*, 10(2), 149. <https://doi.org/10.1037/stl0000281>
- Leung, K. K., Lue, B. H. y Lee, M. B. (2003). Development of a teaching style inventory for tutor evaluation in problem-based learning. *Medical Education*, 37(5), 410-416. <https://doi.org/10.1046/j.1365-2923.2003.01493.x>
- Lozano, L. M., García-Cueto, E. y Muñiz, J. (2008). Effect of the number of response categories on the reliability and validity of rating scales. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, 4(2), 73-79. <https://doi.org/10.1027/1614-2241.4.2.73>
- Lozano-Rodríguez, A. y Flores-Fahara, M. (2010). Online teaching styles: A study in distance education. *International Journal of University Teaching and Faculty*, 1(2), 1-14. <https://repositorio.tec.mx/handle/11285/578001?locale-attribute=es>
- Lyubomirsky, S., King, L. y Diener, E. (2005). The benefits of frequent positive affect: Does happiness lead to success? *Psychological Bulletin*, 131(6), 803-855. <https://doi.org/10.1037/0033-2909.131.6.803>
- Lyubomirsky, S. y Lepper, H. S. (1999). A measure of subjective happiness: Preliminary reliability and construct validation. *Social Indicators Research*, 46(2), 137-155. <https://doi.org/10.1023/A:1006824100041>
- Malay, J., Kassab, S. E., Merghani, T. H., Rathana, R. y Sreejith, A. (2022). Faculty and students' perceptions about online teaching styles of faculty in large group lectures. *Advances in Medical Education and Practice*, 13, 1261-1266. <https://doi.org/10.2147/AMEP.S377351>
- Malay, J., Rangraze, I. R., Merghani, T. H. y Kassab, S. E. (2024). Description of medical students' behavioral, cognitive, and psychological engagement with faculty online teaching styles. *Advances in Medical Education and Practice*, 15, 121-132. <https://doi.org/10.2147/AMEP.S448779>
- Merino-Barrero, J. A., Valero-Valenzuela, A. y Moreno-Murcia, J. A. (2017). Análisis psicométrico del Cuestionario Estilos de Enseñanza en Educación Física (EEEF). *Revista Internacional de Medicina y Ciencias de la Actividad Física y del Deporte*, 17(66), 225-241. <https://doi.org/10.15366/rimcafd2017.66.002>
- Mori, G., Bactindon, E., Dublado, K. K., Garcia, J. y Gimo, J. (2023). Online teaching styles and age, sex, and degree program: A correlational-descriptive study. *Ho Chi Minh City Open University Journal of Science - Social Sciences*, 13(1), 144-159. <https://doi.org/10.46223/HCMCOUJS.soci.en.13.1.2590.2023>
- Muthén, B. y Kaplan, D. (1985). A comparison of some methodologies for the factor analysis of non-normal Likert variables. *British Journal of Mathematical and Statistical Psychology*, 38(2), 171-189. <https://doi.org/10.1111/j.2044-8317.1985.tb00832.x>
- Navarro-Patón, R., Arufe-Giráldez, V., Mecías-Calvo, M. y Lago-Ballesteros, J. (2020). Análisis de las preferencias metodológicas del profesorado gallego de Educación Física en Educación Infantil y Primaria en función del género y los años de experiencia docente. *Publicaciones*, 50(3), 163-178. <https://doi.org/10.30827/publicaciones.v50i3.15775>

- Noguez Ortiz, J. C. (2024). Importancia de los estilos de aprendizaje como estrategia en la enseñanza en una universidad privada en México. *MLS Educational Research*, 8(2), 375-390. <https://doi.org/10.29314/mlser.v8i2.2301>
- Palacios-Rodríguez, A., Llorente-Cejudo, C., Lucas, M. y Bem-haja, P. (2025). Macroassessment of teachers' digital competence. DigCompEdu study in Spain and Portugal [Macroevaluación de la competencia digital docente. Estudio DigCompEdu en España y Portugal]. *RIED-Revista Iberoamericana de Educación a Distancia*, 28(1), 177-196. <https://doi.org/10.5944/ried.28.1.41379>
- Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y. y Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5), 879-903. <https://doi.org/10.1037/0021-9010.88.5.879>
- Polit, D. F., Beck, C. T. y Owen, S. V. (2007). Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Research in Nursing & Health*, 30(4), 459-467. <https://doi.org/10.1002/nur.20199>
- Rakha, A. H. (2023). The impact of Blackboard Collaborate breakout groups on the cognitive achievement of physical education teaching styles during the COVID-19 pandemic. *PLoS ONE*, 18(1), Article e0279921. <https://doi.org/10.1371/journal.pone.0279921>
- Romero-Félix, K. N. (2023). *Relaciones entre los estilos de enseñanza en línea y relaciones de cuidado en el clima del aula virtual* [Tesis doctoral, Instituto Tecnológico de Sonora]. Repositorio institucional del Instituto Tecnológico de Sonora. <http://sib2.itson.mx/cgi-bin/koha/opac-detail.pl?biblionumber=261851>
- Sanje, G. y Varnali, K. (2014). The effects of teaching style and internet self-efficacy on instructors' attitudes toward online education in higher education. *American International Journal of Contemporary Research*, 4(7), 35-43. <https://bit.ly/4gbQzYR>
- Scherer, R., Siddiq, F. y Tondeur, J. (2019). The technology acceptance model (TAM): A meta-analytic structural equation modeling approach to explaining teachers' adoption of digital technology in education. *Computers & Education*, 128, 13-35. <https://doi.org/10.1016/j.compedu.2018.09.009>
- Shaari, A. S., Yusoff, N. M., Ghazali, I. M., Osman, R. H. y Dzahir, N. F. M. (2014). The relationship between lecturers' teaching style and students' academic engagement. *Procedia-Social and Behavioral Sciences*, 118, 10-20. <https://doi.org/10.1016/j.sbspro.2014.02.002>
- Silva, B. D. (2017). Estilos de e-tutoria: Um estudo no âmbito da UAB no Ceará. *Revista de Estilos de Aprendizaje*, 10(19). <https://doi.org/10.55777/rea.v10i19.1075>
- Silva-Vera, F., Esteves-Fajardo, Z. I. y Melgar-Ojeda, K. A. (2023). Formación integral del estudiante: Análisis comparativo en modalidad presencial y virtual. *Revista Arbitrada Interdisciplinaria Koinonía*, 8(1), 172-191. <https://doi.org/10.35381/r.k.v8i1.2779>
- Soleimani, N. (2020). ELT teachers' epistemological beliefs and dominant teaching style: A mixed method research. *Asian-Pacific Journal of Second and Foreign Language Education*, 5, Article 1. <https://doi.org/10.1186/s40862-020-00094-y>
- Soussi, K. (2020). Web-based learning: Characteristics, practices, challenges, and recommendations. *International Journal of Science and Research*, 9(3), 936-943. <https://bit.ly/4oZFeVU>
- Sternberg, R. J. (1997). *Thinking styles*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511584152>

- Usart, M., Lázaro, J. L. y Gisbert, M. (2021). Validation of a tool for self-evaluating teacher digital competence. *Educación XX1*, 24(1), 353-373. <https://doi.org/10.5944/educxx1.27080>
- Valdés-Cuervo, A. A., García Vázquez, F. I., Torres Acuña, G. M., Urías Murrieta, M. y Grijalva Quiñonez, C. S. (2019). *Medición en investigación educativa con apoyo del SPSS y el AMOS*. Clave Editorial. <https://bit.ly/44FCluc>
- Valdez, L. A. (2022). *El clima social escolar y la agresión proactiva y reactiva en adolescentes: El rol mediador de la felicidad subjetiva* [Tesis de maestría, Instituto Tecnológico de Sonora]. Biblioteca del Instituto Tecnológico de Sonora. http://biblioteca.itson.mx/dac_new/tesis/2553.pdf
- Vieira Barros, D. M., Alonso García, C. y Ferreira do Amaral, S. (2008). Estilo de uso do espaço virtual. *Revista de Estilos de Aprendizaje*, 1(1), 88-108. <https://doi.org/10.55777/rea.viii.865>
- Vikas, S. y Mathur, A. (2022). An empirical study of student perception towards pedagogy, teaching style and effectiveness of online classes. *Education and Information Technologies*, 27(1), 589-610. <https://doi.org/10.1007/s10639-021-10793-9>
- Wang, S., Chiu, T. K. F., Zhao, L. y He, W. (2024). Teacher motivating styles and undergraduates' self-regulation of their online learning experience. *Distance Education*. Advance online publication. <https://doi.org/10.1080/01587919.2024.2383764>
- Yirci, R., Özalp, U. y Kocabas, I. (2018). Relationship between mentor teachers' mentoring functions and beginning teachers' subjective happiness. *Educational Administration: Theory & Practice*, 24(2), 387-435. <https://doi.org/10.14527/kuey.2018.010>
- Zante, B. y Klasen, J. M. (2021). Learner-centered education: ICU residents' expectations of teaching style and supervision level. *BMC Medical Education*, 21(1), Article 399. <https://doi.org/10.1186/s12909-021-02844-z>
- Zinn, L. M. (2004). Identifying your philosophical orientation. En M. W. Galbraith (Ed.), *Adult learning methods: A guide for effective instruction* (3rd ed., pp. 37-72). Krieger Publishing.

Fecha de recepción del artículo: 1 de diciembre de 2024

Fecha de aceptación del artículo: 7 de marzo de 2025

Fecha de aprobación para maquetación: 26 de marzo de 2025

Fecha de publicación en OnlineFirst: 9 de abril de 2025

Fecha de publicación: 1 de julio de 2025

Despertando el compromiso estudiantil: el poder transformador de H5P en la educación superior

Igniting student engagement: H5P's transformative potential in higher education



- FRIDEL JULIO RAMOS-AZCUY - Pontificia Universidad Católica del Ecuador, PUCE (Ecuador)
- MARÍA RODRÍGUEZ-GÁMEZ - Universidad Técnica de Manabí, UTM (Ecuador) y Pontificia Universidad Católica del Ecuador, PUCE (Ecuador)
- JEOVANNY MOISÉS BENAVIDES-BAILÓN - Universidad Técnica de Manabí, UTM (Ecuador) y Pontificia Universidad Católica del Ecuador, PUCE (Ecuador)
- MARÍA MARGOTH BONILLA-JIMÉNEZ - Pontificia Universidad Católica del Ecuador, PUCE (Ecuador)
- ÁNGEL ENRIQUE ARROBA-CÁRDENAS - Pontificia Universidad Católica del Ecuador, PUCE (Ecuador)

RESUMEN

El compromiso estudiantil, clave para el éxito en la educación en línea, exige estrategias efectivas para fomentarlo, particularmente en entornos híbridos. Este estudio evaluó el impacto de actividades interactivas con H5P en el compromiso de 87 estudiantes de pregrado de dos universidades ecuatorianas, mediante un diseño cuasi-experimental pretest-postest con grupos control y experimental. La recolección de datos se realizó a través de un cuestionario de 12 ítems que abarcaba las dimensiones cognitiva, afectiva y conductual del compromiso y de preguntas abiertas. Los resultados sugieren que la intervención con H5P mejoró significativamente la comprensión conceptual, la aplicación del conocimiento, la profundización del aprendizaje y el disfrute de las actividades, aunque no se observaron efectos significativos en la relevancia percibida ni en la colaboración. Estos hallazgos indican que la efectividad de H5P es limitada si no se integra en un diseño instruccional estratégico más amplio. Se recomienda complementar la herramienta con estrategias pedagógicas que promuevan activamente la colaboración, la conexión del contenido con los intereses del estudiante y la adaptación al contexto, incluyendo enfoques como el trabajo en equipo, la especialización de contenidos, los debates basados en el material interactivo y la co-creación. Investigaciones futuras deberían emplear diseños experimentales con asignación aleatoria, muestras más amplias y diversas, y mediciones multidimensionales del compromiso, para optimizar el uso de herramientas tecnológicas en la promoción efectiva y sostenible del compromiso estudiantil en diversos contextos educativos.

Palabras clave: H5P, compromiso estudiantil; enseñanza a distancia; elaboración de medios de enseñanza; tecnología de la educación.

ABSTRACT

Student engagement is a key factor for success in online education, and there is a persistent need to identify and implement effective strategies to foster it, particularly in the increasingly common hybrid learning environments. Addressing this need, the present study evaluated the impact of interactive activities, designed using the H5P tool, on the engagement levels of 87 undergraduate students from two Ecuadorian universities. A quasi-experimental pretest-posttest design was employed to compare control and experimental groups. Data were collected through a 12-item questionnaire assessing cognitive, affective, and behavioral dimensions of engagement, supplemented by open-ended questions to gather qualitative data. The H5P intervention significantly improved cognitive aspects, such as concept understanding, knowledge application, and perceived depth of learning, as well as enjoyment; however, it did not significantly affect content relevance or collaboration. These findings suggest that while H5P can be a valuable tool for fostering specific components of student engagement, particularly cognitive and affective engagement, its effectiveness is limited when considered in isolation. Therefore, to maximize its impact, it is crucial to complement H5P with additional pedagogical strategies that actively promote collaboration, critical thinking, and connect the learning material with students' existing interests, experiences, and real-world applications. H5P offers considerable potential in online education, but requires a pedagogically informed, context-sensitive, and holistic approach. Future research is strongly recommended, employing rigorous experimental designs, larger and more diverse sample sizes, and multidimensional measurements of engagement, to provide deeper insights into optimizing the use of technology to effectively and sustainably foster all dimensions of student engagement, leading to improved learning outcomes.

Keywords: H5P; engagement; distance study; creation of teaching aids; educational technology.

INTRODUCCIÓN

El proceso de enseñanza-aprendizaje virtual en la educación superior enfrenta el desafío de mantener el interés y el compromiso de los estudiantes. Factores como la estructura del curso, el instructor, el estilo de enseñanza, el ambiente de aprendizaje y las características del alumnado influyen decisivamente en su compromiso y sus resultados (Mazman Akar, 2024; Taylor et al., 2018). Este desafío demanda estrategias pedagógicas innovadoras que integren tecnologías educativas y un diseño instruccional efectivo. Dicha integración debe facilitar el aprendizaje activo y la reflexión metacognitiva del estudiante a través de retroalimentación oportuna y específica sobre sus fortalezas y debilidades (Amhag, 2020; Baleni, 2015; Morris et al., 2021; Rahmi et al., 2024; Theelen y van Breukelen, 2022).

La innovación en la educación en línea debe centrarse en estrategias pedagógicas que promuevan la construcción activa del conocimiento, alineado con los principios de las principales teorías pedagógicas del aprendizaje. En este sentido, Lamtara (2023) propone combinar actividades como la gamificación, a través de elementos lúdicos y desafíos, en una estrategia pedagógica híbrida dirigida a incrementar la motivación y el compromiso de los estudiantes, estimulando procesos cognitivos como la atención, la memoria y la resolución de problemas. Lo anterior, en línea con los postulados del cognitivismo, está enfocado en atender las diferentes formas de aprender de los estudiantes, para favorecer el aprendizaje significativo (Méndez Mantuano et al., 2021; Parson y Major, 2020).

Asimismo, al crear experiencias de aprendizaje significativas y relevantes, se fomenta un mayor compromiso con el aprendizaje, mientras que la variedad de actividades y formatos evita la monotonía y mantiene a los estudiantes concentrados (Kang y Furtak, 2021). Esto fomenta el desarrollo de habilidades del siglo XXI, la preparación efectiva para el mundo laboral y la formación de estudiantes reflexivos de por vida con motivación para aprender (Bailey et al., 2021; Bajaber, 2024; Kuh, 2009).

