



8. El uso de ChatGPT para personalizar el aprendizaje en estudiantes universitarios: una revisión sistemática de la literatura con el modelo PRISMA

The Use of ChatGPT to Personalize Learning for University Students: A PRISMA-Guided Systematic Literature Review

Karla Karina Ruiz Mendoza ¹ @  Rubi Surema Peniche Cetzal ² @ 

^{1 y 2} Universidad Autónoma de Baja California, Tijuana y Ensenada, México.

RESUMEN

Esta revisión sistemática (PRISMA 2020) examina de qué manera ChatGPT se emplea para personalizar el aprendizaje en educación superior (2022–2025) y qué condiciones favorecen su adopción. Se realizaron búsquedas en Web of Science, Scopus y ERIC, complementadas con búsquedas hacia atrás y hacia adelante por citas. Se incluyeron 38 estudios empíricos revisados por pares con uso explícito de ChatGPT y evidencia de personalización. La producción se intensifica desde 2023 y se distribuye en múltiples disciplinas. Los hallazgos sugieren que la estrategia más robusta es la retroalimentación individual y oportuna, así como apoyos activados por información del estudiante, con mejoras reportadas en desempeño, comprensión, autorregulación y compromiso. Otras modalidades parciales —ajuste del tono/persona, itinerarios de aprendizaje y explicaciones personalizadas— refinan la experiencia, aunque con evidencia aún heterogénea. A la par, persiste un déficit de transparencia técnica (versión y parámetros del modelo), lo que limita la replicabilidad y comparación. Se recomienda una personalización con gobernanza: supervisión humana, lineamientos de transparencia e integridad, alfabetización en IAGen e infraestructura equitativa; en el aula, diseños centrados en retroalimentación, metas y secuencias adaptadas, activación de habilidades de orden superior y trazabilidad técnica. Estudios futuros deberían privilegiar diseños longitudinales y experimentales, junto con reportes estandarizados de configuración y versiones de modelos.

Palabras clave: inteligencia artificial; tecnología educacional; educación superior; estudiante universitario; enseñanza individualizada

The use of ChatGPT to personalize learning for university students: A PRISMA-guided systematic literature review

ABSTRACT

This PRISMA 2020-guided systematic review examines how ChatGPT is used to personalize learning in higher education (2022–2025) and what conditions enable its adoption. Searches across Web of Science, Scopus, and ERIC were complemented with backward and forward citation chasing. We included 38 peer-reviewed empirical studies featuring explicit ChatGPT use and evidence of personalization. Publication output has accelerated since 2023 and spans multiple disciplines. Findings indicate that the most consistent pathway is individualized, timely feedback and data-triggered support, with reported gains in achievement, comprehension, self-regulation, and engagement. Partial modalities—tone/persona adjustment, learning itineraries, and personalized explanations—further refine the experience, albeit with heterogeneous evidence. Technical transparency remains limited, as model versions and key parameters are seldom reported, constraining replicability and cross-study comparison. Governance-aware personalization is therefore recommended: human oversight, transparency and integrity policies, AI literacy, and equitable infrastructure; instructionally, prioritize feedback-first designs, goal-linked sequences and pacing, higher-order thinking activation, and technical traceability. Future research should employ longitudinal and experimental designs and standardized reporting of model configurations and versions.

Keywords: Artificial Intelligence; Educational Technology; Higher Education; University Student; Individualized Instruction

O uso do ChatGPT para personalizar a aprendizagem em estudantes universitários: uma revisão sistemática da literatura com o modelo PRISMA

RESUMO

Esta revisão sistemática (PRISMA 2020) analisa como o ChatGPT é utilizado para personalizar a aprendizagem no ensino superior (2022–2025) e quais condições favorecem sua adoção. Foram realizadas buscas na Web of

Science, Scopus e ERIC, complementadas por encadeamento de citações. Incluíram-se 38 estudos empíricos revisados por pares com uso explícito do ChatGPT e evidência de personalização. A produção cresce desde 2023 e é interdisciplinar. Os achados indicam que a via mais consistente é o feedback individual e oportuno e os apoios acionados por dados do estudante, com melhorias relatadas em desempenho, compreensão, autorregulação e engajamento. Modalidades parciais —ajuste de tom/persona, itinerários de aprendizagem e explicações personalizadas— refinam a experiência, embora com evidência ainda heterogênea. Persiste um déficit de transparência técnica (versão e parâmetros do modelo), o que limita a replicabilidade e a comparação. Recomenda-se uma personalização com governança: supervisão humana, políticas de transparência e integridade, letramento em IAGen e infraestrutura equitativa. Pesquisas futuras devem priorizar desenhos longitudinais/experimentais e relato padronizado de versões e configurações de modelos.

Palavras-chave: inteligência artificial; tecnologia educacional; ensino superior; estudante universitário; instrução individualizada

L'utilisation de ChatGPT pour personnaliser l'apprentissage chez les étudiants universitaires : une revue systématique de la littérature selon le modèle PRISMA

RÉSUMÉ

Cette revue systématique (PRISMA 2020) examine comment ChatGPT est utilisé pour personnaliser l'apprentissage dans l'enseignement supérieur (2022–2025) et quelles conditions facilitent son adoption. Des recherches ont été menées dans Web of Science, Scopus et ERIC, complétées par un chaînage de citations. Trente-huit études empiriques évaluées par les pairs et présentant un usage explicite de ChatGPT avec des indices de personnalisation ont été incluses. La production s'intensifie depuis 2023 et couvre plusieurs disciplines. Les résultats suggèrent que la voie la plus robuste est le feedback individualisé et opportun, ainsi que des soutiens déclenchés par des données de l'étudiant, avec des gains rapportés en performance, compréhension, autorégulation et engagement. Des modalités partielles —ajustement du ton/de la persona, itinéraires d'apprentissage et explications personnalisées— affinent l'expérience, mais avec des preuves encore hétérogènes. Un déficit de transparence technique persiste (version et paramètres du modèle), limitant la reproductibilité et la comparaison.

Une personnalisation avec gouvernance est recommandée : supervision humaine, politiques de transparence et d'intégrité, littératie en IAGén et infrastructure équitable. Les recherches futures devraient privilégier des devis longitudinaux/expérimentaux et un rapport standardisé des versions et configurations de modèles.

Mots clés: intelligence artificielle; technologie éducative; enseignement supérieur; étudiant universitaire; enseignement individualisé

1. INTRODUCCIÓN

Con el artículo mediático *Your Brain on ChatGPT: Accumulation of Cognitive Debt when using an AI assistant for essay writing task* (Kosmyna et al., 2025), la pregunta principal fue si la inteligencia artificial generativa (IAGén) nos hará más estúpidos. Este estudio y otros advierten que delegar la escritura y el razonamiento en un asistente puede reducir el esfuerzo cognitivo y debilitar la evaluación crítica, acumulando una posible “deuda cognitiva” (Kosmyna et al., 2025; Risko & Gilbert, 2016; Stadler et al., 2024). En educación superior, la cuestión no es solo si se usa, sino cómo: bajo qué condiciones opera como andamiaje para aprender y no como sustituto del pensamiento.