Lo propuesto hasta este punto responde a un enfoque cognitivista que puede enriquecerse mediante estrategias pedagógicas como el aprendizaje basado en proyectos o en problemas, así como con actividades que promuevan la colaboración para facilitar la construcción de redes de conocimiento, al conectar ideas y conceptos previos de manera significativa, lo cual es un principio clave del conectivismo (Downes, 2022; Ortiz y Corrêa, 2020; Safarifard et al., 2024) y de las teorías socioculturales del aprendizaje (Kang y Furtak, 2021).

En este contexto, H5P (HTML 5 Package) emerge como una herramienta digital innovadora, que se alinea perfectamente con este enfoque al fomentar precisamente la interactividad y la participación activa. Integrable en plataformas como Moodle, permite a los docentes diseñar actividades diversas, incluyendo cuestionarios, presentaciones y videos interactivos con retroalimentación inmediata.

En el metaanálisis realizado por Ploetzner (2024), sobre la efectividad de los videos con interacción mejorada en el aprendizaje, se encontró que estos tienen una mayor efectividad en comparación con los videos que solo integran funciones de navegación. Esto se fundamenta, desde una perspectiva fisiológica, por Gellisch et al. (2023) cuando plantean que las actividades interactivas en la educación en línea generan respuestas psicobiológicas más fuertes en los estudiantes en comparación con actividades con poca o nula interactividad. Este nivel de respuestas psicobiológicas “se asocian a una mayor atención y mayores niveles de compromiso tanto con el trabajo del curso como con el material de aprendizaje” (Gellisch et al., 2023, p. 11).

Fredricks et al. (2004) plantean que el compromiso estudiantil es un constructo multidimensional que incluye aspectos comportamentales, cognitivos y emocionales. El compromiso implica conductas observables como la participación en clase, la persistencia en tareas desafiantes y una conexión emocional con el aprendizaje. Por su parte, Kuh (2009) hace énfasis en la importancia de las experiencias extracurriculares y la conexión con la comunidad universitaria para fomentar el compromiso. Tanto Fredricks et al. (2004) como Kuh (2009) confirman la correlación positiva entre el compromiso y el rendimiento y sugieren que un mayor compromiso mejora la probabilidad de los estudiantes para alcanzar sus objetivos educativos y personales.

En este estudio, el compromiso se define como

la energía y el esfuerzo que emplean los estudiantes dentro de su comunidad de aprendizaje, observable a través de una serie de indicadores conductuales, cognitivos o afectivos a lo largo de un continuo. Está determinado por una serie de influencias estructurales e internas, incluida la compleja interacción de las relaciones, las actividades de aprendizaje y el entorno de aprendizaje. (Bond et al., 2020, p. 3)

Para Fredricks et al. (2004), el compromiso se articula a través de tres dimensiones interrelacionadas: conductual, afectivo y cognitivo. De tal manera, el compromiso conductual se relaciona con la participación y el esfuerzo, con la implicación de los estudiantes en las actividades de aprendizaje (Fredricks et al., 2004). El compromiso afectivo se vincula a las relaciones de los estudiantes con su entorno educativo, incluyendo a sus pares y profesores. Esta dimensión del compromiso abarca las expectativas, motivaciones y suposiciones de los estudiantes en relación con su aprendizaje (Redmond et al., 2018) y está asociada con determinados elementos que impactan en la motivación para aprender como el sentido de pertenencia y ciertos estados emocionales (Abdool et al., 2017; Ali et al., 2020; Mulrooney y Kelly, 2020). El compromiso cognitivo está relacionado con el esfuerzo que realizan los estudiantes para comprender conceptos e ideas complejos incluyendo la reflexión sobre su propio aprendizaje (Fredricks et al., 2004, 2016).

Con el propósito de optimizar el uso de tecnologías educativas para potenciar la correlación positiva entre el nivel de compromiso de los estudiantes y sus resultados

de aprendizaje, Chi y Wylie (2014) establecieron una taxonomía de cuatro niveles de compromiso (interactivo, constructivo, activo y pasivo) aplicable a la educación en línea. Cada uno de los niveles responde a un conjunto de procesos subyacentes a la construcción activa del conocimiento.

El compromiso pasivo se caracteriza por la mera recepción de la información, donde el estudiante no realiza acciones adicionales como tomar notas o hacer preguntas. En contraste, el compromiso activo implica que el estudiante interactúe físicamente o manipule la información de alguna forma, evidenciando un mayor involucramiento. El compromiso constructivo se alcanza cuando, mediante un procesamiento cognitivo más profundo, el estudiante genera nuevos conocimientos o productos —por ejemplo, explicando los conceptos en sus propias palabras, creando representaciones visuales o formulando preguntas que permitan profundizar en el tema. Finalmente, el compromiso interactivo se produce cuando dos o más estudiantes dialogan y colaboran, construyendo de manera conjunta nuevas ideas y enriqueciendo el proceso de aprendizaje a través del intercambio de perspectivas.

Los niveles de compromiso de Chi y Wylie (2014) son de gran importancia como guías del diseño instruccional, así como para establecer condiciones de control óptimas para estudios experimentales. Además, en cada uno de estos niveles es posible incrementar la profundidad del aprendizaje mediante el aumento de la integración de herramientas tecnológicas.

En este sentido, Puentedura (2014) propone un modelo para categorizar el grado de esta integración en las categorías: sustitución, aumento, modificación y redefinición. Si al integrar tecnología no se produce un cambio funcional en la actividad de aprendizaje, se considera que esta integración es una sustitución. Si la actividad de aprendizaje se mantiene igual pero la integración tecnológica aporta una mejora funcional, la integración se considera un aumento. Por su parte, si la integración tecnológica implica el rediseño de la actividad de aprendizaje, se considera una modificación. Por último, si la integración conlleva a la creación de actividades innovadoras, se considera una redefinición.

Esta interrelación entre los modelos de Chi y Wylie (2014) y Puentedura (2014) devela la amplia gama de combinaciones tecnológicas que posibilitan profundizar en el aprendizaje considerando que el estudiante mantenga el mismo nivel de compromiso. De tal forma, la versatilidad y variedad de componentes interactivos de H5P la consolidan como una herramienta clave para implementar estrategias pedagógicas innovadoras en el proceso formativo.

Por otra parte, Bond et al. (2020) realizaron un profundo y exhaustivo estudio de los instrumentos utilizados para explorar el impacto de la tecnología educativa en el compromiso estudiantil. Así, identificaron un conjunto de criterios que permiten evaluar este compromiso a partir de las tres dimensiones que lo integran (Tabla 1).

Tabla 1
Criterios para evaluar el compromiso estudiantil

Dimensión	Criterios	Descripción
Cognitiva	Aprendizaje Profundo, Autorregulación	Relacionados con el pensamiento, la comprensión y la gestión del propio aprendizaje.
Afectiva	Interés, Motivación, Entusiasmo, Actitud positiva hacia el aprendizaje, Disfrute	Relacionados con las emociones, las actitudes y los sentimientos hacia el aprendizaje.
Conductual	Participación/Interacción/ Involucramiento, Logro, Interacción positiva con profesores y compañeros, Aprendizaje de pares	Relacionados con las acciones y comportamientos observables del estudiante.

Nota: elaboración propia a partir de Bond et al. (2020).

El aprendizaje profundo se trata de ir más allá de la simple memorización, es decir, involucra la comprensión de conceptos de manera significativa (Finn y Zimmer, 2012). Mediante la autorregulación el estudiante demuestra su capacidad para reflexionar acerca de las actividades y estrategias implementadas para su aprendizaje (Cleary y Zimmerman, 2012). Estos criterios se refuerzan mutuamente y el grado en que se presenten en los estudiantes indican el nivel de su compromiso cognitivo.

Por otra parte, el compromiso afectivo, se pone de manifiesto mediante emociones y actitudes interrelacionadas que incluyen el interés, entendido como la curiosidad y relevancia que se atribuye a una tarea. También incluyen la motivación intrínseca, que impulsa a la acción desde un deseo genuino; el entusiasmo, expresado en alegría y excitación; una actitud favorable hacia el aprendizaje, basada en la confianza en las propias capacidades; y el disfrute, que evidencia la satisfacción y el placer derivados de la actividad, consolidando así el vínculo emocional con el aprendizaje (Brookfield, 2009; Reeve, 2012; Skinner y Pitzer, 2012).

Finalmente, la participación regular en clase, la realización de las actividades y la interacción con los materiales por parte de los estudiantes demuestran compromiso conductual. También la calidad de las tareas, la persistencia para superar obstáculos y el logro de metas académicas; así como la iniciativa para aclarar dudas, el respeto hacia los demás y hacia las normas de la clase, la disposición para el trabajo colaborativo y el intercambio de ideas y perspectivas con los compañeros para enriquecer el aprendizaje son indicadores de compromiso conductual (Martín y Borup, 2022; Pekrun y Linnenbrink-García, 2012).

Varios estudios (Jacob y Centofanti, 2024; Rossetti-López et al., 2023; Sharmin et al., 2024) analizan la percepción de los estudiantes sobre el uso de actividades

interactivas creadas con H5P en Moodle en los que se encontró una valoración positiva de estos recursos, especialmente en términos de facilitar el aprendizaje y mantener la atención. Por su parte, el estudio de Jacob y Centofanti (2024) se propuso evaluar si la implementación de actividades de aprendizaje, enriquecidas con componentes interactivos de H5P, podía mejorar los resultados de aprendizaje en estudiantes de pregrado, no pudiendo aportar evidencia de mejoría en el rendimiento de los estudiantes asignable a estas. También Sinnayah et al. (2021) condujeron una investigación dirigida a explorar el uso de H5P como plataforma para fomentar el aprendizaje autodirigido en estudiantes de fisiología donde encontraron un elevado nivel de compromiso con el aprendizaje. Además, la mayoría de los estudiantes afirmó que su conocimiento mejoró gracias a la práctica repetida, facilitada por las funcionalidades de H5P.

Los estudios analizados, con la excepción de Jacob y Centofanti (2024), utilizan metodologías no experimentales, por lo que les resulta complicado relacionar clara y directamente los hallazgos con las distintas implementaciones pedagógicas que se valen de H5P para enriquecer la experiencia de aprendizaje. Además, en la literatura no se aprecian suficientes estudios con diseño experimental que relacionen la integración de H5P en el proceso de aprendizaje con una mejora en el compromiso de los estudiantes. Por lo tanto, este estudio tuvo como objetivo evaluar el impacto de las actividades diseñadas con H5P en el compromiso de estudiantes universitarios en cursos híbridos. El estudio se realizó en dos universidades ecuatorianas con programas de estudio principalmente presenciales, aunque con algunas asignaturas híbridas que carecían de actividades interactivas H5P. Para guiar la investigación, se planteó la siguiente hipótesis: los estudiantes que participan en actividades interactivas basadas en H5P, integradas en una metodología pedagógica innovadora y enriquecidas con medios audiovisuales, muestran un mayor compromiso con su aprendizaje que aquellos que no participan en dichas actividades.

MÉTODO

Diseño de la investigación

La investigación realizada fue de tipo mixta, longitudinal y cuasi-experimental de grupos intactos, lo que permitió viabilizar el estudio en un contexto en el que los grupos estaban previamente establecidos y no era factible la asignación aleatoria. Se optó por la ruta de investigación mixta dada la necesidad de contar con datos numéricos y cualitativos para evaluar el efecto de esta intervención. A través de un diseño pretest-posttest, se comparó el nivel de compromiso de los estudiantes antes y después de la implementación de las actividades H5P distribuidos en grupos de control y experimentales.

Teniendo en cuenta la complejidad del diseño longitudinal de la investigación, se seleccionaron dos profesores de cada universidad con dos grupos de segundo y

tercer año que impartían su asignatura en la modalidad híbrida con una frecuencia de una hora semanal. Lo anterior permitió garantizar la viabilidad del estudio y el cumplimiento del cronograma al facilitar la coordinación, la recopilación de datos y el seguimiento de los participantes a lo largo del tiempo. Además, la similitud en la media de las edades de los estudiantes permitió mantener un nivel de control adecuado pues estos compartían características similares en términos de madurez cognitiva y experiencias académicas. También, se tuvo en cuenta que ambos profesores tuvieran similares niveles de experiencia en el desarrollo de actividades interactivas con H5P y en el proceso de enseñanza-aprendizaje mediado por tecnologías educativas para minimizar el impacto de variables relacionadas con el docente.

Participantes

La muestra estuvo conformada por 87 estudiantes universitarios, de los cuales el 76 % eran mujeres. Los participantes provinieron de dos universidades de Portoviejo, Manabí: la Pontificia Universidad Católica del Ecuador Sede Manabí (PUCESM) y la Universidad Técnica de Manabí (UTM). Específicamente, se incluyeron 40 estudiantes de segundo año de las carreras Derecho y Negocios Internacionales de la PUCESM, media = $(19,4 \pm 0,51)$ años y 47 estudiantes de tercer año de Ingeniería Eléctrica de la UTM media = $(20,5 \pm 0,7)$ años.

Instrumentos

Para evaluar el compromiso estudiantil, se diseñó un cuestionario de 12 ítems basado en el constructo teórico de compromiso aportado por Fredricks et al. (2004) y enriquecido en Fredricks et al. (2016), además, en los indicadores que proponen Bond y Bedenlier (2019) para medir las dimensiones cognitiva, afectiva y conductual del compromiso. El cuestionario incluyó cuatro ítems para cada dimensión del compromiso (cognitiva, afectiva y conductual). Cada ítem se correspondía con una declaración afirmativa que el estudiante debía calificar usando una escala Likert de 5 niveles.

Para medir la dimensión cognitiva el cuestionario incluía los ítems siguientes:

- Entiendo claramente los conceptos que se explican en este curso (CC1).
- Soy capaz de aplicar lo que aprendo en nuevas situaciones (CC2).
- Las actividades que realizo me permiten profundizar en mi conocimiento (CC3).
- Me siento seguro/a al responder preguntas sobre el contenido (CC4).

Para la dimensión afectiva:

- Disfruto participando en las actividades de este curso (CA1).
- Me siento motivado/a a aprender los contenidos (CA2).
- Creo que los temas que se tratan son relevantes para mí (CA3).
- Estoy interesado/a en aprender más sobre este contenido (CA4).

Y para la dimensión conductual:

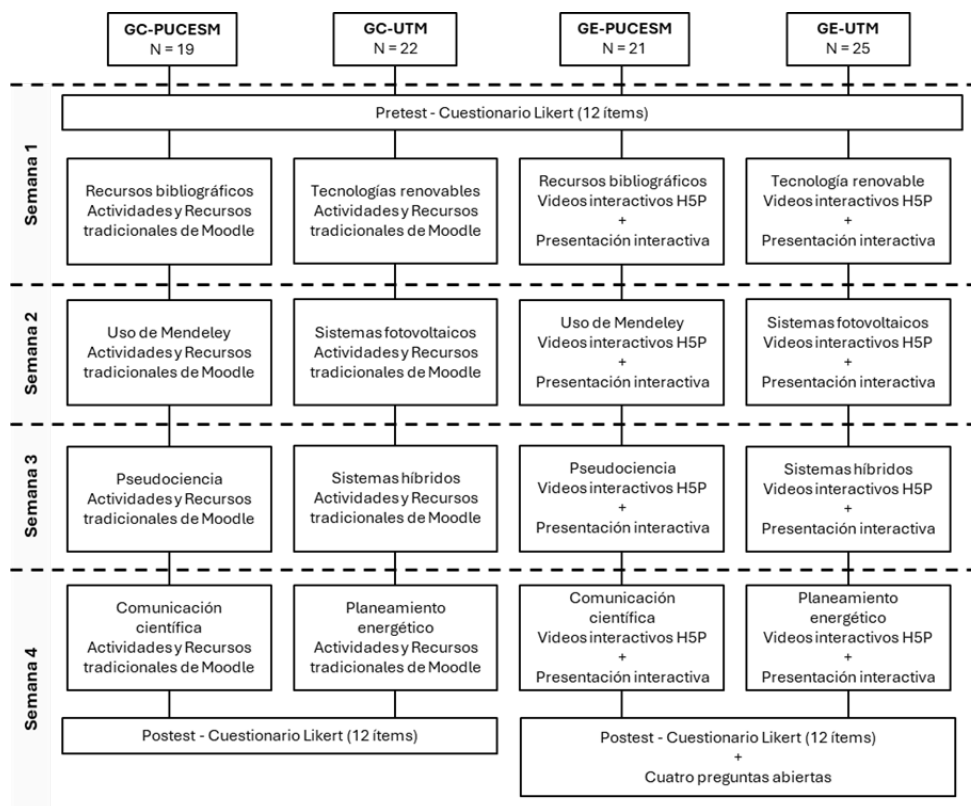
- Participo activamente en las clases y actividades (CCond1).
- Dedico tiempo a estudiar fuera de clase (CCond2).
- Colaboro con mis compañeros en las tareas (CCond3).
- Busco oportunidades para aprender más por mi cuenta (CCond4).