Ante esto, la personalización del aprendizaje se entiende como un enfoque centrado en el estudiante que ajusta contenidos, ritmo y apoyos según necesidades, intereses y capacidades, con el propósito de mejorar eficacia y equidad (Hattie, 2008). Aunque sus raíces pueden rastrearse en la tradición de la experiencia guiada y la mediación sociocultural (Dewey, 1916; Vygotsky, 1978), el debate contemporáneo en educación superior se centra en cómo operacionalizar esa adaptación con criterios pedagógicos y evidencia.

En cuanto a su efectividad, la evidencia indica que la instrucción personalizada mejora el rendimiento cuando se acompaña de objetivos claros, andamiaje y retroalimentación de alta calidad (Bloom, 1984; Hattie, 2008). No obstante, su implementación en educación superior enfrenta restricciones de escala (tamaño de grupos, heterogeneidad disciplinar y tiempos docentes), lo que ha impulsado el uso de tecnologías digitales y analíticas de aprendizaje para sostener decisiones pedagógicas diferenciadas (Liu et al., 2020).

Para precisar el componente tecnológico que posibilita muchas de estas propuestas, conviene definir qué se entiende por IAGén. En términos opera-

tivos, la IAGen alude a sistemas capaces de producir contenido nuevo (texto, imagen, audio o código) a partir de patrones aprendidos; en el ámbito del lenguaje natural, esta capacidad se materializa principalmente mediante modelos de lenguaje de gran tamaño (LLM) basados en arquitecturas Transformer.

En la familia ChatGPT de OpenAI, la primera versión pública (afinada a partir de la serie GPT-3.5) se lanzó en 2022 (OpenAI, 2022). En marzo de 2023 OpenAI presentó GPT-4, con multimodalidad, mayor ventana de contexto y mejoras de rendimiento respecto a GPT-3.5 (OpenAI, 2023). Estudios comparativos confirman que estas diferencias de versión son sustantivas; por ejemplo, Meyer et al. (2024) documentaron mejoras consistentes de GPT-4 sobre GPT-3.5 en exámenes médicos estandarizados. En consecuencia, la investigación educativa requiere reportar con precisión la versión y las condiciones de uso del modelo.

A partir de esta caracterización, ChatGPT y otros modelos generativos se perfilan como herramientas para ofrecer retroalimentación inmediata, andamiaje textual y apoyo adaptativo en tareas de escritura, resolución de problemas y actividades formativas, favoreciendo rutas de aprendizaje más personalizadas (Naznin et al., 2025). Sin embargo, revisiones y estudios advierten riesgos de dependencia, errores del modelo, sesgos y problemas de integridad y privacidad que exigen políticas institucionales, formación docente y diseños instruccionales responsables (Fuchs, 2023; Imran & Almusharraf, 2023; Labadze et al., 2023). En consecuencia, integrar ChatGPT en educación superior implica equilibrar su potencial de personalización con prácticas críticas, verificables y éticas.

En este sentido, el objetivo de este estudio es explorar cómo se utiliza ChatGPT y cuáles son las recomendaciones derivadas para personalizar el aprendizaje en estudiantes universitarios, a partir de una revisión sistemática de la literatura basada en el modelo PRISMA 2020.

- R1.** ¿Conocen los modelos que han utilizado?
- R2.** ¿Cómo se utiliza ChatGPT para personalizar el aprendizaje en la universidad?
- R3.** ¿Cuáles son las recomendaciones que se pueden obtener para personalizar el aprendizaje en la universidad?

2. METODOLOGÍA

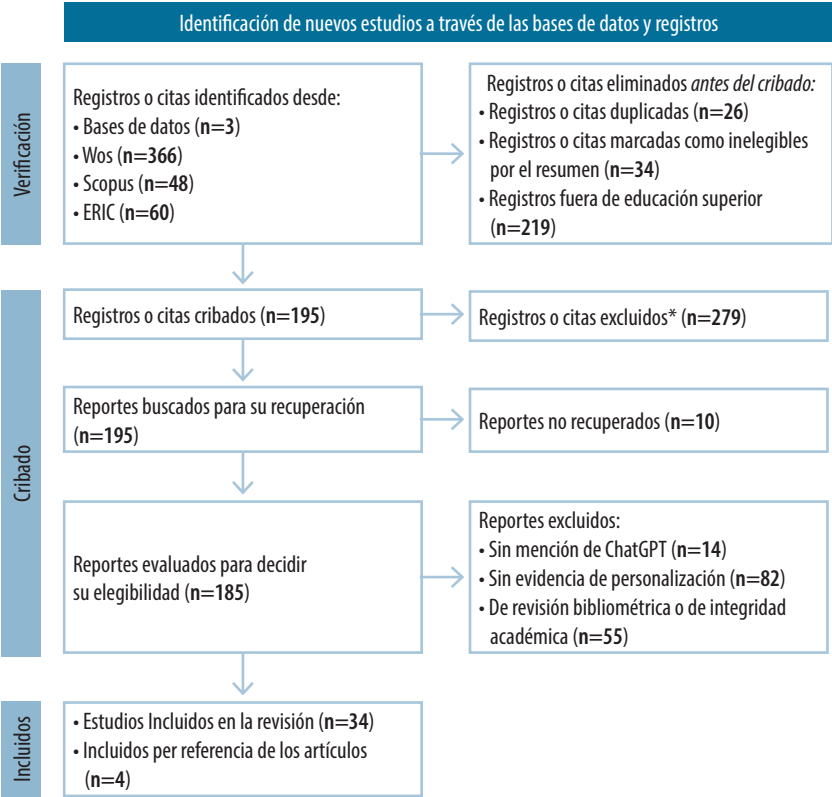
Para atender las preguntas, se realizó una revisión sistemática de la literatura (RSL) guiada por PRISMA 2020 para asegurar transparencia en la búsqueda, el cribado, la elegibilidad y la síntesis (Page et al., 2021). El protocolo se elaboró a priori; sin embargo, no se registró en un repositorio de protocolos (p. ej., OSF), lo que se consigna como limitación y se compensa con la descripción detallada de la estrategia de búsqueda y de los criterios de inclusión/exclusión.

La búsqueda cubrió desde enero de 2022 hasta septiembre de 2025 en Web of Science (n = 366), Scopus (n = 48) y ERIC (n = 60). Se priorizó el acceso abierto cuando estuvo disponible, sin excluir artículos de pago que cumplieran criterios. La fórmula booleana fue "chatgpt" AND ("personalización" OR "aprendizaje personalizado" OR "personalized learning" OR "personalization in learning" OR "individualized teaching" OR "personalized instruction" OR "adaptive learning" OR "differentiated instruction") AND ("higher education" OR "university" OR "university students" OR "college students"). Además, la búsqueda hacia atrás/adelante por referencias y citas permitió incorporar 4 registros pertinentes al flujo PRISMA.

Al final, el flujo PRISMA (véase Figura 1) fue de 474 registros identificados en bases de datos, más 4 registros recuperados mediante búsqueda complementaria por citas; 279 eliminados antes del cribado (duplicados, inelégibles por resumen y fuera de educación superior); 195 cribados y buscados para recuperación (10 no recuperados); 185 evaluados a texto completo; 151 excluidos (sin mención de ChatGPT, sin evidencia de personalización, o revisiones bibliométricas/integridad académica sin componente de personalización); donde se agregaron cuatro más a partir de los artículos revisados; por lo tanto, se incluyeron 38 estudios. La selección y la extracción de datos se realizaron de manera independiente.

Los criterios de inclusión fueron: educación superior; artículo empírico revisado por pares; uso explícito de ChatGPT; evidencia de personalización aunque no fuese mencionada (p. ej., retroalimentación individualizada, tutoría, adaptación de tareas, rutas de aprendizaje). Exclusión: niveles no universitarios; estudios de IAGen que no distinguen ChatGPT; trabajos centrados solo en integridad académica sin componente de personalización; preprints sin arbitraje.

Figura 1. Diagrama de flujo según PRISMA sobre aprendizaje personalizado y ChatGPT



Fuente: Adaptado de Page et al. (2020).

De los artículos incluidos (n = 38), aproximadamente un tercio implementa personalización explícita —adaptación en tiempo real o retroalimentación individual integrada al diseño— (Coban et al., 2025; Ding et al., 2025; Liew et al., 2025; Mahapatra, 2024; McGuire et al., 2024; Oktarin et al., 2024; Rezai et al., 2024; Santhosh et al., 2024; Sayed et al., 2024; Xie et al., 2025; Zekaj, 2023; Boudia & Bengueddach, 2024). Por otro lado, un segundo bloque describe personalización parcial —ajuste del tono/persona del agente, explicaciones personalizadas dentro de la retroalimentación, *co-prompting* o dinámicas de juego— (Perifanou & Economides, 2025; Jin

et al., 2025; Chan & Hu, 2023; Han et al., 2025; French et al., 2023; Husain, 2024; Întorsureanu et al., 2025). Finalmente, el resto no expone personalización de forma explícita, aunque aporta contexto sobre adopción, ética, alfabetización o rendimiento de LLM por disciplina y nivel (Oyelere & Aruleba, 2025; Riedel et al., 2023; Du et al., 2025; Beltozar-Clemente & Díaz-Vega, 2024; Mujtaba, 2024; Wei et al., 2025; Jereb & Urh, 2024; Pikhart et al., 2024; Iskender, 2023; Gencer & Gencer, 2024; Sallam & Al-Salahat, 2023; Morell-Mengual et al., 2025; Alrayes et al., 2024; Buele et al., 2025; Phua et al., 2025; Arbulú Ballesteros et al., 2024; Ngo et al., 2024; Stojanov et al., 2024; Firat, 2023).

3. RESULTADOS

3.1. Características generales de los estudios

La producción científica sobre los usos de ChatGPT para apoyar la personalización en educación superior se intensifica desde 2023. En el periodo analizado se registraron 7 estudios en 2023, 17 en 2024 y 14 en 2025 (corte: septiembre), lo que sugiere un campo en rápida expansión.

Asimismo, la Tabla 1 evidencia esta dispersión editorial: las 38 publicaciones se distribuyen en 33 revistas internacionales. Cinco revistas concentran dos estudios cada una (Frontiers in Education; International Journal of Educational Technology in Higher Education; Sustainability; Education and Information Technologies; y Journal of Information Technology Education: Research); el resto aporta un estudio por revista, lo que refuerza el carácter emergente del campo.

En términos generales, los estudios se concentran principalmente en Asia y Europa, con presencia menor en América, África y Oceanía. Esta distribución sugiere la necesidad de ampliar la evidencia en contextos latinoamericanos para evaluar transferibilidad, gobernanza y equidad.

3.2. Modelos utilizados

Una de las preguntas principales (R1) es si los estudios identifican la versión del modelo utilizado, aspecto crucial dada la velocidad de actualización de los sistemas de OpenAI. La Tabla 2 sintetiza este hallazgo (GPT-4o = 2; gpt-3.5-turbo = 1; "ChatGPT" no especificado = 16; NA = 19; N = 38): la mayoría no reporta versión/modelo o no usa un modelo en su metodología.

Tabla 1. Revistas ubicadas según la RSL	
Revista	Total
Cureus	1
Computers and Education Open	1
International Journal of Technology in Education (IJTE)	1
International Journal of Learning, Teaching and Educational Research	1
EPJ Quantum Technology	1
Frontiers in Education	2
Education and Information Technologies	2
The Electronic Journal of e-Learning	1
Frontiers in Medicine	1
Business Ethics and Leadership	1
Education Sciences	1
Multimodal Technologies and Interaction	1
Acta Psychologica	1
Sustainability	2
Journal of Medical Internet Research	1
Big Data and Cognitive Computing	1
International Journal of Educational Technology in Higher Education	2
IEEE Access	1
Electronics	1
European Journal of Tourism Research	1
Computers in Human Behavior Reports	1
Proceedings of ICRES (conferencia)	1
Edelweiss Applied Science and Technology	1
Nursing Reports	1
International Journal of Engineering Pedagogy (IJEP)	1
Journal of Information Technology Education: Research	2
Journal of Learning Analytics	1
Organizacija	1
Language Testing in Asia	1
Journal of Applied Learning & Teaching	1
Behavioral Sciences	1
Computers and Education: Artificial Intelligence	1
Smart Learning Environments	1

Fuente: Elaboración propia.

En particular, la ausencia de reporte proviene de dos situaciones. Primero, diseños sin uso directo del modelo (p. ej., encuestas de adopción/ética o ensayos), donde el foco es describir actitudes, marcos y riesgos (Alrayes et al., 2024; Arbulú Ballesteros et al., 2024; Chan & Hu, 2023; Firat, 2023; Jereb & Urh, 2024; Mujtaba, 2024; Ngo et al., 2024; Oyelere & Aruleba, 2025; Pikhart et al., 2024; Stojanov et al., 2024). Segundo, falta de transparencia técnica en estudios que sí usaron ChatGPT pero lo reportan como producto genérico (p. ej., uso vía interfaz web en periodos de rápida actualización o límites editoriales) (Coban et al., 2025; Liew et al., 2025; Oktarin et al., 2024; Rezai et al., 2024; Sayed et al., 2024). En contraste, sólo tres trabajos declaran explícitamente la versión: GPT-4o (Du et al., 2025; Xie et al., 2025) y gpt-3.5-turbo (Santhosh et al., 2024).