Procedimiento

El estudio involucró a un total de cuatro grupos de estudiantes: dos de la PUCESM (cursando la materia Fundamentos de Investigación) y dos de la UTM (cursando la materia Fuentes Renovables de Energía). La materia se impartió a cada grupo en la modalidad híbrida mediante la plataforma Moodle. En cada universidad se asignó aleatoriamente uno como grupo de control (GC) y el otro como grupo experimental (GE). En total, participaron 46 estudiantes en los grupos experimentales y 41 en los grupos de control. Al inicio del estudio, se administró el cuestionario de 12 ítems a todos los participantes para evaluar su nivel inicial de compromiso en las dimensiones cognitiva, afectiva y conductual (Figura 1).

Figura 1

Procedimiento experimental para evaluar el impacto de las actividades H5P en el compromiso estudiantil



Durante cuatro semanas los grupos experimentales participaron en un proceso de enseñanza-aprendizaje basado en el uso intensivo de recursos digitales interactivos desarrollados en H5P. Cada semana, se implementó un módulo de aprendizaje en Moodle que incluía varios videos y presentaciones interactivas desarrollados con H5P para cubrir los objetivos del tema de la clase (Figura 1). Si bien la duración del estudio permitió la viabilidad de este, se reconoce que es limitada, pudiendo ser insuficiente para observar efectos a largo plazo en el compromiso de los estudiantes.

Para asegurar la consistencia metodológica de las intervenciones en los grupos experimentales, se siguió la estrategia de abordar los aspectos teóricos de cada uno de los cuatro temas tratados mediante microaprendizaje con videos interactivos H5P incluyendo al menos tres actividades interactivas en cada uno y abordar los aspectos relacionados con la aplicación de la teoría a través de, por lo menos, una presentación

interactiva para la resolución de problemas. Los grupos de control, por su parte, realizaron actividades tradicionales en Moodle, diseñadas para ser equivalentes en términos de contenido y duración a las realizadas por los grupos experimentales, pero sin la integración de H5P.

Al finalizar la intervención, se administró el cuestionario, enriquecido con cuatro preguntas abiertas, para evaluar los cambios en el nivel de compromiso. Las tres primeras preguntas fueron diseñadas para obtener información cualitativa de cada dimensión del compromiso y la última para obtener información general del curso hasta el momento de la intervención:

- ¿Qué tipo de actividades consideras que te han ayudado a consolidar mejor los conocimientos adquiridos en estas semanas?
- ¿Qué aspectos de las actividades realizadas en clase han despertado tu curiosidad y te han motivado a investigar más sobre los temas tratados?
- ¿Cómo crees que las actividades que hemos realizado en clase han contribuido a tu aprendizaje colaborativo y a tu participación activa en el curso?
- ¿Hay algo más que quieras compartir sobre tu experiencia en este curso?

Análisis de Datos

Para el análisis de datos, se utilizó SPSS 25 y R (versión 4.4.3). La fiabilidad del instrumento se evaluó mediante el α de Cronbach, y la validez de contenido mediante la V de Aiken (con participación de ocho expertos). La normalidad se verificó con la prueba de Shapiro-Wilk. En el análisis descriptivo, se calcularon la mediana y el rango intercuartílico (IQR) para cada ítem (pretest y posttest). Para evaluar el efecto de la intervención, se compararon las puntuaciones pretest y posttest mediante la prueba de rangos con signo de Wilcoxon (función `wilcox.test()` en R). Se reportan las medianas pretest y posttest, el estadístico Z, el p-valor bilateral, y las sumas de rangos positivos (R+) y negativos (R-). El tamaño del efecto (r) se calculó como $r = Z/\sqrt{N}$, donde N es el número de pares. Dado que la diferencia se calculó como (pretest – posttest), un valor negativo de r indica que, en promedio, las puntuaciones en el posttest fueron superiores a las del pretest. La magnitud del efecto se interpretó como: despreciable ($|r| < 0,1$), pequeño ($0,1 \leq |r| < 0,3$), moderado ($0,3 \leq |r| < 0,5$) o grande ($|r| \geq 0,5$).

RESULTADOS

Estadísticos descriptivos

Se calculó la validez de contenido mediante la V de Aiken, obteniendo un valor general de 0,92. La fiabilidad del cuestionario, medida a través del alfa de Cronbach, reportó valores de entre 0,70 y 0,83 (Tabla 2).

Tabla 2

Descripción estadística del compromiso estudiantil antes y después de la intervención

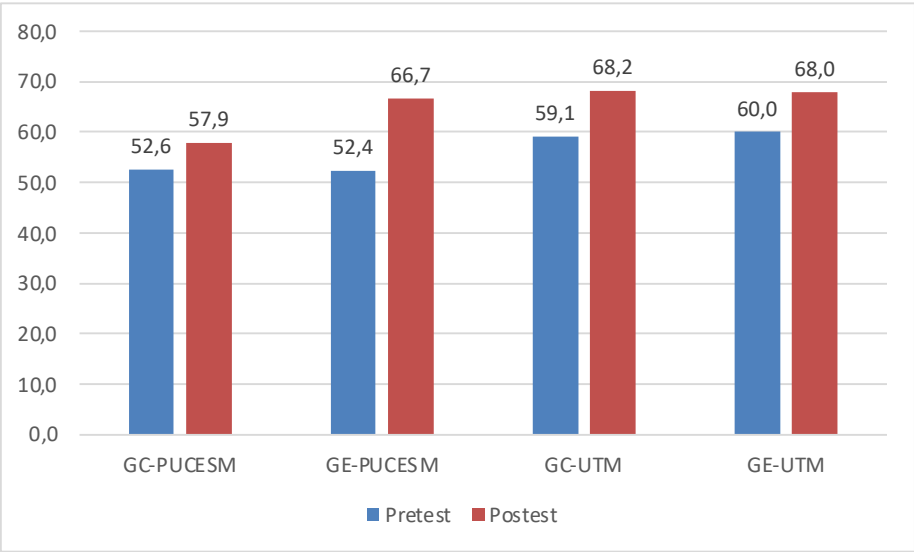
Grupo	Momento	Esta- dístico medido	Criterios para evaluar el compromiso estudiantil											
			Cognitivo				Afectivo				Conductual			
			CC1	CC2	CC3	CC4	CA1	CA2	CA3	CA4	CCondi	CCondi2	CCondi3	CCondi4
Control PUCESM N = 19	Pretest (0,79 ^a)	Mediana	4	4	4	4	4	4	4	4	4	4	4	4
		IQR	1,0	1,0	2,0	0,5	1,5	1,0	1,5	1,0	0,0	1,5	2,0	2,0
		Alfa de C.	0,68				0,72				0,79			
	Postest (0,81 ^a)	Mediana	4	4	4	4	4	4	4	4	3	4	4	5
		IQR	1,0	2,0	1,5	1,0	1,0	1,5	1,0	1,5	1,5	2,0	1,0	1,0
		Alfa de C.	0,81				0,72				0,69			
Experi- mental PUCESM N = 21	Pretest (0,86 ^a)	Mediana	4	4	4	4	4	4	4	4	4	4	4	4
		IQR	1,0	1,0	1,0	1,0	1,0	1,0	1,0	2,0	1,0	1,0	1,0	2,0
		Alfa de C.	0,76				0,74				0,73			
	Postest (0,74 ^a)	Mediana	4	5	5	5	5	4	4	4	4	4	5	5
		IQR	1,0	1,0	1,0	1,0	1,0	1,0	1,0	2,0	1,0	1,0	1,0	1,0
		Alfa de C.	0,69				0,75				0,75			
Control UTM N = 22	Pretest (0,76 ^a)	Mediana	4	4	4	4	4	4	3,5	4	4	4	4	4
		IQR	2,0	2,0	1,0	1,0	1,75	1,0	2,0	1,0	1,75	1,0	0,0	2,0
		Alfa de C.	0,73				0,71				0,72			
	Postest (0,78 ^a)	Mediana	4	4,5	4	4	4	4	4	4	4	4	4	4
		IQR	1,0	1,0	1,5	1,0	1,0	1,0	1,5	1,75	1,0	1,0	1,75	1,0
		Alfa de C.	0,74				0,70				0,74			

Grupo	Momento	Esta- dístico medido	Criterios para evaluar el compromiso estudiantil											
			Cognitivo				Afectivo				Conductual			
			CC1	CC2	CC3	CC4	CA1	CA2	CA3	CA4	CCond1	CCond2	CCond3	CCond4
Experi- mental UTM N = 25	Pretest (0,73 ^a)	Mediana	4	4	4	4	4	4	4	4	4	4	5	4
		IQR	1,0	2,0	1,0	1,0	1,0	2,0	1,0	1,0	1,0	2,0	1,0	1,0
		Alfa de C.	0,75				0,78				0,71			
	Postest (0,70 ^a)	Mediana	4	5	5	4	5	4	4	4	5	4	5	4
		IQR	1,0	1,0	1,0	1,0	1,0	2,0	1,0	1,0	1,0	1,0	1,0	1,0
		Alfa de C.	0,74				0,74				0,74			

Nota: ^a Coeficiente Alfa de Cronbach calculado para todos los ítems del cuestionario.

Los resultados del análisis descriptivo revelan una alta participación de los estudiantes en ambas aplicaciones del cuestionario. Al comparar los resultados del pretest y postest, se observa un aumento en el número de estudiantes que expresaron acuerdo o total acuerdo con los ítems del cuestionario en todos los grupos (Figura 2).

Figura 2
Porcentaje de acuerdo entre grupos experimentales y de control en la aplicación del cuestionario



Nota: Se muestra el menor valor del porcentaje para cada grupo.

Estadísticos Inferenciales

Se llevó a cabo un análisis inferencial comparando las puntuaciones obtenidas en el pretest y el postest para cada ítem del cuestionario, en ambos grupos (control y experimental). El test de Shapiro-Wilk confirmó que los datos no seguían una distribución normal, por tanto, se utilizó la prueba de los rangos con signo de Wilcoxon para muestras relacionadas. Adicionalmente, se calculó el tamaño del efecto para cuantificar la magnitud de las diferencias encontradas y determinar su relevancia práctica. Los resultados de ambas pruebas se presentan en la Tabla 3.

Los resultados de las cuatro preguntas abiertas que integraron el cuestionario utilizado en el postest a los GE arrojaron que el 96 % de los estudiantes respondieron la primera pregunta y muchos valoraron en mayor grado las actividades que les permitieron aplicar los conocimientos de forma práctica y recibir retroalimentación inmediata. El 93 % respondió la segunda pregunta, varios expresaron que la personalización y la interactividad de los contenidos fueron factores clave para despertar su curiosidad y motivación; y algunos sugirieron más ejercicios que involucraran casos de la vida cotidiana. El 87 % respondieron la tercera pregunta, entre ellos identificaron la necesidad de más actividades que fomenten el trabajo en equipo. Por último, los que respondieron la cuarta pregunta realizaron diversas sugerencias para mejorar la experiencia de aprendizaje: como mayor contacto sincrónico con los docentes, personalizar el aprendizaje y resolver problemas técnicos.

Tabla 3
Análisis inferencial pre-post del compromiso estudiantil en grupos de control y experimental mediante prueba de Wilcoxon y tamaño del efecto (r de Rosenthal)

Criterios para evaluar el compromiso estudiantil		Grupo															
		Control PUCESM (N = 19)				Experimental PUCESM (N = 21)				Control UTM (N = 22)				Experimental UTM (N = 25)			
		Z	R-/R+	p-valor	r	Z	R-/R+	p-valor	r	Z	R-/R+	p-valor	r	Z	R-/R+	p-valor	r
Cognitivo	CC1	-1,508	60 130	0,132	-0,35	-2,132	55,5 175,5	0,033	-0,47	-0,755	103,5 149,5	0,450	-0,16	-2,257	78,0 247,0	0,024	-0,45
	CC2	-0,984	72,5 117,5	0,325	-0,23	-2,034	57,5 173,5	0,042	-0,44	-1,127	97,5 155,5	0,260	-0,24	-2,173	90,5 234,5	0,030	-0,43
	CC3	-1,195	66,5 123,5	0,232	-0,27	-2,132	48,0 183,0	0,033	-0,47	-0,943	94,0 159,0	0,346	-0,20	-2,985	54,5 270,5	0,003	-0,60
	CC4	-1,027	76,5 113,5	0,305	-0,24	-1,964	59,0 172,0	0,049	-0,43	-1,069	99,0 154,0	0,285	-0,23	-1,668	108,5 216,5	0,095	-0,33

Criterios para evaluar el compromiso estudiantil		Grupo															
		Control PUCESM (N = 19)				Experimental PUCESM (N = 21)				Control UTM (N = 22)				Experimental UTM (N = 25)			
		Z	R-/R+	p-valor	r	Z	R-/R+	p-valor	r	Z	R-/R+	p-valor	r	Z	R-/R+	p-valor	r
Afectivo	CA1	-0,595	84,5 105,5	0,552	-0,14	-1,999	64,0 167,0	0,046	-0,44	-0,883	92,0 161,0	0,377	-0,19	-2,812	62,5 262,5	0,005	-0,56
	CA2	-0,358	94,5 95,5	0,720	-0,08	-2,138	58,5 172,5	0,033	-0,47	-1,469	95,0 158,0	0,142	-0,31	-1,578	99,5 225,5	0,115	-0,32
	CA3	-1,209	65 125	0,227	-0,28	-1,186	95,0 136,0	0,236	-0,26	-1,160	104,5 148,5	0,246	-0,25	-1,325	116,0 209,0	0,185	-0,26
	CA4	-1,374	64 126	0,169	-0,32	-0,966	95,0 136,0	0,334	-0,21	-0,614	105,5 147,5	0,539	-0,13	-1,983	83,0 242,0	0,047	-0,40
Conductual	CCond1	-1,109	124 66	0,268	-0,25	-1,277	83,5 147,5	0,202	-0,28	-0,790	103,0 150,0	0,429	-0,17	-2,231	84,5 240,5	0,026	-0,45
	CCond2	-0,247	98 92	0,805	-0,06	-1,633	72,5 158,5	0,102	-0,36	-0,714	94,0 159,0	0,475	-0,15	-1,978	95,5 229,5	0,048	-0,40
	CCond3	-0,966	76,5 113,5	0,334	-0,22	-1,355	81,5 149,5	0,175	-0,30	-0,885	105,0 148,0	0,376	-0,19	-1,069	125,5 199,5	0,285	-0,21
	CCond4	-0,964	60,5 129,5	0,335	-0,22	-2,183	61,5 169,5	0,029	-0,48	-0,741	104,0 149,0	0,458	-0,16	-1,210	122,0 203,0	0,226	-0,24

Nota: ^aEn cada celda: R- = suma de rangos negativos; R+ = suma de rangos positivos (prueba de Wilcoxon).