Para fortalecer la replicabilidad, se recomienda reportar de forma estandarizada: versión/modelo, fecha y modalidad de acceso (web/Plus/API), parámetros relevantes (cuando aplique), *prompts* o plantillas utilizadas y salvaguardas de verificación humana.

Tabla 2. Modelos utilizados en los estudios	
Modelo o categoría	Total
GPT-4o (OpenAI)	2
gpt-3.5-turbo (OpenAI)	1
ChatGPT (versión no especificada)	16
NA (no se usó modelo en la metodología)	38

Fuente: Elaboración propia.

3.3. Usos para personalizar el aprendizaje con ChatGPT en la universidad

En cuanto a la personalización del aprendizaje con ChatGPT, la Tabla 3 muestra que la vía más consistente es la retroalimentación individualizada y oportuna, la cual tiende a mejorar desempeño, calidad de productos y alfabetización de retroalimentación (Coban et al., 2025; Ding et al., 2025; Mahapatra, 2024; McGuire et al., 2024; Oktarin et al., 2024; Sayed et al., 2024; Xie et al., 2025). Por ejemplo, cuando la personalización se activa en tiempo real según el estado del estudiante (p. ej., compromiso estimado mediante seguimiento ocular), las intervenciones adaptativas —resúmenes breves y ayudas contextuales— elevan la comprensión y el compromiso (Santhosh et al., 2024).

Algunas otras modalidades que aportan a la experiencia y a la autorregulación incluyen: a) planes o itinerarios por metas —secuencias de actividades personalizadas que fijan objetivos, pasos intermedios y rutas alternativas según brechas diagnosticadas—, b) ajuste lingüístico del agente (registro y complejidad del mensaje) —adecuación de tono, vocabulario y longitud al nivel y a la disciplina del estudiante—, c) explicaciones personalizadas en dashboards —pistas y justificaciones “por qué y para qué” vinculadas al avance, errores y metas, visibles en paneles de progreso—, y d) *co-prompting* para mejorar la calidad de las entradas —coelaboración guiada de instrucciones y plantillas de *prompts* con el docente o con el sistema para obtener respuestas más pertinentes—.

No obstante, sus efectos presentan matices: el tono juvenil —registro informal y cercano al estudiantado— puede favorecer el compromiso, pero también aumentar la carga percibida (Liew et al., 2025; Jin et al., 2025; Perifanou & Economides, 2025). En tareas complejas, ChatGPT actúa como andamiaje cognitivo —apoyos temporales y graduados— que fortalece las HOTS (habilidades de orden superior como análisis, evaluación y creación) y apoya iteraciones de diseño; de forma complementaria, la gamificación —uso de mecánicas de juego como retos, puntos o insignias— mejora la motivación, aunque no siempre implica adaptación fina por estudiante (Du, Du, Zhou, & Bai, 2025; Beltozar-Clemente & Díaz-Vega, 2024).

Asimismo, se identificaron casos distintivos por disciplina (véase la Tabla 4). La personalización se observa con mayor claridad en laboratorios de física con Realidad Aumentada (AR) —donde la retroalimentación redirige la atención hacia lo relevante— y en cursos de inglés como lengua extranjera o segunda lengua (EFL/ESL) mediante planes de estudio y retroalimentación individual que mejoran desempeño, compromiso y variables afectivas (Coban et al., 2025; Ding et al., 2025; Mahapatra, 2024; Oktarin et al., 2024; Sayed et al., 2024).

En formas de estudio y lectura general, la intervención adaptativa basada en seguimiento ocular estima el foco atencional y dispara ayudas justo a tiempo, elevando la comprensión (Santhosh et al., 2024). En estadística y metodología, se ha comparado la retroalimentación de ChatGPT y de pares para fortalecer el juicio evaluativo y la toma de decisiones analíticas (Xie et al., 2025). Por otro lado, en negocios y experiencias de interacción humano-computadora (HCI) la personalización lingüística del chatbot (tono, formalidad y carga afectiva) modula cercanía y emoción, aunque puede aumentar la carga cognitiva; en asignaturas con analíticas de aprendizaje se personalizan explicaciones

Tabla 3. Modalidades de personalización con ChatGPT				
Modalidad	¿Qué personaliza?	Implementación típica	Observaciones	Autores
Retroalimentación individualizada	Tipo, detalle y momento de la retroalimentación según desempeño o producción	Tutor conversacional que comenta respuestas, códigos o textos del estudiante.	↑ rendimiento, ↑ calidad de productos, ↑ feedback literacy, ↑ motivación	Coban <i>et al.</i> , 2025; Ding <i>et al.</i> , 2025; Mahapatra, 2024; McGuire <i>et al.</i> , 2024; Oktarin <i>et al.</i> , 2024; Sayed <i>et al.</i> , 2024; Xie <i>et al.</i> , 2025.
Adaptación en tiempo real por estado del alumno	Intervención cuando baja el compromiso	Detección de compromiso (seguimiento ocular) y disparo de resúmenes o apoyos con ChatGPT	↑ compromiso, ↑ comprensión; intervención oportuna	Santhosh, Dengel, & Ishimaru, 2024.
Planes/itinerarios personalizados	Ritmo, metas y foco de práctica	Sugerencias de objetivos y secuencias de tareas a partir de necesidades del estudiante	↑ autorregulación, ↑ percepción de ajuste a metas	Oktarin <i>et al.</i> , 2024; Rezaei, Namaziandost, & Hwang, 2024.
Personalización lingüística del agente	Tono/persona del chatbot (p. ej., Gen-Z slang vs. estándar)	Ajustes de registro comunicativo y cercanía	↑ compromiso y valencia; posible ↑ de carga percibida	Liew <i>et al.</i> , 2025.
Explicaciones personalizadas dentro de la retroalimentación	Clarificaciones y racionales adaptados al caso del estudiante	Dashboard/análíticas con explicaciones generadas por IA para entender retroalimentación	↑ comprensión de la retroalimentación; adopción depende del ajuste a expectativas	Jin <i>et al.</i> , 2025.
Co-prompting/ andamiaje colaborativo	Apoyos diferenciados en ABP por rol/necesidad	Guías de prompting y codiseño de solicitudes a la IA	↑ eficiencia/estructura; atención a verificación y dependencia	Perifanou & Economides, 2025.
Soporte flexible a tareas complejas (HOTS)	Andamiaje cognitivo durante resolución de problemas	Conversaciones iterativas para idear, evaluar y refinar	↑ diversidad de HOTS; mejoras ligeras en originalidad/utilidad	Du, Du, Zhou, & Bai, 2025.
Gamificación con tutor conversacional	Motivadores y rutas de práctica autodirigidas	Misiones/retos donde el estudiante decide consultas al tutor	↑ rendimiento y ↑ motivación a nivel de curso (no siempre con adaptación por alumno)	Beltozar-Clemente & Díaz-Vega, 2024.

Fuente: Elaboración propia.

para apoyar la lectura de la retroalimentación; y en ABP el *co-prompting* diferencia el aprendizaje según el rol (Liew *et al.*, 2025; Jin *et al.*, 2025; Perifanou & Economides, 2025). Finalmente, en diseño y problemas complejos, ChatGPT puede activar habilidades cognitivas de orden superior (HOTS) y guiar la iteración; en física introductoria, la gamificación con tutor conversacional mejora el rendimiento y la motivación (Du *et al.*, 2025; Beltozar-Clemente & Díaz-Vega, 2024).

Tabla 4. Casos de uso por disciplina			
Disciplina o Curso	Tarea	¿Cómo se usa ChatGPT?	Autores
Física (lab AR)	Resolver/practicar en laboratorio de criptografía	Retroalimentación individual que redirige la atención hacia elementos clave	Coban et al., 2025.
EFL – Escritura	Borradores y revisiones	Planes personalizados y retroalimentación textual	Oktarin et al., 2024; Mahapatra, 2024; Ding et al., 2025.
EFL – Speaking	Preparación y práctica oral	Retroalimentación inmediato e individual sobre desempeño	Sayed et al., 2024.
Lectura/estudio general	Mantener atención y comprensión	Resúmenes activados ante bajo compromiso (seguimiento ocular)	Santhosh et al., 2024.
Negocios / HCI	Interacciones con chatbot	Ajuste del tono/persona (slang vs. estándar)	Liew et al., 2025.
Cursos con analíticas	Comprender la retroalimentación recibida	Explicaciones personalizadas generadas por IA en un dashboard	Jin et al., 2025.
ABP (posgrado)	Resolver proyectos en equipo	Co-prompting y andamiaje de solicitudes según rol	Perifanou & Economides, 2025.
Diseño (problemas complejos)	Ideación y evaluación de soluciones	Andamiaje para activar HOTS (explicar, evaluar, rediseñar)	Du et al., 2025.
Física I (curso)	Estudio con retos gamificados	Tutor conversacional + gamificación (elección flexible del apoyo)	Beltozar-Clemente & Díaz-Vega, 2024.
Estadística / Metodología	Doctorado – Administración Educativa	Caso: retroalimentación de ChatGPT-4o vs. pares para fortalecer juicio evaluativo en análisis estadístico	Xie et al., 2025.

Fuente: Elaboración propia.

Además, como se observa en la Tabla 5, para personalizar de forma efectiva con ChatGPT se identificaron seis componentes. Primero, retroalimentación criterial y explicativa en el momento correcto: comentarios alineados a criterios o rúbricas que explican el porqué y el cómo mejorar, entregados a tiempo para que puedan usarse (Coban et al., 2025; Ding et al., 2025; Mahapatra, 2024; McGuire et al., 2024; Oktarin et al., 2024; Sayed et al., 2024; Xie et al., 2025). Segundo, ayudas de contenido activadas por datos del alumno —pistas justo a tiempo cuando disminuye el compromiso o se detectan errores recurrentes— (Santhosh et al., 2024). Tercero, secuencias y ritmos adaptados a metas y nivel —planes de avance con objetivos intermedios y dosificación

de tareas— (Oktarin et al., 2024; Rezai et al., 2024). Cuarto, tono y persona del agente —rol (tutor, mentor, revisor) y ajuste de registro, cercanía y estilo según el objetivo— (Liew et al., 2025). Quinto, andamiaje de habilidades cognitivas de orden superior (HOTS) mediante preguntas guía, ejemplos modelo o rúbricas parciales (Du et al., 2025). Sexto, una capa motivacional con misiones, retos y recompensas (gamificación) que visibiliza el progreso (Beltozar-Clemente & Díaz-Vega, 2024).

Lo anterior sugiere que estos componentes pueden explicar las mejoras reportadas en rendimiento, comprensión, autorregulación y compromiso; sin embargo, también advierten compensaciones (p. ej., mayor carga percibida cuando se exagera la cercanía del agente). Además, los perfiles de uso y dependencia son heterogéneos, por lo que conviene promover autorregulación y criterios de uso explícitos (Stojanov et al., 2024).

Tabla 5. Adopción, aceptación, uso y ética		
Componente	Prácticas con ChatGPT	Autores
Retroalimentación	Comentarios criterios y explicativos ajustados a la producción del estudiante; momento oportuno	Coban et al., 2025; Ding et al., 2025; Mahapatra, 2024; McGuire et al., 2024; Oktarin et al., 2024; Sayed et al., 2024; Xie et al., 2025.
Contenido y ayudas	Resúmenes cuando cae el compromiso; explicaciones contextualizadas	Santhosh et al., 2024; Jin et al., 2025.
Secuencia y ritmo	Planes y metas individualizadas para avanzar según necesidad	Oktarin et al., 2024; Rezai et al., 2024.
Tono y persona	Ajuste del registro (cercanía/juventud vs. estándar) según perfil	Liew et al., 2025.
HOTS y estrategias	Andamiaje para idear, evaluar, comparar y refinar	Du et al., 2025.
Motivación y compromiso	Misiones, retos y recompensas con tutor conversacional	Beltozar-Clemente & Díaz-Vega, 2024.
Perfiles de uso y dependencia	Identifica patrones de relianza (baja/alta) y delegación de tareas; insumo para autorregulación y políticas de uso	Stojanov et al., 2024.

Fuente: Elaboración propia.

3.4. Recomendaciones para personalizar el aprendizaje con ChatGPT en la universidad

En conjunto, como se observa y detalla en la Tabla 6, los estudios convergen en una personalización del aprendizaje con gobernanza: usar ChatGPT como apoyo y nunca como sustituto del juicio docente, acompañada de supervisión humana y lineamientos institucionales claros (Alrayes et al., 2024; Gencer & Gencer, 2024; Jereb & Urh, 2024; Riedel et al., 2023; Wei et al., 2025). Este marco se sostiene en tres pilares: (a) políticas de transparencia, integridad y privacidad, con mitigación activa de sesgos y alucinaciones; (b) alfabetización en IAGen para docentes y estudiantes (evaluación crítica de salidas, límites del modelo, buenas prácticas de *prompting*); y (c) infraestructura y acceso equitativo para reducir brechas y asegurar la usabilidad de las herramientas (Chan & Hu, 2023; Jin et al., 2025; Liew et al., 2025; Morell-Mengual et al., 2025; Sallam & Al-Salahat, 2023; Sayed et al., 2024).