DISCUSIÓN

Las diferencias observadas entre las mediciones pretest y posttest en los criterios de la dimensión cognitiva, evaluados en ambos grupos de control, no alcanzaron el nivel de significación estadística. Además, los tamaños del efecto, inferiores a 0,5 en cada caso, sugieren que los cambios en las puntuaciones fueron de baja magnitud. Esto indica que, aunque hubo ciertas variaciones en la distribución de los rangos, estas no fueron lo suficientemente consistentes para considerarse significativas.

En la dimensión afectiva, para ambos grupos de control, los p-valores no evidencian diferencias significativas y los tamaños del efecto son pequeños. Esto sugiere que las percepciones sobre la motivación, el interés y la relevancia del curso se mantuvieron relativamente estables entre las mediciones pretest y posttest.

En la dimensión conductual, las diferencias tampoco fueron significativas y los tamaños del efecto fueron reducidos en ambos grupos. Estos resultados reflejan la estabilidad relativa de las percepciones y comportamientos medidos en ambos grupos de control, lo que era previsible en ausencia de factores que promovieran cambios sustanciales en las dimensiones evaluadas.

Por otra parte, en la dimensión cognitiva, la intervención mostró un impacto positivo en ambos grupos experimentales, aunque con variaciones en la significancia y magnitud del efecto según el ítem. Respecto a la comprensión de los conceptos explicados en el curso (CC1), ambos grupos experimentales presentaron mejoras estadísticamente significativas a pesar de que tanto la mediana (4) como el IQR (1,0) se mantuvieron constantes entre el pretest y el postest. En cuanto a la capacidad de aplicar el conocimiento a nuevas situaciones (CC2), ambos grupos experimentales mostraron mejoras significativas, con un aumento de la mediana de 4 a 5. Sin embargo, mientras que en PUCESM el IQR se mantuvo en 1,0, en UTM el IQR disminuyó de 2,0 a 1,0, lo que indica no solo una mejora en la capacidad percibida, sino también una mayor homogeneidad en las respuestas después de la intervención.

La percepción de profundización del conocimiento a través de las actividades (CC3) mejoró significativamente en ambos grupos, con un aumento de la mediana de 4 a 5 en ambos casos. El tamaño del efecto en la UTM, $r = -0,60$, sugiere que la intervención tuvo un impacto práctico considerable en la percepción de los estudiantes de la UTM sobre este ítem. Finalmente, en relación con la seguridad al responder preguntas (CC4), solo el grupo PUCESM mostró una mejora estadísticamente significativa, con un aumento en la mediana (de 4 a 5). Aunque en la UTM la mediana de CC4 también aumentó, el cambio no fue significativo.

El análisis de la dimensión afectiva en los grupos experimentales revela un panorama mixto. Si bien ambos grupos experimentales mostraron mejoras estadísticamente significativas en el disfrute de las actividades del curso (CA1), el impacto de la intervención fue más fuerte en el GE UTM ($r = -0,56$, $p = 0,005$) que en el GE PUCESM ($r = -0,44$, $p = 0,046$). Esto sugiere que la intervención, tal como se implementó en la UTM, fue particularmente exitosa para generar una experiencia de aprendizaje más placentera para los estudiantes.

En cuanto a la motivación para aprender los contenidos (CA2), el grupo experimental de la PUCESM mostró una mejora significativa ($r = -0,47$, $p = 0,033$). Esto sugiere una mayor efectividad de la intervención en la PUCESM para promover la motivación intrínseca de los estudiantes, a diferencia de la UTM, donde no hubo un cambio significativo.

Mientras que la relevancia percibida (CA3) no cambió significativamente en ninguno de los grupos, el interés en aprender más (CA4) aumentó significativamente en la UTM ($r = -0,40$, $p = 0,047$), con un tamaño de efecto moderado. La ausencia de cambio en la relevancia percibida sugiere que la intervención no logró conectar el contenido del curso con las necesidades o experiencias de los estudiantes. Sin embargo, el aumento en el interés en aprender más en el GE UTM, indica que,

aunque los estudiantes no vieron una conexión inmediata, la intervención pudo despertar curiosidad o mostrar el valor potencial del contenido a largo plazo.

La ausencia de una mejora significativa en la relevancia percibida (CA3) en ambos grupos experimentales podría deberse a factores preexistentes, como las percepciones iniciales de los estudiantes sobre las materias involucradas y el grado de alineación entre el currículo y sus intereses individuales. Estos factores podrían haber influido en la forma en que los estudiantes percibieron la relevancia del contenido del curso, independientemente de la intervención (Sailer et al., 2024). Además, es posible que la ausencia de mejora en CA3 esté influida por el elevado valor que los estudiantes otorgaron al acompañamiento y a la personalización del aprendizaje, tal como se refleja en sus respuestas. Por ejemplo, un estudiante manifestó: “Sería bueno que la profesora dedicara más tiempo presencial o en línea para responder preguntas y resolver dudas”, mientras que otro señaló: “Considero que deberían hacer más sesiones en vivo para poder preguntar directamente lo que no entiendo”.

Teniendo en cuenta lo anterior, la sistematicidad en la presencia sincrónica de los profesores y en el acompañamiento, combinada con la experiencia de los estudiantes para aprender en línea, pudieron incidir en que no se aprovechara por completo el impacto de la sinergia entre los vínculos interpersonales (relaciones entre estudiantes y profesores) e intrapersonales (motivación, autoeficacia) en el componente afectivo del compromiso estudiantil (Martin et al., 2017; Redmond et al., 2018). Lo anterior constituye una invitación a explorar estrategias que involucren, por ejemplo, la integración de actividades H5P y foros para fomentar la discusión y el intercambio de ideas entre todos los participantes.

Por último, en la dimensión conductual, los resultados difieren entre los grupos experimentales. Respecto a la participación activa en clases y actividades (CCond1), en el GE PUCESM no se observó una mejora atribuible a la intervención. En contraste, en el GE UTM se registró una mejora estadísticamente significativa ($r = -0,45$, $p = 0,026$), evidenciada en el aumento de la mediana de 4 a 5, lo que sugiere un mayor involucramiento y participación activa de los estudiantes como resultado de la intervención. En el caso de la dedicación de tiempo al estudio fuera de clase (CCond2), ninguno de los dos grupos experimentales, GE PUCESM y GE UTM, mostró una mejora estadísticamente significativa. Lo anterior pudiera deberse a la dificultad para modificar hábitos de estudio establecidos, la carga académica de otras asignaturas y factores contextuales que afectaron la disponibilidad de la infraestructura tecnológica, limitando la efectividad de las actividades H5P, disminuyendo así el impacto potencial de la intervención en el estudio fuera de clase.

En cuanto a la colaboración con compañeros en las tareas (CCond3), no se observaron cambios estadísticamente significativos tras la intervención en ninguno de los dos grupos. Estos hallazgos sugieren que la intervención no tuvo un impacto consistente en fomentar la colaboración entre compañeros. Esto podría atribuirse a la reducida duración de la intervención, así como con las posibles preferencias

individuales por el trabajo solitario. En cuanto a la búsqueda de oportunidades para aprender más por cuenta propia (CCond4), el GE PUCESM no solo registró un aumento estadísticamente significativo en la mediana (de 4 a 5, $p = 0,029$), sino también una disminución en el IQR (de 2 a 1), lo que sugiere una mayor homogeneidad en las respuestas y un efecto positivo más consistente de la intervención en este grupo. En contraste, en GE UTM no hubo cambios significativos. Esta diferencia podría deberse a que la motivación para el aprendizaje autónomo es intrínseca y puede ser más difícil de influir a través de intervenciones puntuales, especialmente si los estudiantes ya poseen niveles basales de autonomía (Bakker et al., 2015; Safarifard et al., 2024).

Finalmente, cabe recalcar que la evaluación de los criterios de esta dimensión se basó en el autoinforme, y aunque el instrumento utilizado estaba respaldado por un constructo teórico sólido con fuerte consenso entre especialistas y académicos, esto podría haber afectado la estimación del verdadero nivel de colaboración y búsqueda de aprendizaje autónomo.

En general, los hallazgos del estudio corroboran las teorías cognitivas y socioculturales del aprendizaje (Kang y Furtak, 2021; Ortiz y Corrêa, 2020; Safarifard et al., 2024), y coinciden con Bakker et al. (2015) quienes sugieren que entornos de aprendizaje enriquecidos con recursos y oportunidades para el crecimiento pueden fomentar un mejor estado de compromiso lo que predice un buen rendimiento académico. También, estos resultados respaldan la idea constructivista de que las actividades interactivas basadas en H5P tienen el potencial para promover un aprendizaje más profundo y significativo, al facilitar que los estudiantes construyan su propio conocimiento a través de la exploración y la experimentación, lo que fomenta el compromiso cognitivo (Murillo Sevillano et al., 2023).

Según el modelo de Puentedura (2014), la integración de H5P en el diseño de las actividades puede clasificarse como una “modificación”, ya que implica un rediseño de las tareas tradicionales, lo que a su vez fomenta un mayor nivel de interacción y compromiso por parte de los estudiantes. En sintonía con el modelo de Chi y Wylie (2014), los resultados sugieren que las actividades basadas en videos y presentaciones interactivas con H5P, promovieron una transición hacia niveles más profundos de compromiso cognitivo.

Por último, estos resultados abren nuevas líneas de investigación, relacionadas con el estudio del impacto a largo plazo de diseños instruccionales que emplean intensivamente H5P y en la exploración de combinaciones sinérgicas de H5P con un abanico más amplio de otros recursos educativos digitales.

CONCLUSIONES

Este estudio, que estuvo centrado en evaluar el impacto de actividades basadas en H5P en el compromiso estudiantil en cursos híbridos de dos universidades ecuatorianas, encontró resultados mixtos y dependientes de la dimensión específica

del compromiso. Si bien la intervención demostró un efecto positivo en aspectos cognitivos clave como la comprensión de conceptos, la aplicación de conocimientos, la percepción de profundización y en el disfrute de las actividades (dimensión afectiva), no se observaron mejoras significativas en la relevancia percibida del contenido (dimensión afectiva) ni en la colaboración entre compañeros (dimensión conductual). Por lo tanto, si bien H5P puede ser un componente útil en el diseño de experiencias de aprendizaje, su integración debe ser estratégica y complementarse con otros enfoques pedagógicos, especialmente si se busca promover el trabajo colaborativo y la conexión del contenido con los intereses de los estudiantes. Estos hallazgos son particularmente relevantes para el contexto latinoamericano, donde la investigación sobre el impacto de estas herramientas en el compromiso estudiantil está en desarrollo.

En este sentido, se proponen diseños instruccionales que integren estratégicamente H5P con otros enfoques pedagógicos. Una opción sería diseñar escenarios complejos en H5P que requieran la resolución colaborativa de problemas, donde los estudiantes deban trabajar en equipo, discutir y llegar a consensos. Otra posibilidad es implementar una metodología de aprendizaje cooperativo, utilizando módulos H5P para que pequeños grupos se especialicen en diferentes contenidos, buscando así fomentar la interdependencia y el intercambio de conocimientos. Adicionalmente, se podría utilizar el contenido interactivo de H5P como punto de partida para debates estructurados, en foros o en sesiones sincrónicas, con el fin de conectar el material con discusiones significativas y aumentar su relevancia percibida. Finalmente, otra alternativa es involucrar a los estudiantes en la co-creación de actividades H5P, lo que podría promover la colaboración y una comprensión más profunda del contenido.

Las principales limitaciones de este estudio radican en su diseño cuasi-experimental, que dificulta el establecimiento de relaciones causales definitivas; la falta de control sobre variables como el rendimiento previo y la familiaridad con la tecnología; y la duración relativamente corta de la intervención. Para futuras investigaciones, se recomienda adoptar diseños metodológicos más rigurosos, como estudios experimentales con asignación aleatoria o diseños cuasi-experimentales con grupos de control más equivalentes, que permitan establecer con mayor certeza la relación causal entre las intervenciones y los cambios en el compromiso estudiantil. Es crucial ampliar el tamaño muestral para mejorar la potencia estadística y la generalización de los resultados, así como utilizar un enfoque multidimensional en la medición del compromiso, combinando cuestionarios con observaciones, análisis de datos de aprendizaje y entrevistas cualitativas para una comprensión más completa. Se deben extender los estudios a lo largo del tiempo, con diseños longitudinales, para evaluar la sostenibilidad de los efectos. Además, es fundamental investigar el impacto de diferentes diseños instruccionales que integren la tecnología, manteniendo H5P como componente central, a la vez que se comparan diversos enfoques pedagógicos para identificar las estrategias más efectivas en el fomento de

las distintas dimensiones del compromiso estudiantil. Finalmente, se debe ampliar el contexto geográfico y cultural de estas investigaciones.

REFERENCIAS

- Abdool, P. S., Nirula, L., Bonato, S., Rajji, T. K. y Silver, I. L. (2017). Simulation in undergraduate psychiatry: Exploring the depth of learner engagement. *Academic Psychiatry*, 41(2), 251-261. <https://doi.org/10.1007/s40596-016-0633-9>
- Ali, I., Narayan, A. K. y Sharma, U. (2020). Adapting to COVID-19 disruptions: Student engagement in online learning of accounting. *Accounting Research Journal*, 34(3), 261-269. <https://doi.org/10.1108/ARJ-09-2020-0293>
- Amhag, L. (2020). Student reflections and self-assessments in vocational training supported by a mobile learning hub. *International Journal of Mobile and Blended Learning*, 12(1), 1-16. <https://doi.org/10.4018/IJMBL.2020010101>
- Bailey, D., Almusharraf, N. y Hatcher, R. (2021). Finding satisfaction: Intrinsic motivation for synchronous and asynchronous communication in the online language learning context. *Education and Information Technologies*, 26(3), 2563-2583. <https://doi.org/10.1007/s10639-020-10369-z>
- Bajaber, S. S. A. (2024). Factors influencing students willingness to continue online learning as a lifelong learning: A path analysis based on MOA theoretical framework. *International Journal of Educational Research Open*, 7, 100377. <https://doi.org/10.1016/j.ijedro.2024.100377>
- Bakker, A. B., Sanz Vergel, A. I. y Kuntze, J. (2015). Student engagement and performance: A weekly diary study on the role of openness. *Motivation and Emotion*, 39(1), 49-62. <https://doi.org/10.1007/s11031-014-9422-5>
- Baleni, Z. (2015). Online formative assessment in higher education: Its pros and cons. *Electronic Journal of e-Learning*, 13(4), 281-291.
- Bond, M. y Bedenlier, S. (2019). Facilitating student engagement through educational technology: Towards a conceptual framework. *Journal of Interactive Media in Education*, 2019(1), Artículo 11. <https://doi.org/10.5334/jime.528>
- Bond, M., Buntins, K., Bedenlier, S., Zawacki-Richter, O. y Kerres, M. (2020). Mapping research in student engagement and educational technology in higher education: A systematic evidence map. *International Journal of Educational Technology in Higher Education*, 17, Artículo 2. <https://doi.org/10.1186/s41239-019-0176-8>
- Brookfield, S. D. (2009). Self-directed learning. En R. Maclean y D. Wilson (Eds.), *International handbook of education for the changing world of work: Bridging academic and vocational learning* (pp. 2615-2627). Springer Netherlands. https://doi.org/10.1007/978-1-4020-5281-1_172
- Chi, M. T. H. y Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, 49(4), 219-243. <https://doi.org/10.1080/00461520.2014.965823>
- Cleary, T. J. y Zimmerman, B. J. (2012). A cyclical self-regulatory account of student engagement: Theoretical foundations and applications. En S. L. Christenson, A. L. Reschly y C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 237-257). Springer. https://doi.org/10.1007/978-1-4614-2018-7_11