En la implementación didáctica, la personalización se operacionaliza mediante *prompts* guiados por el docente y retroalimentación justo a tiempo ajustada al nivel, ritmo y estilo lingüístico del estudiante; se enfatiza diseñar tareas que activen HOTS/metacognición y que hagan visible el razonamiento del modelo para sostener la autonomía y evitar sobrerreliancia (McGuire et al., 2024; Pikhart et al., 2024). Paralelamente, se pide rediseñar la evaluación hacia desempeños auténticos y, cuando se usen pruebas objetivas, fortalecer su robustez frente a la IA (p. ej., revisión de MCQ), además de explorar ChatGPT como coevaluador con criterios y supervisión (Beltozar-Clemente & Díaz-Vega, 2024; Coban et al., 2025; Jereb & Urh, 2024; Riedel et al., 2023). Para experiencias más ricas, varios trabajos proponen adaptatividad, multimodalidad (VR/AR, simulaciones) y gamificación cuidadosamente curada, manteniendo registros de versiones y parámetros del LLM para trazabilidad y mejora continua (Du et al., 2025; Han et al., 2025; Santhosh et al., 2024).

Además, recomiendan analítica de aprendizaje y reportes estandarizados (documentar versión, temperatura, protocolos éticos), así como investigación rigurosa y replicable con diseños longitudinales y experimentales, muestras diversas y comparadores entre modelos y contextos disciplinares (Boudia & Bengueddach, 2024; French et al., 2023; Liew et al., 2025; Ngo et al., 2024; Zekaj, 2023). También subrayan hojas de ruta institucionales para integrar LLM en el currículo, con atención a la soste-

Tabla 6. Recomendaciones para la personalización del aprendizaje con ChatGPT	
Tipo de recomendación Descripción	Autores
Uso complementario con supervisión humana (no sustituto) ChatGPT como apoyo/tutor, nunca reemplazo del juicio docente; evitar sobredependencia.	Alrayes et al., 2024; Beltozar-Clemente & Díaz-Vega, 2024; Boudia & Bengueddach, 2024; Gencer & Gencer, 2024; Iskender, 2023; Jereb & Urh, 2024; Perifanou & Economides, 2025; Rezaei et al., 2024; Riedel et al., 2023; Wei et al., 2025; Intorsureanu et al., 2025.
Políticas, lineamientos y divulgación institucional Reglas claras de uso, transparencia/registro del uso de IAGen en cursos y evaluación.	Chan & Hu, 2023; Jereb & Urh, 2024; Liew et al., 2025; Morell-Mengual et al., 2025; Ngo et al., 2024; Perifanou & Economides, 2025; Rezaei et al., 2024; Riedel et al., 2023; Sayed et al., 2024; Wei et al., 2025)
Ética, privacidad, sesgos y seguridad Salvaguardas de datos; mitigación de sesgos y alucinaciones; uso responsable.	(Alrayes et al., 2024; Buele et al., 2025; Chan & Hu, 2023; Gencer & Gencer, 2024; Jin et al., 2025; Liew et al., 2025; Morell-Mengual et al., 2025; Rezaei et al., 2024; Sayed et al., 2024; Wei et al., 2025; Intorsureanu et al., 2025)
Alfabetización en IA y formación docente-estudiantil Capacitar en límites, sesgos, evaluación crítica y buenas prácticas de prompting.	(Alrayes et al., 2024; Arbulú Ballesteros et al., 2024; Chan & Hu, 2023; Jereb & Urh, 2024; Jin et al., 2025; Liew et al., 2025; Oktarin et al., 2024; Perifanou & Economides, 2025; Rezaei et al., 2024; Sallam & Al-Salahat, 2023; Sayed et al., 2024; Wei et al., 2025; Stojanov et al., 2024)
Diseño instruccional adaptativo y con HOTS Tareas que demanden razonamiento, metacognición y reflexión sobre el output de IA.	(Beltozar-Clemente & Díaz-Vega, 2024; Morell-Mengual et al., 2025; Pikhart et al., 2024; Riedel et al., 2023; Wei et al., 2025)
Retroalimentación personalizada y prompts guiados Retroalimentación “just-in-time”, rúbricas claras; diseño de prompts granulares y guiados por el docente.	(Boudia & Bengueddach, 2024; Coban et al., 2025; Han et al., 2025; Iskender, 2023; Jereb & Urh, 2024; McGuire et al., 2024; Perifanou & Economides, 2025; Ding et al., 2025; Mahapatra, 2024; Xie et al., 2025)
Personalización/adaptatividad de la experiencia Rutas personalizadas, ajuste de estilo lingüístico y nivel de detalle; control del usuario.	(Du et al., 2025; French et al., 2023; Gencer & Gencer, 2024; Han et al., 2025; Husain, 2024; Phua et al., 2025; Santhosh et al., 2024; Wei et al., 2025)
Transparencia y explicabilidad del modelo Aumentar interpretabilidad para docentes/estudiantes; rendición de cuentas.	(Husain, 2024; Jin et al., 2025; Liew et al., 2025; Morell-Mengual et al., 2025; Rezaei et al., 2024)
Infraestructura, acceso equitativo y brecha digital Conectividad, ancho de banda, acceso a modelos potentes; entornos GenAI-friendly.	(Chan & Hu, 2023; Gencer & Gencer, 2024; Morell-Mengual et al., 2025; Oktarin et al., 2024; Phua et al., 2025; Sayed et al., 2024)
Rediseño de evaluación (auténtica, MCQ, coevaluación) Evaluaciones centradas en desempeño/razonamiento; MCQ robustas; ChatGPT como coevaluador con criterios.	(Beltozar-Clemente & Díaz-Vega, 2024; Coban et al., 2025; Jereb & Urh, 2024; Perifanou & Economides, 2025; Riedel et al., 2023; Sayed et al., 2024)
Análítica de aprendizaje y trazabilidad de LLM Medir impacto (objetivo y subjetivo); documentar versión, temperatura y parámetros.	(Boudia & Bengueddach, 2024; Du et al., 2025; Han et al., 2025; Rezaei et al., 2024)
Multimodalidad y XR (VR/AR, laboratorios) Integrar LLM en AR/VR y simulaciones; intervenciones multimodales curadas.	(Coban et al., 2025; Du et al., 2025; Phua et al., 2025; Santhosh et al., 2024)
Gamificación y compromiso Combinar ChatGPT con mecánicas de juego y verificación formativa para motivar.	(Boudia & Bengueddach, 2024; French et al., 2023; Han et al., 2025; Santhosh et al., 2024; Sayed et al., 2024)
Roadmap institucional e integración curricular Plan estratégico, áreas prioritarias, integración transversal en el currículo.	Morell-Mengual et al., 2025; Riedel et al., 2023; Wei et al., 2025.
Investigación rigurosa y reporte Diseños longitudinales/experimentales, muestras diversas, comparadores; estandarizar métricas y reportes.	Boudia & Bengueddach, 2024; Buele et al., 2025; Coban et al., 2025; French et al., 2023; Iskender, 2023; Liew et al., 2025; Morell-Mengual et al., 2025; Ngo et al., 2024; Sallam & Al-Salahat, 2023; Santhosh et al., 2024; Sayed et al., 2024; Zekaj, 2023; Ding et al., 2025; Mahapatra, 2024; Stojanov et al., 2024; Xie et al., 2025.
Comunidades y colaboración Espacios compartidos de prompting, revisión por pares y comunidades de práctica.	Boudia & Bengueddach, 2024; Chan & Hu, 2023; Mujtaba, 2024; Pikhart et al., 2024; Sayed et al., 2024.
Sostenibilidad (costos/energía) Adopción equilibrada con mirada a costos, eficiencia y huella energética.	Chan & Hu, 2023; Liew et al., 2025; Perifanou & Economides, 2025; Zekaj, 2023.
Credenciales y reconocimiento de logros con IA Explorar acreditación/certificación de productos/competencias apoyados por IA.	Ngo et al., 2024.
Usabilidad/UI y legibilidad Mejorar interfaz, legibilidad y modos offline para accesibilidad y satisfacción.	Buele et al., 2025; Phua et al., 2025; Sayed et al., 2024.

Source: Own elaboration, based on cited authors.

nibilidad (costos/energía) y a la mejora de usabilidad/UI y a la legibilidad —todo ello para maximizar los beneficios de la personalización sin comprometer pensamiento crítico, creatividad ni equidad (Morell-Mengual et al., 2025; Buele et al., 2025; Phua et al., 2025; Wei et al., 2025).

4. DISCUSIÓN Y CONCLUSIONES

El corpus ($n = 38$) confirma un crecimiento acelerado de publicaciones desde 2023, con un pico en 2024 y una amplia dispersión editorial. Esto sugiere un campo interdisciplinario —tecnología educativa, medicina, ingeniería, psicología y ciencias sociales— sin un foro único que concentre la evidencia.

En respuesta a R2, predomina una combinación de diseños. Por un lado, intervenciones y cuasi-experimentos que miden efectos en aprendizaje o procesos (retroalimentación individualizada, adaptatividad basada en seguimiento ocular, andamiaje de habilidades de orden superior, ajuste del tono del chatbot y evaluación automatizada de escritura) (Coban et al., 2025; Ding et al., 2025; Du et al., 2025; Liew et al., 2025; Mahapatra, 2024; Oktarin et al., 2024; Santhosh et al., 2024; Sayed et al., 2024; Xie et al., 2025). Por otro, encuestas sobre adopción, alfabetización y ética que contextualizan condiciones de uso (Alrayes et al., 2024; Chan & Hu, 2023; Han et al., 2025; Jereb & Urh, 2024; Morell-Mengual et al., 2025; Ngo et al., 2024; Oyelere & Aruleba, 2025; Phua et al., 2025; Arbulú Ballesteros et al., 2024; Stojanov et al., 2024).

En paralelo, persiste un déficit de transparencia técnica: la versión del modelo y su modalidad de acceso raramente se reportan, lo que dificulta la replicabilidad, comparación entre hallazgos y evaluación de sesgos. Esta revisión refuerza, por tanto, la necesidad de reportes estandarizados (versión, acceso, *prompts* y salvaguardas).

En conjunto, los hallazgos empíricos son coherentes con la teoría clásica que sustenta la personalización como adaptación guiada y con la evidencia acumulada sobre el impacto de la retroalimentación: ajustar tareas, apoyos y tiempos constituye un mecanismo plausible para mejorar resultados y equidad (Bloom, 1984; Dewey, 1916; Hattie, 2008; Vygotsky, 1978).

Respecto a R1, se confirma un déficit de transparencia técnica: sólo tres estudios reportan explícitamente la versión del modelo empleado (Du et al., 2025; Santhosh et al., 2024; Xie et al., 2025). Este hallazgo refuerza la necesidad de reportes estandarizados, particularmente en un campo donde la actualización del modelo puede modificar tipos de respuesta, sesgos y resultados.

En términos de efectividad, la personalización más consistente se logra mediante retroalimentación individualizada y adaptaciones en el momento oportuno. En laboratorios de física con realidad aumentada, la retroalimentación mediada por ChatGPT puede redirigir la atención de forma visual hacia elementos clave y mejorar el rendimiento (Coban et al., 2025); en lectura y estudio general, la detección de compromiso mediante seguimiento ocular activa ayudas y resúmenes personalizados que elevan la comprensión y el compromiso (Santhosh et al., 2024).

De manera similar, en escritura académica (EFL/ESL) se reportan mejoras cuando ChatGPT se utiliza como herramienta de evaluación y retroalimentación formativa (Ding et al., 2025; Mahapatra, 2024; Oktarin et al., 2024). En estadística y metodología, la comparación entre retroalimentación de ChatGPT y de pares sugiere efectos en el juicio evaluativo y la calidad del análisis (Xie et al., 2025). Finalmente, en tareas de diseño, ChatGPT opera como andamiaje de habilidades de orden superior (HOTS) que amplía la diversidad de operaciones cognitivas y, en menor medida, la originalidad y utilidad de los productos (Du et al., 2025; Sayed et al., 2024).

Además de la retroalimentación y la adaptatividad, se identifican modalidades parciales que afinan la experiencia más que el contenido: ajuste del tono y de la persona del chatbot, personalización lingüística y emocional, co-diseño de *prompts* (*co-prompting*) y gamificación. Estas estrategias suelen mejorar el nivel de motivación y percepción de utilidad, aunque la evidencia es aún heterogénea y depende del diseño instruccional y del contexto (Beltozar-Clemente & Díaz-Vega, 2024; Jin et al., 2025; Liew et al., 2025; Perifanou & Economides, 2025).

Además, los estudios convergen en una personalización con gobernanza: uso complementario de ChatGPT —no sustitutivo del juicio docente— con supervisión humana, políticas de transparencia e integridad, alfabetización en IAGen y condiciones de acceso equitativo (Du et al., 2025; McGuire et al., 2024; Stojanov et al., 2024). En la práctica didáctica, se traduce en retroalimentación criterial, explicativa y oportuna; rutas con metas y secuencias ajustadas; tareas que activen habilidades de orden superior y hagan visible la lógica de la salida para evitar sobrerrelianza; ajuste lingüístico con control de carga; y analítica y trazabilidad de versiones y parámetros del modelo, junto con evaluación y rediseño hacia desempeños auténticos (Ding et al., 2025; Mahapatra, 2024; Sayed et al., 2024; Xie et al., 2025).

Es decir, para llevar el uso de este tipo de LLM a la práctica (R3), las recomendaciones priorizan: (1) diseñar actividades *feedback-first* donde ChatGPT entregue retroalimentación alineada a criterios y el estudiante deba revisarla, justificarla y aplicar cambios (Ding et al., 2025; Mahapatra, 2024; McGuire et al., 2024; Xie et al., 2025); (2) combinar la salida del modelo con verificación humana y evaluación auténtica para evitar dependencia y favorecer el aprendizaje (Du et al., 2025; Sayed et al., 2024); y (3) institucionalizar políticas de integridad, transparencia y capacitación, junto con soporte tecnológico equitativo y monitoreo de perfiles de uso (Stojanov et al., 2024). Finalmente, se propone reorientar la interacción hacia un cuestionamiento continuo (mayéutica socrática), de modo que el estudiante actúe como supervisor constante de la salida del modelo —verifica, contrasta y justifica— guiado por el docente como experto, y orientado a niveles de análisis, evaluación y creación en la taxonomía de Bloom (Bloom et al., 1956; Yang et al., 2005).