- Downes, S. (2022). Connectivism. *Asian Journal of Distance Education*, 17(1), 114-138.
- Finn, J. D. y Zimmer, K. S. (2012). Student engagement: What is it? Why does it matter? En S. L. Christenson, A. L. Reschly y C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 97-131). Springer. https://doi.org/10.1007/978-1-4614-2018-7_5
- Fredricks, J. A., Blumenfeld, P. C. y Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, 74(1), 59-109. <https://doi.org/10.3102/00346543074001059>
- Fredricks, J. A., Filsecker, M. y Lawson, M. A. (2016). Student engagement, context, and adjustment: Addressing definitional, measurement, and methodological issues. *Learning and Instruction*, 43, 1-4. <https://doi.org/10.1016/j.learninstruc.2016.02.002>
- Gellisch, M., Morosan-Puopolo, G., Wolf, O. T., Moser, D. A., Zaehres, H. y Brand-Saberi, B. (2023). Interactive teaching enhances students' physiological arousal during online learning. *Annals of Anatomy*, 247, 152050. <https://doi.org/10.1016/j.aanat.2023.152050>
- Jacob, T. y Centofanti, S. (2024). Effectiveness of H5P in improving student learning outcomes in an online tertiary education setting. *Journal of Computing in Higher Education*, 36, 469-485. <https://doi.org/10.1007/s12528-023-09361-6>
- Kang, H. y Furtak, E. M. (2021). Learning theory, classroom assessment, and equity. *Educational Measurement: Issues and Practice*, 40(3), 73-82. <https://doi.org/10.1111/emip.12423>
- Kuh, G. D. (2009). What student affairs professionals need to know about student engagement. *Journal of College Student Development*, 50(6), 683-706. <https://doi.org/10.1353/csd.0.0099>
- Lamtara, S. (2023). Faculty's pedagogical knowledge for technology-enhanced learning in higher education. *International Journal of Higher Education Pedagogies*, 4(2), 1-24. <https://doi.org/10.33422/ijhep.v4i2.412>
- Martin, F. y Borup, J. (2022). Online learner engagement: Conceptual definitions, research themes, and supportive practices. *Educational Psychologist*, 57(3), 162-177. <https://doi.org/10.1080/00461520.2022.2089147>
- Martin, K., Goldwasser, M. y Galentino, R. (2017). Impact of cohort bonds on student satisfaction and engagement. *Current Issues in Education*, 19(3). <https://cie.asu.edu/ojs/index.php/cieatasu/article/view/1550>
- Mazman Akar, S. G. (2024). Why do students disengage from online courses? *The Internet and Higher Education*, 62, 100948. <https://doi.org/10.1016/j.iheduc.2024.100948>
- Méndez Mantuano, M. O., Egíez Caviades, E. C., Ochoa Ladines, K. V., Plúas Rogel, D. R. y Paredes Yuqui, C. E. (2021). Análisis del conductismo, cognitivismo, constructivismo y su interrelación con el conectivismo en la educación postpandemia. *South Florida Journal of Development*, 2(5), 6850-6863. <https://doi.org/10.46932/sfjdv2n5-038>
- Morris, R., Perry, T. y Wardle, L. (2021). Formative assessment and feedback for learning in higher education: A systematic review. *Review of Education*, 9(3), Artículo e3292. <https://doi.org/10.1002/rev3.3292>
- Mulrooney, H. M. y Kelly, A. F. (2020). Covid 19 and the move to online teaching: Impact on perceptions of belonging in staff and students in a UK widening participation university. *Journal of Applied Learning and Teaching*, 3(2), 17-30. <https://doi.org/10.37074/jalt.2020.3.2.15>
- Murillo Sevillano, L. N. de J., Vintimilla Burgos, N. P. y Murillo Sevillano, I. M.

- (2023). La educación virtual e híbrida. Consideraciones desde la Universidad de Guayaquil. *Conrado*, 19(90), 429-438. http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1990-86442023000100429&lng=es&nrm=iso&tlng=en
- Ortiz, J. A. T. y Corrêa, T. H. B. (2020). Pedagogical aspects of connectivism and its relationship with social networks and learning ecologies. *Revista Brasileira de Educação*, 25, Artículo e250026. <https://doi.org/10.1590/s1413-24782020250026>
- Parson, L. y Major, C. (2020). Learning theory through a social justice lens. En C. H. Major y L. S. Parson (Eds.), *Teaching and learning for social justice and equity in higher education* (pp. 7-38). Springer International Publishing. https://doi.org/10.1007/978-3-030-44939-1_2
- Pekrun, R. y Linnenbrink-Garcia, L. (2012). Academic emotions and student engagement. En S. L. Christenson, A. L. Reschly y C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 259-282). Springer. https://doi.org/10.1007/978-1-4614-2018-7_12
- Ploetzner, R. (2024). The effectiveness of enhanced interaction features in educational videos: A meta-analysis. *Interactive Learning Environments*, 32(5), 1597-1612. <https://doi.org/10.1080/10494820.2022.2123002>
- Puentedura, R. R. (2014). *SAMR: Getting to transformation*. Hippasus. <http://www.hippasus.com/rrpweblog/archives/2013/04/16/SAMRGettingToTransformation.pdf>
- Rahmi, U., Fajri, B. y Azrul, A. (2024). Effectiveness of interactive content with H5P for Moodle-Learning Management System in blended learning. *Journal of Learning for Development*, 11(1), 66-81. <https://doi.org/10.56059/jl4d.v11i1.1135>
- Redmond, P., Abawi, L.-A., Brown, A., Henderson, R. y Heffernan, A. (2018). An online engagement framework for higher education. *Online Learning*, 22(1), 183-204. <https://doi.org/10.24059/olj.v22i1.1175>
- Reeve, J. (2012). A self-determination theory perspective on student engagement. En S. L. Christenson, A. L. Reschly y C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 149-172). Springer. https://doi.org/10.1007/978-1-4614-2018-7_7
- Rossetti-López, S. R., Coronado-García, M. A. y Rojas-Rodríguez, I. S. (2023). Percepción de los estudiantes sobre el uso de actividades interactivas con H5P. *Revista Iberoamericana de Educación Superior*, 14(40), 59-76. <https://doi.org/10.22201/iisue.20072872e.2023.40.1545>
- Safarifard, R., Lavasani, M. G., Hejazi, E. y Thani, F. N. (2024). Pedagogical aspect of e-learning in higher education: A systematic literature review. *Knowledge Management & E-Learning*, 16(3), 521-546. <https://doi.org/10.34105/j.kmel.2024.16.024>
- Sailer, M., Maier, R., Berger, S., Kastorff, T. y Stegmann, K. (2024). Learning activities in technology-enhanced learning: A systematic review of meta-analyses and second-order meta-analysis in higher education. *Learning and Individual Differences*, 112, 102446. <https://doi.org/10.1016/j.lindif.2024.102446>
- Sharmin, N., Pandya, J., Stevenson, T. R. y Chow, A. K. (2024). Interactive H5P content for increased student engagement in a dental hygiene program. *Canadian Journal of Dental Hygiene*, 58(2), 88. <https://pmc.ncbi.nlm.nih.gov/articles/PMC11223635/>
- Sinnayah, P., Salcedo, A. y Rekhari, S. (2021). Reimagining physiology education with interactive content developed in H5P. *Advances in Physiology Education*, 45(1), 71-76. <https://doi.org/10.1152/advan.00021.2020>
- Skinner, E. A. y Pitzer, J. R. (2012). Developmental dynamics of student

- engagement, coping, and everyday resilience. En S. L. Christenson, A. L. Reschly y C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 21-44). Springer. https://doi.org/10.1007/978-1-4614-2018-7_2
- Taylor, F., Welton, R. y Ellicott, D. (2018). The role of technology in enhancing engagement and building communities of practice within tourism education. En L. Gómez Chova, A. López Martínez y I. Candel Torres (Eds.), *ICERI2018 proceedings: 11th International Conference of Education, Research and Innovation, Seville, Spain* (pp. 934-938). IATED Academy. <https://doi.org/10.21125/iceri.2018.0122>
- Theelen, H. y van Breukelen, D. H. J. (2022). The didactic and pedagogical design of e-learning in higher education: A systematic literature review. *Journal of Computer Assisted Learning*, 38(5), 1286-1303. <https://doi.org/10.1111/jcal.12705>

Fecha de recepción del artículo: 1 de diciembre de 2024

Fecha de aceptación del artículo: 12 de marzo de 2025

Fecha de aprobación para maquetación: 31 de marzo de 2025

Fecha de publicación en OnlineFirst: 24 de abril de 2025

Fecha de publicación: 1 de julio de 2025

Aula invertida gamificada: aplicación de una estrategia didáctica para trabajar el pensamiento computacional en futuros docentes

Gamified flipped classroom: application of a teaching strategy to develop computational thinking in future teachers



 Jesús Carpena Arias - *Universitat Jaume I, UJI (España)*
 Francesc M. Esteve Mon - *Universitat Jaume I, UJI (España)*

RESUMEN

Este estudio analiza el impacto de la gamificación y el aprendizaje invertido en el desarrollo del pensamiento computacional en estudiantes del Grado en Educación Infantil. Para ello, se evaluó la estrategia del aula invertida gamificada. La muestra incluyó un total de 86 estudiantes. El grupo experimental compuesto por 36 estudiantes han recibido una intervención didáctica basada en el aula invertida gamificada y completaron tanto el pretest como el postest. El grupo control, compuesto por 50 estudiantes, siguieron una metodología tradicional durante seis semanas, y completaron también el pretest como el postest, permitiendo una comparación de los efectos de ambas metodologías. Para evaluar la mejora del pensamiento computacional en los estudiantes, se utilizó el Test de Pensamiento Computacional (TPC). Este instrumento se administró tanto antes como después de la implementación de la estrategia didáctica, actuando como pretest y postest, respectivamente. Los análisis estadísticos de los datos recopilados indicaron un incremento en las puntuaciones de pensamiento computacional en el grupo experimental. Este hallazgo sugiere que la combinación de la metodología de aula invertida con elementos de gamificación puede ser eficaz para mejorar el pensamiento computacional en estudiantes universitarios del grado de maestro/a de educación infantil. Por lo tanto, el estudio aporta evidencia de que la implementación de estrategias pedagógicas innovadoras, como el aula invertida gamificada, puede tener un impacto significativo en el desarrollo del pensamiento computacional en estudiantes de educación superior. Sin embargo, es necesaria más investigación para confirmar y expandir estos resultados en diferentes contextos y con muestras más grandes.

Palabras clave: aula invertida gamificada; pensamiento computacional; educación superior.

ABSTRACT

This study analyses the impact of gamification and flipped learning on the development of computational thinking in students enrolled in the Bachelor's Degree in Early Childhood Education. The gamified flipped classroom strategy was evaluated for this purpose. The sample included a total of 86 students. The experimental group, consisting of 36 students, received a didactic intervention based on the gamified flipped classroom approach and completed both the pre-test and post-test. The control group, consisting of 50 students, followed a traditional methodology for six weeks and also completed both the pre-test and post-test, allowing for a comparison of the effects of both methodologies. To assess students' development in computational thinking, the Computational Thinking Test (CTT) was used. This instrument was administered both before and after the implementation of the didactic strategy, serving as the pre-test and post-test, respectively. Statistical analyses of the collected data indicated an increase in computational thinking scores in the experimental group. This finding suggests that combining the flipped classroom methodology with gamification elements can be effective in enhancing computational thinking in university students pursuing a degree in Early Childhood Education. Therefore, the study provides evidence that implementing innovative pedagogical strategies, such as the gamified flipped classroom, can have a significant impact on the development of computational thinking in higher education students. However, further research is needed to confirm and expand these findings in different contexts and with larger samples.

Keywords: gamified flipped classroom; computational thinking; higher education.

INTRODUCCIÓN

La educación superior está experimentando una transformación significativa con la incorporación de metodologías innovadoras como la gamificación y el aula invertida. Estas estrategias pedagógicas están redefiniendo los paradigmas de enseñanza y aprendizaje, ofreciendo nuevas perspectivas y desafíos (Soriano-Sánchez y Jiménez-Vázquez, 2023). El uso del flipped learning (FL) y la gamificación están cobrando mayor protagonismo en los centros educativos (Ekici, 2021).

A pesar de ello, existe poca investigación que combine estos dos enfoques (FL y gamificación) y analice las posibles mejoras en la motivación y el rendimiento académico en la etapa universitaria (Carpena Arias y Esteve Mon, 2022a). La finalidad de este trabajo es aplicar una intervención didáctica en estudiantes universitarios utilizando la metodología gamificación y el FL para trabajar el pensamiento computacional (PC).

Gamificación y aula invertida en educación superior

La educación superior está introduciendo estrategias didácticas como la gamificación y el aula invertida, las cuales buscan mejorar la motivación y el rendimiento académico de los estudiantes (Candel et al., 2024). Estas metodologías se han combinado eficazmente en lo que se conoce como aula invertida gamificada, a partir de ahora AIG, una estrategia pedagógica que incorpora elementos lúdicos dentro de un marco de aprendizaje activo y centrado en el estudiante (Carpena Arias y Esteve Mon, 2024).

El concepto de “aula invertida” fue introducido inicialmente bajo el término *inverted classroom* por Lage et al. (2000). Sin embargo, no fue hasta 2009 cuando Bergmann y Sams (2009) popularizaron definitivamente el término. Ellos comenzaron a grabar sus lecciones para que los estudiantes que no podían asistir físicamente a clase pudieran acceder a los contenidos en sus hogares. Esto permitía a todos los estudiantes mantenerse al día con el material enseñado y seguir el ritmo académico de la clase. Esta metodología transforma el enfoque tradicional educativo al requerir que los estudiantes primero trabajen previamente el material didáctico, utilizando para ello recursos digitales, lecturas o vídeos antes de asistir a clase (Bergmann y Sams, 2014). Esta preparación individual permite que, durante el tiempo en el aula, se dediquen a actividades más dinámicas y aplicadas como debates, trabajos en grupo y resolución de problemas, todo ello supervisado y orientado por el docente (Santiago y Bergmann, 2018).

Además, un estudio llevado a cabo por Bishop y Verleger (2013), demostró que el aula invertida mejora significativamente la calidad de las interacciones de los estudiantes. Al trasladar la instrucción directa fuera del aula, los estudiantes tienen más oportunidades de interactuar con sus profesores durante el tiempo de clase, lo que facilita una comprensión más profunda y un aprendizaje más significativo. Por otro

lado, según Touron y Santiago (2014), el modelo FL fomenta una mayor interacción entre estudiantes y profesores, permitiendo una atención más personalizada y un aprendizaje más profundo.

La implementación del aprendizaje invertido, respaldado por dispositivos digitales, en estudiantes del grado de educación primaria, favorece el fomento de la creatividad, el pensamiento crítico, la comunicación, la colaboración y las habilidades sociales, según Martín y Tourón (2017). Estos beneficios se ven también en la formación en línea para estudiantes del grado de maestro/a de educación infantil, donde el aula invertida muestra mejoras en el rendimiento académico en el área de matemáticas, así como un aumento en la motivación (Sacristán et al., 2017). Además, es importante considerar otras estrategias pedagógicas complementarias, como la gamificación para ampliar estos beneficios. La gamificación en educación, entendida como el uso de elementos y principios del diseño de juegos en contextos no lúdicos, tiene como objetivo motivar y aumentar la participación de los estudiantes, mejorando así su proceso de aprendizaje (Prieto, 2020). En este sentido, la gamificación se aplica en la educación para hacerla más atractiva e inmersiva, utilizando mecánicas de juego como puntos, insignias y tableros de clasificación, que fomentan la competencia saludable y guían a los estudiantes hacia logros específicos (Werbach y Hunter, 2015). En 2014, Marczewski propuso una distinción en los enfoques de gamificación, diferenciando entre gamificación superficial y profunda, basada en el tipo de motivación que se busca estimular. La gamificación superficial, a menudo centrada en el contenido, utiliza elementos externos como puntos, insignias y tableros de clasificación para fomentar la motivación extrínseca. En cambio, la gamificación profunda se enfoca en una integración holística de las dinámicas, mecánicas y componentes esenciales del juego para cultivar una motivación intrínseca más sostenible y profunda (Marczewski, 2014).

El uso de la gamificación y el aula invertida ha emergido como una estrategia metodológica en la educación superior, orientada hacia el fomento de un aprendizaje más interactivo y motivador. Esta combinación pedagógica ha sido objeto de numerosas investigaciones, que apuntan a sus beneficios tanto en la motivación de los estudiantes como en el rendimiento académico Carpena Arias et al. (2022). El estudio de Zainuddin (2018) investigó cómo la gamificación en el contexto del aula invertida puede mejorar tanto el rendimiento de aprendizaje como la motivación percibida por los estudiantes. Por otro lado, Zamora-Polo et al. (2019) examinaron cómo los estudiantes no científicos pueden beneficiarse de una pedagogía activa que fusiona la gamificación con el aula invertida, revelando mejoras significativas en el compromiso y la comprensión de conceptos científicos generales. Similarmente, Gómez-Carrasco et al. (2019) documentaron los efectos positivos de un programa de gamificación y aula invertida en profesores en formación, resaltando un aumento en la motivación y la percepción del aprendizaje.

Investigaciones adicionales, como las de Sailer y Sailer (2021) destacan cómo la incorporación de gamificación en actividades de aula invertida puede potenciar la

interacción y la colaboración entre estudiantes, lo que a su vez mejora los resultados de aprendizaje. Además, el trabajo de Elzeke et al. (2022) sobre la efectividad del aula invertida gamificada en estudiantes de enfermería subraya la mejora de las competencias clínicas y la motivación de aprendizaje, ofreciendo evidencia de la adaptabilidad de estas metodologías a distintas disciplinas. Además, una revisión sistemática llevada a cabo por Ekici (2021) sobre el uso de la gamificación en el aprendizaje invertido indica que la combinación de estas metodologías está ganando terreno en diversos contextos educativos, demostrando resultados prometedores en términos de motivación y rendimiento académico.

La implementación de metodologías como la gamificación y el aula invertida en la educación superior ha demostrado un impacto positivo en las interacciones entre compañeros, fomentando un ambiente de colaboración y trabajo en equipo. (Carpena Arias y Esteve Mon, 2022a). Por tanto, podemos definir la estrategia metodológica de AIG como la estrategia pedagógica que combina los elementos propios de la gamificación y el FL para fomentar una mayor participación y motivación en los estudiantes.

Pensamiento computacional

El término pensamiento computacional fue introducido por primera vez por Wing (2006), definiéndolo como la capacidad para resolver problemas, diseñar sistemas y comprender el comportamiento humano a través de los principios fundamentales de la ciencia de la computación. Wing argumentó que el pensamiento computacional es tan importante como aprender matemáticas y escribir para la formación en el siglo XXI, posicionándolo como una habilidad esencial que todos deberían poseer. En los siguientes años, esta primera definición ha sido objeto de debate, tratando de llegar a un consenso sobre qué es el pensamiento computacional y qué elementos lo incorporan (Serrano, 2022). Después de varios años sin lograr un acuerdo sobre la definición del pensamiento computacional, finalmente se ha alcanzado un consenso, definiéndolo como: “la capacidad (humana) de resolver problemas y expresar ideas haciendo uso de conceptos, prácticas y perspectivas propias de las Ciencias de la Computación” (Román-González, 2022).

La creciente importancia del pensamiento computacional ha llevado a que se incluya en la educación formal de diversas maneras. Por ejemplo, en España, el Real Decreto 95/2022, del 1 de febrero, que establece las enseñanzas mínimas de la Educación Infantil (EI), ha planteado por primera vez la incorporación del Pensamiento Computacional (PC) como parte del currículo de esta etapa educativa. Este enfoque también se refleja en el informe INTEF (2021), que concluye que el uso de actividades desenchufadas en educación infantil y primaria es una estrategia muy efectiva para trabajar el PC al mismo tiempo que se abordan contenidos de diferentes áreas del currículo.

Según Serrano y Ortuño (2021), los docentes no poseen la formación adecuada para integrar el PC de manera efectiva y apropiada. Además, la inclusión del PC en los currículos ya establecidos puede resultar compleja, especialmente en áreas que tradicionalmente no han estado vinculadas con la informática. En relación con la formación docente, los profesores tienden a especializarse en programación y robótica, pero no en la didáctica del PC desconectado, lo cual es esencial para fomentar el PC, como señalan Ortuño y Serrano (2024).

En su revisión sistemática, Collado-Sánchez et al. (2023) propone adaptar los programas de formación docente para incluir habilidades de PC y así garantizar una educación alineada con las demandas tecnológicas contemporáneas. En su estudio, Villalustre-Martínez (2024) analiza el nivel de pensamiento computacional en futuros maestros, destacando la influencia del género y la experiencia previa en programación robótica. El estudio identifica diferencias significativas entre hombres y mujeres, así como entre estudiantes con y sin experiencia previa, lo que subraya la necesidad de adaptar la formación docente a estos perfiles.

METODOLOGÍA

Esta investigación es parte de un proyecto más amplio que utiliza la metodología de investigación en diseño educativo (EDR, por sus siglas en inglés), una estrategia sistemática enfocada en la mejora de las prácticas educativas a través del desarrollo iterativo y la evaluación de intervenciones pedagógicas, destinada a solucionar problemas complejos en entornos reales de enseñanza y aprendizaje (Plomp y Nieveen, 2009). Además de enfocarse en la invención de soluciones educativas avanzadas, la metodología EDR busca generar conocimiento teórico significativo que sea transferible a distintos entornos educativos.

En etapas precedentes a este estudio, hemos trabajado en dos fases. En primer lugar, realizamos una revisión de la literatura con el objetivo de construir una base teórica en torno a la combinación de las metodologías de gamificación y el modelo de aula invertida en el contexto de la educación superior (Carpena Arias y Esteve Mon, 2022a). Luego, durante la segunda etapa, procedimos a crear y validar un enfoque didáctico que aplicaba el concepto de aula invertida, enriquecido con elementos de gamificación. Este proceso incluyó consultas con expertos en tecnología educativa para asegurar que nuestra propuesta fuese relevante y coherente (Carpena Arias y Esteve Mon, 2024). La etapa final se centró en evaluar cómo los estudiantes percibían la efectividad y aplicabilidad de nuestra intervención, mediante entrevistas personales. Esta evaluación nos permitió tener una aproximación de la percepción de los estudiantes en el proceso de aprendizaje. En este estudio trataremos de comprobar si la intervención AIG tiene o no un impacto en el rendimiento de las habilidades de PC en estudiantes del grado de maestro/a de Educación Infantil.

El objetivo principal de este estudio es determinar la eficacia de una intervención didáctica que integra elementos de gamificación y aula invertida, examinando su

influencia en las habilidades del PC de estudiantes universitarios. Esto se logrará mediante la comparación de los resultados de pretest y postest entre un grupo experimental y un grupo control. En consonancia a este objetivo principal, la pregunta de investigación es la siguiente: ¿Es más efectiva la metodología de aula invertida gamificada que la metodología tradicional en el desarrollo del pensamiento computacional en estudiantes del grado de maestro/a en Educación Infantil, según los resultados del test TPC?

Contexto y participantes

Los participantes en este estudio son estudiantes matriculados en la asignatura Tecnologías de la Información y la Comunicación en Educación de cuarto curso del Grado de Maestro/a en Educación Infantil de una universidad pública de España. Los participantes en el estudio fueron voluntarios y toda la información recopilada a través de los test se mantuvo de forma anónima. Antes de su participación, cada individuo firmó un consentimiento informado, en el que se explicaban los procedimientos del estudio, su derecho a la privacidad y la posibilidad de retirarse en cualquier momento sin consecuencias. Esta investigación sigue las pautas éticas establecidas por la British Educational Research Association (BERA, 2018), garantizando el respeto a los principios de confidencialidad, consentimiento informado y bienestar de los participantes. La muestra ha estado compuesta por 86 estudiantes en el estudio, con edades comprendidas entre 20 y 40 años distribuidos en dos grupos. Grupo experimental: 36 estudiantes (9 hombres y 27 mujeres), recibieron la intervención basada en gamificación y aula invertida para trabajar el pensamiento computacional y completaron tanto el pretest como el postest. Grupo control: 50 estudiantes (49 mujeres y 1 hombre), participaron en un enfoque convencional de enseñanza y completaron también el pretest y el postest. Para evitar sesgos en la asignación de los grupos, se intentó equilibrar la distribución de los participantes. Además, el tamaño de la muestra se determinó con un nivel de confianza del 95 % y un margen de error del 6 %, lo que ha resultado en un cálculo de 86 participantes, coincidiendo exactamente con los estudiantes que completaron el estudio.

La intervención se ha llevado a cabo durante un periodo de seis semanas. En este tiempo, los estudiantes se sumergieron en una estrategia metodológica que integraba el modelo del aula invertida con elementos de gamificación. Este enfoque promovió que los estudiantes asumieran un rol activo en su proceso de aprendizaje, adquiriendo conocimientos teóricos de manera autónoma fuera del aula, mediante la consulta de recursos digitales y la realización de tareas preparatorias. Posteriormente, aplicaban dichos conocimientos en actividades prácticas durante las sesiones presenciales en la universidad. La gamificación se incorporó como un elemento dinamizador, buscando potenciar la motivación y el compromiso de los estudiantes hacia las actividades y el estudio mediante la utilización de mecánicas y dinámicas lúdicas en un contexto educativo.

Intervención y metodología gamificada

MyClassGame es una innovadora plataforma de gamificación educativa de código abierto que permite a los docentes de todas las etapas educativas crear y diseñar entornos gamificados personalizados. Esta herramienta facilita la integración de metodologías activas, haciendo del proceso de enseñanza una experiencia más motivadora. Los docentes pueden configurar elementos como puntos de experiencia, niveles y misiones que evalúan a los estudiantes mediante rúbricas, todo dentro de un marco lúdico que busca mejorar la motivación y el aprendizaje. En este estudio, hemos adoptado los principios de gamificación propuestos por Werbach y Hunter (2015), centrados en las dinámicas, mecánicas y componentes, comúnmente referidos mediante el acrónimo MDA; así como las fases que propone Ocón-Galilea (2017) para diseñar nuestras actividades gamificadas. A continuación, detallamos cómo se han aplicado estos elementos en nuestra intervención.

Mecánicas de Juego

Las mecánicas constituyen los elementos fundamentales del juego, definiendo sus reglas y funcionamiento. En nuestra intervención, hemos incorporado las siguientes mecánicas:

- Colaboración: fomentamos el trabajo en equipo en la mayoría de las actividades de nuestro proyecto.
- Desafíos: implementamos actividades progresivas que requerían ser superadas por los grupos para avanzar de nivel.
- Recompensas: utilizamos puntos de experiencia (XP) y poderes como incentivos para los estudiantes al superar retos.

Dinámicas de Juego

Las dinámicas representan los aspectos más abstractos del entorno gamificado. En nuestro trabajo, hemos empleado:

- Narrativa: la historia “El apocalipsis zombie” sirvió como eje central de nuestro Entorno de Aprendizaje Gamificado (GLE). Se creó un video introductorio que establecía el contexto: una pandemia que ha llevado a los estudiantes a convertirse en “alumnos zombies” con deficiencias en competencias digitales. La misión era superar desafíos relacionados con el pensamiento computacional para evitar el analfabetismo digital.
- Progresión: establecimos niveles de progresión (1 aficionado, 2 debutante, 3 amateur, 4 estrella, 5 leyenda) para marcar el avance de los estudiantes.

Componentes de Juego

Los componentes son los elementos concretos y específicos del juego. En nuestra intervención, incluimos:

- Avatar: los estudiantes crearon sus propios avatares.
- Combate: una batalla final contra los zombies.
- Misiones: actividades específicas que debían completarse.
- Colecciones: los estudiantes coleccionaban poderes en su tablero de recompensas.

Figura 1
Gestión de la clase gamificada

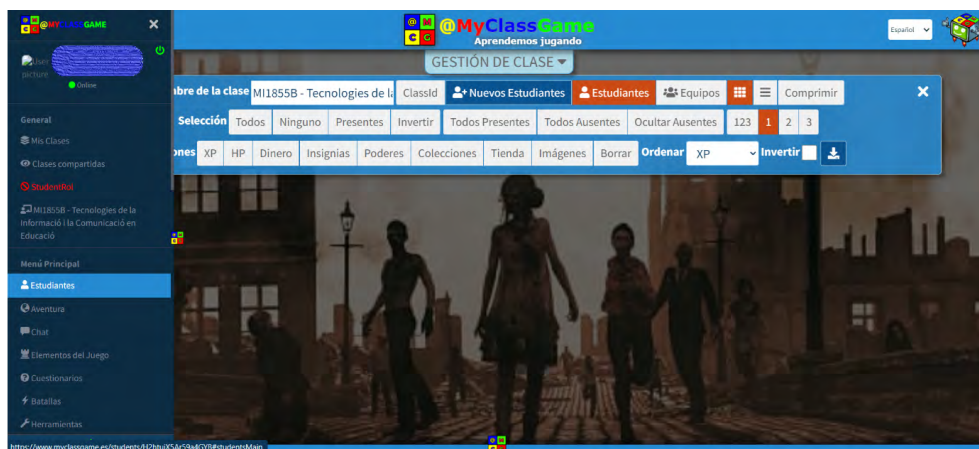


Figura 2

Los componentes del juego



Comportamientos y Penalizaciones

Los comportamientos deseables y las penalizaciones fueron clave en nuestra metodología:

- Comportamientos premiados: colaboración, participación activa, asistencia, puntualidad y persistencia en el trabajo, recompensados con puntos de experiencia (XP).
- Penalizaciones: se retiraban 50 puntos de vida (Hp) por llegar tarde, interrumpir en clase o no cumplir con las tareas.

Eventos Narrativos

La narrativa se dividió en eventos clave para mantener el interés y la motivación:

1. El Reencuentro: lucha contra los zombies (1ª misión).
2. La Huida: escapar de la prisión zombie (2ª misión).
3. La Gran Batalla: batalla a muerte (3ª misión).
4. El Desenlace: vuelta a la universidad 5.0 (4ª misión).

Cada evento requería superar retos específicos relacionados con el pensamiento computacional para avanzar. La intervención educativa AIG se desarrolla en cinco

fases estructuradas a lo largo de seis semanas (Carpena Arias y Esteve Mon, 2022b). Comenzamos con una fase de investigación y puesta en práctica que dura una semana, donde los estudiantes exploran fuera del aula conceptos de pensamiento computacional y preparan exposiciones de lo investigado. La siguiente fase tiene una duración de dos semanas. Dentro del aula, se organizan debates sobre los beneficios y desafíos del uso de PCs y robótica en educación. Posteriormente, se trabajan habilidades propias del PC mediante plataformas como code.org y Scratch. Los estudiantes que más participan son premiados con insignias que pueden intercambiar por recompensas. Posteriormente, los estudiantes reflexionan escribiendo en sus blogs personales sobre los aprendizajes obtenidos. La tercera fase, que también dura dos semanas, se centra en la elaboración de productos. Los estudiantes trabajan en equipos para crear materiales educativos vinculados al pensamiento computacional y diseñan unidades didácticas adaptadas a la educación infantil, donde los equipos más participativos reciben poderes extras. La cuarta fase es la presentación del producto final. Aquí, dentro del aula, los equipos presentan y demuestran los materiales creados, y aquellos con más privilegios escogen primero su turno de presentación. Finalmente, la quinta fase, que dura una semana, consiste en la valoración del proyecto. Los estudiantes realizan autoevaluaciones y coevaluaciones sobre la participación en el proyecto y el desempeño de sus compañeros utilizando una rúbrica diseñada por el docente. Esta estructura fomenta un aprendizaje integral, colaborativo y reflexivo sobre el pensamiento computacional y su aplicación práctica.

Instrumento y análisis de datos

Para evaluar las capacidades de pensamiento computacional de los estudiantes, se empleó el Test de Pensamiento Computacional (TPC) desarrollado por Román (2015). El instrumento fue aplicado antes de la intervención (Pretest) y después de la misma (Postest), con el fin de evaluar el desarrollo en las competencias de pensamiento computacional adquiridas por los estudiantes. Este test (TPC) evalúa elementos fundamentales como secuenciación, iteraciones, condicionales, funciones y el manejo de variables. El propósito principal del TPC es estimar la habilidad en pensamiento computacional de los participantes, enfocándose en cómo los estudiantes aplican conceptos computacionales para resolver problemas. Este test está compuesto por 28 preguntas objetivas de elección múltiple ofreciendo cuatro opciones de respuesta y debe completarse en un tiempo máximo de 45 minutos. Este test tiene un coeficiente alfa de Cronbach de 0,74, indicando una fiabilidad aceptable en términos de consistencia interna. Las preguntas fueron administradas a través de un formulario de Google Forms al inicio y al término de la intervención en ambos grupos. Los datos fueron recogidos mediante Google Forms y se analizaron usando el software estadístico Jamovi. En el estudio del TPC de Román-González (2015), se compararon los grupos, Grupo Control y Grupo Experimental, antes y después de la intervención (pretest y postest), aplicando análisis descriptivo. Esto

permitió calcular tanto la ganancia de conocimiento (gain score) como la mejora de aprendizaje ajustada (normalized learning gain).

RESULTADOS

En primer lugar, se presentan los resultados descriptivos en el nivel inicial de PC de todos los estudiantes diferenciado por género. Como se observa en la Tabla 1, la media para todos los participantes es de 20.1, con una media de 20.2 para las mujeres y ligeramente inferior para los hombres, con 19.1. La mediana se mantiene en 20.0 en general, y es idéntica para las mujeres, mientras que para los hombres es ligeramente menor, con 19.0. Por otro lado, la desviación estándar (SD), que indica la variabilidad de los datos, es de 4.56 para todos los participantes, 4.45 para las mujeres, y es más alta para los hombres, con 5.47.

Tabla 1
Estadísticos descriptivos del pretest del test PC

	Sexo	Nivel inicial
N	Todos	86
	Mujeres	76
	Hombres	10
Media	Todos	20.1
	Mujeres	20.2
	Hombres	19.1
Mediana	Todos	20.0
	Mujeres	20.0
	Hombres	19.0
SD	Todos	4.56
	Mujeres	4.45
	Hombres	5.47
Mínimo	Todos	9
	Mujeres	9
	Hombres	10
Máximo	Todos	30
	Mujeres	30
	Hombres	26

A continuación, la Tabla 2 muestra los resultados del pretest y posttest del nivel de PC tras la implementación de la intervención AIG (grupo experimental) y la metodología tradicional (grupo control).

Tabla 2

Estadísticos descriptivos de ambos grupos del pretest y posttest

	Grupo experimental				Grupo control			
	Pretest M (SD)	Posttest M (SD)	Gan.	Ap. N.	Pretest M (SD)	Posttest M (SD)	Gan.	Ap. N.
PC Total	20.8 (5.08)	23.8 (4.60)	3.06	26,1	20.8 (4.34)	21.1 (4.43)	0.320	10,8
Mujeres	20.6 (4.53)	23.6 (4.81)	3.06	25,8	20.9 (4.33)	21.2 (4.39)	0.347	10,9
Hombres	20.9 (5.67)	23.9 (4.53)	3.05	25,8	16.0 (5.0)	15.0 (5.0)	-1.00	6,25

En el grupo experimental, el pretest mostró una media de PC de 20.8 (SD = 5.08) que ascendió a 23.8 (SD = 4.60) en el posttest, lo que se traduce en una ganancia media de 3.06 puntos y un aprendizaje normalizado de 26.1 para el total de participantes. Desglosando por género, las mujeres pasaron de una media de 20.6 (SD = 4.53) a 23.6 (SD = 4.81), con una ganancia idéntica de 3.06 y un aprendizaje normalizado de 25.8, mientras que los hombres mostraron una mejora de 20.9 (SD = 5.67) a 23.9 (SD = 4.53), con una ganancia de 3.05 y un aprendizaje normalizado de 25.8. En el grupo control, el pretest reveló una media de 20.8 (SD = 4.34), que incrementó ligeramente a 21.1 (SD = 4.43) en el posttest, con una mínima ganancia promedio de 0.320 y un aprendizaje normalizado de 10.8 para el total de participantes. Analizando por género, las mujeres en el grupo control partieron de una media de 20.9 (SD = 4.33) y llegaron a 21.2 (SD = 4.39), teniendo una ganancia de 0.347 y un aprendizaje normalizado de 10.9; por su parte, los hombres tuvieron una reducción desde una media de 16.0 (SD = 5.0) a 15.0 (SD = 5.0), con una pérdida media de -1.00 y un aprendizaje normalizado de 6.25.

Finalmente, para profundizar en este análisis comparado, en la Figura 3 y en la Tabla 3 se muestran las medias marginales estimadas de los grupos experimental y control tras el ajuste por pretest (Tabla 3), así como su representación gráfica (Figura 3). Estos resultados permiten visualizar la diferencia entre ambos grupos y el efecto de la intervención.

Figura 3
Medias marginales estimadas para la variable posttest en el análisis de covarianza (ANCOVA)

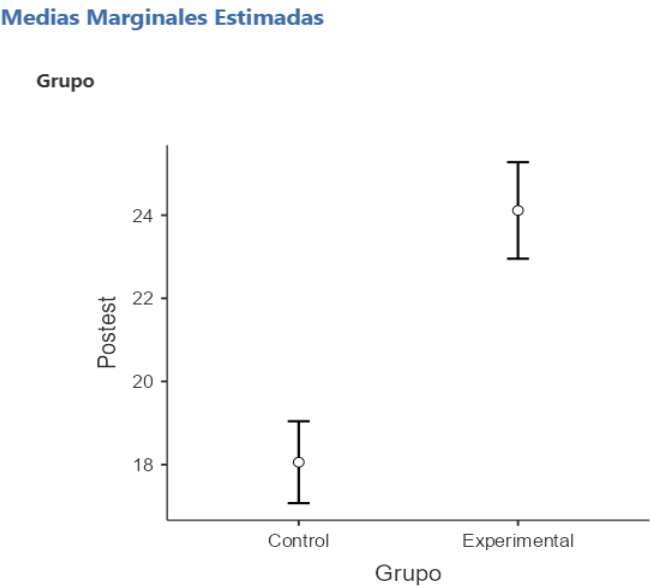


Tabla 3
Medias marginales estimadas para el Posttest tras el ajuste por Pretest

Medidas Marginales Estimadas - Grupo				
Grupo	Media	EE	Intervalo de Confianza al 95 %	
			Inferior	Superior
Control	18.1	0.495	17.1	19.0
Experimental	24.1	0.584	23.0	25.3

El análisis de covarianza (ANCOVA) mostró una diferencia significativa entre los grupos experimental y control después de ajustar por el pretest. En la comparación de medias marginales estimadas (Tabla 3), el grupo experimental obtuvo una media ajustada de 24.1 (IC 95 %: 23.0 - 25.3), mientras que el grupo control presentó una media ajustada de 18.1 (IC 95 %: 17.1 - 19.0). La diferencia entre los grupos fue de 6.0 puntos, indicando un mejor desempeño del grupo experimental. El efecto del grupo en el modelo ANCOVA fue significativo ($p < .001$), lo que sugiere que

la intervención aplicada tuvo un impacto sobre la variable dependiente. Además, como se observa en la Figura 3, los intervalos de confianza de ambas medias no se solapan, lo que refuerza la existencia de una diferencia real entre los grupos y la robustez del efecto encontrado.

DISCUSIÓN Y CONCLUSIONES

El presente estudio se centró en evaluar la efectividad de la gamificación combinada con el aprendizaje invertido para potenciar el pensamiento computacional en estudiantes universitarios del grado de Maestro/a en Infantil. La muestra contó con la participación de 86 estudiantes de entre 20 y 40 años, distribuidos en dos grupos. El grupo experimental, conformado por 36 estudiantes, recibió una intervención basada en gamificación y aula invertida, completando tanto el pretest como el postest. El grupo control, integrado por 50 estudiantes, siguió una enseñanza tradicional y también realizó ambas pruebas. El tamaño de la muestra se calculó con un nivel de confianza del 95 % y un margen de error del 6 %, obteniendo un resultado de 86 participantes, cantidad que coincide exactamente con los estudiantes que finalizaron el estudio. Esto asegura que la muestra es representativa y respalda la fiabilidad de los hallazgos.

La intervención educativa AIG se ha llevado a cabo en cinco fases durante seis semanas. En la primera semana, los estudiantes investigaron conceptos relacionados con el pensamiento computacional fuera del aula. Las dos semanas siguientes se dedicaron a debatir sobre el uso de PCs y robótica, con prácticas en plataformas como code.org y Scratch. La tercera fase involucró la creación de materiales educativos en equipo. En la cuarta fase, los equipos presentaron sus productos finales, y en la última semana realizaron autoevaluaciones y coevaluaciones, fomentando un aprendizaje colaborativo y reflexivo sobre el pensamiento computacional. Este enfoque fomentó que los estudiantes adoptaran un papel activo en su aprendizaje, obteniendo habilidades sobre el PC de manera autónoma. La composición de la muestra refleja la realidad del Grado en Educación Infantil, donde la matrícula estaba mayoritariamente formada por mujeres. Si bien esta distribución no afecta la validez de los resultados, es un factor a considerar en su interpretación y en la posibilidad de extrapolarlos a contextos con una representación de género más equilibrada.

En relación con la pregunta de esta investigación, los resultados indican que, en general, en los dos grupos (experimental y control) existe una mejora tras la intervención con ambas metodologías utilizadas (tradicional y AIG). Sin embargo, los resultados de la investigación demuestran que el grupo que participó en la metodología AIG obtuvo mejores resultados que el grupo control. Es decir, existe una mejora en las habilidades de pensamiento computacional tras la intervención educativa en el grupo experimental, lo que sugiere que el uso del AIG puede ser más efectivo que la metodología tradicional para trabajar las habilidades del PC. Esto coincide con los hallazgos de estudios como el de Zainuddin (2018), donde

se observó un mayor rendimiento y compromiso en estudiantes que participaron en un aula invertida gamificada. La incorporación de componentes de la estrategia metodológica AIG ha tenido un efecto positivo sobre los resultados del grupo experimental, estos resultados coinciden con los obtenidos por Recabarren et al. (2021). Los resultados de este estudio están también en consonancia con otras investigaciones, como las de Huang y Hew (2019), que han demostrado los beneficios de la gamificación y el aprendizaje invertido en términos de mejora de los resultados académicos. La mejora observada en el pensamiento computacional entre los participantes del grupo experimental respalda la idea de que estas metodologías pedagógicas innovadoras pueden ser particularmente efectivas en el ámbito de la educación superior especialmente en futuros docentes.

Los resultados obtenidos indican que la intervención aplicada al grupo experimental tuvo un efecto positivo en comparación con el grupo control. La diferencia de 6.0 puntos en la media ajustada, junto con la significancia estadística ($p < .001$), sugiere que el tratamiento implementado contribuyó a una mejora sustancial en el desempeño de los participantes. Estos hallazgos son consistentes con estudios previos que han reportado mejoras similares tras la aplicación de metodologías específicas en contextos similares. La ausencia de solapamiento en los intervalos de confianza refuerza la robustez del efecto observado y reduce la probabilidad de que la diferencia sea atribuible al azar. Sin embargo, es importante considerar que, si bien la diferencia entre los grupos es clara desde un punto de vista estadístico, su impacto práctico debe evaluarse en función del contexto y la magnitud del efecto. Estudios futuros podrían analizar el impacto a largo plazo de la intervención y explorar la influencia de otras variables que pudieran estar mediando en los resultados obtenidos. Estos resultados pueden implicar que, en el contexto de la formación de docentes, la integración de estrategias didácticas invertidas y gamificadas fomenta no solo la adquisición de conocimientos, sino también la capacidad de aplicarlos en situaciones prácticas. Esto concuerda con una investigación parecida que subraya los beneficios de usar un modelo de aula invertida gamificada en la educación universitaria (Carpena Arias y Esteve Mon, 2022a). La combinación de pensamiento computacional y gamificación en la formación docente se presenta como una estrategia prometedora, pero es necesario un diseño cuidadoso y posiblemente una intervención más prolongada para maximizar sus beneficios (Bueno-Baquero et al., 2024).

A pesar de los resultados, es importante reconocer las limitaciones del estudio. La naturaleza cuantitativa del análisis podría ser complementada con datos cualitativos, como entrevistas para obtener una comprensión más profunda de cómo los estudiantes experimentan y se involucran con la metodología de aula invertida gamificada. Además, para futuras investigaciones, será esencial basar nuestra intervención didáctica en el desarrollo del pensamiento computacional en algún modelo existente, como por ejemplo el modelo 5PC. Este modelo propone una secuencia didáctica de cinco fases que permite integrar el PC en la formación

docente. Estas son: Problemas reales, Actividades desenchufadas, Programación de dispositivos digitales, Programación de objetos físicos y Exposición de recursos (Serrano y Serrano, 2024). Este enfoque garantizará que las estrategias pedagógicas empleadas estén fundamentadas en teorías y prácticas educativas probadas, lo que permitirá una implementación más efectiva y resultados medibles en el aprendizaje del pensamiento computacional. La adopción de modelos teóricos y prácticos validados proporcionará una estructura más sólida para evaluar y mejorar continuamente las estrategias de enseñanza del PC.

En conclusión, este estudio demuestra cómo la combinación de las metodologías didácticas de aula invertida y gamificación puede ser una estrategia eficaz para potenciar el pensamiento computacional en estudiantes de educación superior, especialmente en la formación inicial docente. El AIG no solo facilita la adquisición de habilidades del PC, sino que también promueve la autonomía y la motivación intrínseca de los estudiantes. Sin embargo, se requiere más investigación para confirmar y expandir estos resultados en diferentes contextos educativos de la formación inicial del profesorado y con muestras más amplias, para comprender completamente los mecanismos subyacentes y optimizar la adaptación de estos enfoques a diversas situaciones de enseñanza y aprendizaje. Por otro lado, será necesario incorporar un enfoque metodológico específico del PC en la formación inicial para poderlo combinar con el AIG y de esta manera comprobar si mejoran aún más las habilidades del pensamiento computacional.

REFERENCIAS

- BERA [British Educational Research Association]. (2018). *Pautas éticas para la investigación educativa*. BERA.
- Bergmann, J. y Sams, A. (2009). Remixing chemistry class: Two Colorado teachers make vodcasts of their lectures to free up class time for hands-on activities. *Learning & Leading with Technology*, 36(4), 22-27.
- Bergmann, J. y Sams, A. (2014). *Flipped learning: Gateway to student engagement*. International Society for Technology in Education.
- Bishop, J. L. y Verleger, M. A. (2013). The flipped classroom: A survey of the research. *ASEE Annual Conference & Exposition*, 1-18. <https://doi.org/10.18260/1-2--22585>
- Bueno-Baquero, A., del Olmo Muñoz, J., González-Calero, J. A. y Cózar-Gutiérrez, R. (2024). Enfoques gamificados de pensamiento computacional en formación docente. *Revista de Educación*, 405, 71-102. <https://doi.org/10.4438/1988-592X-RE-2024-405-628>
- Candel, E. C., de-la-Peña, C. y Yuste, B. C. (2024). Pre-service teachers' perception of active learning methodologies in history: Flipped classroom and gamification in an e-learning environment. *Education and Information Technologies*, 29, 3365-3387. <https://doi.org/10.1007/s10639-023-11924-0>
- Carpena Arias, J. y Esteve Mon, F. M. (2022a). Aula invertida gamificada como estrategia pedagógica en la educación superior: Una revisión sistemática. *EDUTEC. Revista Electrónica de Tecnología Educativa*, 80, 84-98. <https://doi.org/10.21556/edutec.2022.80.2435>

- Carpena Arias, J. y Esteve Mon, F. M. (2022b). Diseño y validación de una intervención educativa gamificada para trabajar el pensamiento computacional en la formación inicial docente. En XXV Congreso Internacional EDUTEC 2022. Repositorio Institucional UIB. <http://hdl.handle.net/11201/160517>
- Carpena Arias, J., Esteve Mon, F. M. y Valdeolivas, M. G. (2022). La experiencia de aula invertida gamificada en futuros docentes para trabajar el pensamiento computacional: Percepción de los estudiantes. En E. A. Aveleyra y M. A. Proyetti Martino (Eds.), *Escenarios y recursos para la enseñanza con tecnología: Desafíos y retos* (p. 817). Octaedro.
- Carpena Arias, J. y Esteve Mon, F. M. (2024). Diseño y validación de una estrategia didáctica gamificada para desarrollar el pensamiento computacional en futuros docentes. *RiiTE: Revista Interuniversitaria de Investigación en Tecnología Educativa*, 16, 62-79. <https://doi.org/10.6018/riite.591681>
- Collado-Sánchez, M., Pinto-Llorente, A. M. y García-Peñalvo, F. J. (2023). Pensamiento computacional en el profesorado de primaria: Una revisión sistemática. *Campus Virtuales*, 12(2), 147-162. <https://doi.org/10.54988/cv.2023.2.1418>
- Ekici, M. (2021). A systematic review of the use of gamification in flipped learning. *Education and Information Technologies*, 26(3), 3327-3346. <https://doi.org/10.1007/s10639-020-10394-y>
- Elzeky, M. E. H., Elhabashy, H. M. M., Ali, W. G. M. y Allam, S. M. E. (2022). Effect of gamified flipped classroom on improving nursing students' skills competency and learning motivation: A randomized controlled trial. *BMC Nursing*, 21, Artículo 316. <https://doi.org/10.1186/s12912-022-01096-6>
- Gómez-Carrasco, C. J., Monteagudo-Fernández, J., Moreno-Vera, J. R. y Sainz-Gómez, M. (2019). Effects of a gamification and flipped-classroom program for teachers in training on motivation and learning perception. *Education Sciences*, 9(4), Artículo 299. <https://doi.org/10.3390/educsci9040299>
- Huang, B. y Hew, K. F. (2019). Implementing a theory-driven gamification model in higher education flipped courses: Effects on out-of-class activity completion and quality of artifacts. *Computers & Education*, 125, 254-272. <https://doi.org/10.1016/j.compedu.2018.06.018>
- INTEF. (2021). *Escuela de pensamiento computacional e inteligencia artificial 20/21: Enfoques y propuestas para su aplicación en el aula. Resultados de la investigación* [Informe técnico]. Ministerio de Educación y Formación Profesional. <https://code.intef.es/epcia/informes-epcia/>
- The jamovi project. (2022). *jamovi* [Software]. <https://www.jamovi.org>
- Lage, M. J., Platt, G. J. y Treglia, M. (2000). Inverting the classroom: A gateway to creating an inclusive learning environment. *The Journal of Economic Education*, 31(1), 30-43. <https://doi.org/10.1080/00220480009596759>
- Marczewski, A. (2014, 25 de marzo). Gamification design vs game design. *Gamified UK*. <http://www.gamified.uk/2014/03/25/gamification-design-vs-game-design/>
- Martín, D. y Tourón, J. (2017). El enfoque flipped learning en estudios de magisterio: Percepción de los alumnos. *RIED-Revista Iberoamericana de Educación a Distancia*, 20(2), 187-211. <https://doi.org/10.5944/ried.20.2.17704>
- Ocón-Galilea, R. (2017). La gamificación en educación y su trasfondo pedagógico. *Revista Electrónica de e-learning Innova*, (187), 1-10. Universidad Complutense de Madrid. <https://webs.ucm.es/BUCM/revcul/e-learning-innova/187/art2664.pdf>

- Ortuño, G. y Serrano, J. L. (2024). Implementación y formación del profesorado de educación primaria en pensamiento computacional: Una revisión sistemática. *RIED. Revista Iberoamericana de Educación a Distancia*, 27(1), 255-287. <https://doi.org/10.5944/ried.27.1.37572>
- Plomp, T. y Nieveen, N. (2009). *An introduction to educational design research*. Netherlands Institute for Curriculum Development (SLO).
- Prieto, J. M. (2020). Una revisión sistemática sobre gamificación, motivación y aprendizaje en universitarios. *Teoría de la Educación. Revista Interuniversitaria*, 32(1), 73-99. <https://doi.org/10.14201/teri.20625>
- Real Decreto 95/2022, de 1 de febrero, por el que se establece la ordenación y las enseñanzas mínimas de la Educación Infantil. *Boletín Oficial del Estado*, núm. 28, 2 de febrero de 2022. <https://www.boe.es/boe/dias/2022/02/02/pdfs/BOE-A-2022-1654.pdf>
- Recabarren, M., Corvalán, B. y Villegas, M. (2021). Exploring the differences between gamer and non-gamer students in the effects of gamification on their motivation and learning. *Interactive Learning Environments*, 31(6), 3529-3542. <https://doi.org/10.1080/10494820.2021.1933543>
- Román-González, M. (2015). Computational thinking test: Design guidelines and content validation. En *EDULEARN15 Proceedings: 7th International Conference on Education and New Learning Technologies* (pp. 2436-2444). IATED. <https://library.iated.org/view/ROMANGONZALEZ2015COM>
- Román-González, M. (2022, 22 de julio). Pensamiento computacional: Un constructo que llega a la madurez. *Aula Magna 2.0*. <https://cuedespyd.hypotheses.org/11109>
- Sacristán, M., Martín, D., Navarro, E. y Tourón, J. (2017). Flipped classroom y didáctica de las matemáticas en la formación online de maestros de educación infantil. *Revista Electrónica Interuniversitaria de Formación del Profesorado*, 20(3), 1-14. <https://doi.org/10.6018/reifop.20.3.292551>
- Sailer, M. y Sailer, M. (2021). Gamification of in-class activities in flipped classroom lectures. *British Journal of Educational Technology*, 52(1), 75-90. <https://doi.org/10.1111/bjet.12948>
- Santiago, R. y Bergmann, J. (2018). *Aprender al revés: Flipped Classroom 3.0 y metodologías activas en el aula*. Paidós.
- Serrano, J. L. (2022). *Pensamiento computacional en educación: Kit de conocimientos para antes de comprar y programar un robot*. RC Libros.
- Serrano, J. L. y Ortuño, G. (2021). Percepciones del profesorado en formación sobre el desarrollo del pensamiento computacional desde el Modelo 5PC. *EduTec. Revista Electrónica de Tecnología Educativa*, 78, 212-230. <https://doi.org/10.21556/edutec.2021.78.2173>
- Soriano-Sánchez, J. y Jiménez-Vázquez, D. (2023). Prácticas educativas innovadoras en la educación superior: Una revisión sistemática. *Revista Innova Educación*, 5(1), 23-37. <https://doi.org/10.35622/j.rie.2023.05.002>
- Tourón, J. y Santiago, R. (2014). *The flipped classroom: Cómo convertir la escuela en un espacio de aprendizaje*. Grupo Océano.
- Villalustre-Martínez, L. (2024). Análisis del nivel de pensamiento computacional de los futuros maestros: Una propuesta diagnóstica para el diseño de acciones formativas. *Píxel-Bit. Revista de Medios y Educación*, 69, 169-194. <https://doi.org/10.12795/pixelbit.101205>
- Werbach, K. y Hunter, D. (2015). *The gamification toolkit: Dynamics*,

- mechanics, and components for the win.* Wharton Digital Press.
- Wing, J. M. (2006). Computational thinking. *Communications of the ACM*, 49(3), 33-35. <https://doi.org/10.1145/1118178.1118215>
- Zainuddin, Z. (2018). Students' learning performance and perceived motivation in gamified flipped-class instruction. *Computers & Education*, 126, 75-88. <https://doi.org/10.1016/j.compedu.2018.07.003>
- Zamora-Polo, F., Corrales-Serrano, M., Sánchez-Martín, J. y Espejo-Antúnez, L. (2019). Non-scientific university students training in general science using an active-learning merged pedagogy: Gamification in a flipped classroom. *Education Sciences*, 9(4), Artículo 297. <https://doi.org/10.3390/educsci9040297>

Fecha de recepción del artículo: 1 de diciembre de 2024

Fecha de aceptación del artículo: 12 de marzo de 2025

Fecha de aprobación para maqueta: 10 de abril de 2025

Fecha de publicación en OnlineFirst: 25 de abril de 2025

Fecha de publicación: 1 de julio de 2025

La Política Editorial de la *RIED. Revista Iberoamericana de Educación a Distancia*, se concreta en los siguientes criterios:

- **De la AIESAD.** La RIED. *Revista Iberoamericana de Educación a Distancia* se configura como el instrumento de la Asociación Iberoamericana de Educación Superior a Distancia (AIESAD) para la difusión de trabajos de carácter científico, experiencias, convocatorias e información bibliográfica, dentro del ámbito de la enseñanza/aprendizaje abierto y a distancia en sus diferentes formulaciones y presentaciones.
- **Arbitrada.** La RIED es una publicación arbitrada que utiliza el sistema de evaluación externa de revisión por pares (doble ciego), identificándose cada trabajo con un DOI (*Digital Object Identifier System*).
- **Periodicidad y formato.** La RIED. *Revista Iberoamericana de Educación a Distancia*, tiene una periodicidad semestral (un volumen anual con dos números). Se edita en doble versión: impresa (ISSN: 1138-2783) y electrónica (E-ISSN: 1390-33061).
- **Idioma de los trabajos.** Podrán presentarse trabajos en lengua española, portuguesa e inglesa.
- **Requisitos.** Toda propuesta de colaboración deberá reunir los siguientes requisitos:
 - hacer referencia al campo de especialización propio de la RIED;
 - estar científicamente fundada y gozar de unidad interna;
 - suponer una ayuda para la profundización en las diversas dimensiones y ámbitos de la educación abierta y a distancia y de las TIC aplicadas a la educación.
 - Se primarán los trabajos sujetos al modelo IMyD (*Introducción, Metodología, Resultados y Discusión*) y que puedan tener incidencia en la educación superior.
- **Trabajo original.** Los trabajos enviados a la RIED para su publicación deberán constituir una colaboración original no publicada previamente en soporte alguno, ni encontrarse en proceso de publicación o valoración en cualquiera otra revista o proyecto editorial.
- **Normas de redacción y presentación.** Los trabajos deberán atenerse a las normas de redacción y presentación de carácter formal de la RIED. Las colaboraciones enviadas a la RIED que no se ajusten a ellas serán desestimadas.
- **Recepción de originales.** La Secretaría de la RIED acusará la recepción del manuscrito enviado por el autor/es. El Consejo de Redacción revisará el artículo enviado informando al autor/es, en caso necesario, si se adecua al campo temático de la revista y al cumplimiento de las normas y requisitos formales de redacción y presentación. En el caso de que todos los aspectos sean favorables, se procederá a la revisión por pares del artículo.
- **Revisión externa.** Antes de la publicación, los manuscritos enviados serán valorados de forma anónima por dos miembros del Comité Científico o Evaluadores Externos (revisión por pares), por el sistema de doble ciego que, en su caso, realizarán sugerencias para la revisión y mejora en vistas a la elaboración de una nueva versión. Para la publicación definitiva se requiere la valoración positiva de ambos revisores. En caso de controversia evidente por parte de éstos, se requerirá de una tercera valoración para su aceptación, modificación o rechazo definitivos de la publicación.
- **Criterios de Evaluación del Comité Científico y Evaluadores Externos.** Los criterios de valoración de cada artículo que justifican la decisión de aceptación/modificación/rechazo se basan en los siguientes ejes:
 - interés del campo de estudio al ámbito de los formatos educativos no presenciales, prioritariamente con posible incidencia en la educación superior.
 - relevancia, originalidad e información valiosa de las aportaciones,
 - aplicabilidad de los resultados para la resolución de problemas.
 - actualidad y novedad,
 - avance del conocimiento científico,
 - fiabilidad y validez científica: calidad metodológica contrastada,
 - correcta organización, redacción y estilo de la presentación del material.
- **Información.** La Secretaría de la RIED informará a los autores de la decisión de aceptación, modificación y rechazo de cada uno de los artículos. La corrección de pruebas de imprenta la hará la RIED cotejando con el original.
- **Política de privacidad:** Se mantendrá y preservará en todos los casos y circunstancias el anonimato de los autores y el contenido de los artículos desde la recepción del manuscrito hasta su publicación. La información obtenida en el proceso de revisión y evaluación tendrá carácter confidencial.
- **Fuentes.** Los autores citarán debidamente las fuentes de extracción de datos, figuras e información de manera explícita y tangible tanto en la bibliografía, como en las referencias. Si el incumplimiento se detectase durante el proceso de revisión o evaluación se desestimará automáticamente la publicación del artículo.
- **Responsabilidad.** RIED no se hará responsable de las ideas y opiniones expresadas en los trabajos publicados. La responsabilidad plena será de los autores de los mismos.
- **Licencia.** Los textos publicados en esta revista están sujetos a una Licencia Creative Commons Atribución 4.0 Internacional". Puede copiarlos, distribuirlos, comunicarlos públicamente, hacer obras derivadas y usos comerciales siempre que reconozca los créditos de las obras (autoría, nombre de la revista, institución editora) de la manera especificada por los autores o por la propia RIED.

OTRAS INFORMACIONES DE INTERÉS

- Declaración ética sobre publicación y malas prácticas: <https://revistas.uned.es/index.php/ried/about#custom-o>
- Directrices para autores. Normas para publicar en RIED: <https://revistas.uned.es/index.php/ried/about/submissions#authorGuidelines>
- Lista de comprobación previa de los envíos: <https://revistas.uned.es/index.php/ried/about/submissions>

Revista Iberoamericana de Educación a Distancia

ChatGPT, IA generativa y tecnología educativa: aportes y desafíos en la educación virtual

ESTUDIOS E INVESTIGACIONES:

Asistentes ChatGPT en educación superior en línea y satisfacción del alumnado: un caso de estudio

Agente de Inteligencia Artificial Generativa en investigación científica. Un análisis explicativo del aprendizaje en el aula

Diseño y análisis de narrativas enriquecidas con inteligencia artificial. Proyecto de transferencia UNED - UTMACH

Inteligencia Artificial y chatbots para una educación superior sostenible: una revisión sistemática

Percepciones de futuros docentes y pedagogos sobre uso responsable de la IA. Un instrumento de medida

Del optimismo a la confianza: el impacto de ChatGPT en la confianza de los estudiantes en el aprendizaje asistido por IA

Efectos de la conciencia y confianza en la disposición estudiantil para usar ChatGPT: modelo TAMECM integrado

Marcos normativos para una IA ética y confiable en la educación superior: estado de la cuestión

Desafíos de la Inteligencia Artificial generativa en educación superior: fomentando su uso crítico en el estudiantado

IA generativa y pensamiento crítico en la educación universitaria a distancia: desafíos y oportunidades

El estudiantado de secundaria ante la RV en materias STEM. Efecto de la variable de género

Impacto de variables sociodemográficas en la percepción del alumnado universitario sobre la educación en línea post-COVID-19

Adopción del aprendizaje en línea en la educación superior de China: la perspectiva de los docentes

Propiedades psicométricas de una escala para medir estilos de enseñanza en línea en docentes universitarios

Despertando el compromiso estudiantil: el poder transformador de H5P en la educación superior

Aula invertida gamificada: aplicación de una estrategia didáctica para trabajar el pensamiento computacional en futuros docentes