Y en respuesta a nuestra pregunta principal —si la IAGen “nos hará estúpidos”— esta se puede responder como advertencia: cuando el estudiante delega sin contraste la producción y la validación, se incentiva la descarga cognitiva y puede acumularse deuda cognitiva (Kosmyrna et al., 2025; Gerlich, 2025; Risko & Gilbert, 2016). No obstante, la evidencia revisada sugiere que, con diseños “*feedback-first*”, tareas de orden superior y supervisión humana, ChatGPT puede convertirse en un socio de pensamiento que amplía modos de procesar información; incluso, con la emergencia de agentes, el aprendizaje podría articularse como cognición distribuida, siempre que el estudiante orqueste, verifique y decida, guiado por el docente como supervisor experto. En ese tránsito, recuperar el cuestionamiento socrático —preguntar antes de aceptar, contrastar antes de concluir— es crucial para sostener agencia, metacognición y pensamiento crítico (Bloom et al., 1956; Stadler et al., 2024; Yang et al., 2005).

En cuanto a limitaciones, aunque se siguió PRISMA 2020, el protocolo no se registró y la búsqueda se restringió a tres bases (Web of Science, Scopus y ERIC), complementada con encadenamiento de citas; por ello, es posible que existan estudios no recuperados. Además, la heterogeneidad de diseños y métricas limita comparaciones directas y la posibilidad de meta-análisis. Como agenda inmediata, se requieren diseños longitudinales y experimentales, así como reportes estandarizados de versión y modelo, acceso, parámetros y *prompts*, para sostener replicabilidad. Asimismo, resulta prioritario evaluar impactos en equidad, sesgos y sostenibilidad en contextos diversos —en particular en América Latina— para establecer recomendaciones transferibles.

REFERENCIAS

- Alrayes, A., Henari, T. F., & Ahmed, D. A. (2024). ChatGPT in education: Understanding the Bahraini academics' perspective. *The Electronic Journal of e-Learning*, 22(2), 112–134. <https://doi.org/q7tn>
- Arbulú Ballesteros, M. A., Acosta Enríquez, B. G., Ramos Farroñán, E. V., García Juárez, H. D., Cruz Salinas, L. E., Blas Sánchez, J. E., Arbulú Castillo, J. C., Licapa-Redolfo, G. S., & Farfán Chilicaus, G. C. (2024). The sustainable integration of AI in higher education: Analyzing ChatGPT acceptance factors through an extended UTAUT2 framework in Peruvian universities. *Sustainability*, 16(23), 10707. <https://doi.org/q7tp>
- Beltazar-Clemente, S., & Díaz-Vega, E. (2024). Physics XP: Integration of ChatGPT and gamification to improve academic performance and motivation in Physics 1 course. *International Journal of Engineering Pedagogy (iJEP)*, 14(6), 82–92. <https://doi.org/q7tq>
- Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive domain. David McKay.
- Bloom, B. S. (1984). The 2-sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher*, 13(6), 4–16. <https://doi.org/ddj5p7>
- Boudia, C., & Bengueddach, A. (2024). Innovating higher education: The rise of ChatGPT in collaborative learning – A short review. In M. Shelley & O. T. Ozturk (Eds.), *Proceedings of ICRES 2024—International Conference on Research in Education and Science* (pp. 1347–1370). ISTES.
- Buele, J., Sabando-García, Á. R., Sabando-García, B. J., & Yáñez-Rueda, H. (2025). Ethical use of generative artificial intelligence among Ecuadorian university students. *Sustainability*, 17(10), 4435. <https://doi.org/q7tr>
- Chan, C. K. 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/gshsfg>
- Coban, A., Dzotjan, D., Küchemann, S., Durst, J., Kuhn, J., & Hoyer, C. (2025). AI support meets AR visualization for Alice and Bob: Personalized learning based on individual ChatGPT feedback in an AR quantum cryptography experiment for physics lab courses. *EPJ Quantum Technology*, 12, 15. <https://doi.org/q7ts>

- Dewey, J. (1916). *Democracy and education: An introduction to the philosophy of education*. Macmillan.
- Ding, L., Zou, D., & Kohnke, L. (2025). ChatGPT as an automated writing evaluation tool: How students perceive it and how it affects their writing. *Education and Information Technologies*, 30, 26777–26799. <https://doi.org/q7tt>
- Du, X., Du, M., Zhou, Z., & Bai, Y. (2025). Facilitator or hindrance? The impact of AI on university students' higher-order thinking skills in complex problem solving. *International Journal of Educational Technology in Higher Education*, 22, 39. <https://doi.org/hbjd8c>
- Firat, M. (2023). What ChatGPT means for universities: Perceptions of scholars and students. *Journal of Applied Learning & Teaching*, 6(1), 57–63. <https://doi.org/g82g8p>
- French, F., Levi, D., Maczo, C., Simonaityte, A., Triantafyllidis, S., & Varda, G. (2023). Creative use of OpenAI in education: Case studies from game development. *Multimodal Technologies and Interaction*, 7(8), 81. <https://doi.org/gsqbns>
- 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, 1166682. <https://doi.org/gtjfh3>
- Gencer, G., & Gencer, K. (2024). A comparative analysis of ChatGPT and medical faculty graduates in medical specialization exams: Uncovering the potential of artificial intelligence in medical education. *Cureus*, 16(8), e66517. <https://doi.org/q7tv>
- Gerlich, M. (2025). AI tools in society: Impacts on cognitive offloading and the future of critical thinking. *Societies*, 15(1), 6. <https://doi.org/g8znms>
- Han, S., Kang, H. S., Gimber, P., & Lim, S. (2025). Nursing students' perceptions and use of generative artificial intelligence in nursing education. *Nursing Reports*, 15(2), 68. <https://doi.org/q7tw>
- Hattie, J. (2008). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. Routledge.
- Husain, A. J. A. (2024). Potentials of ChatGPT in computer programming: Insights from programming instructors. *Journal of Information Technology Education: Research*, 23, Article 2. <https://doi.org/q7tx>

- Imran, M., & Almusharraf, N. (2023). Analyzing the role of ChatGPT as a writing assistant at higher education level: A systematic review of the literature. *Contemporary Educational Technology*, 15(4), ep464. <https://doi.org/gtc4m4>
- Întorsureanu, I., Oprea, S.-V., Bâra, A., & Vespan, D. (2025). Generative AI in education: Perspectives through an academic lens. *Electronics*, 14(5), 1053. <https://doi.org/q7tz>
- Iskender, A. (2023). Holy or unholy? Interview with OpenAI's ChatGPT. *European Journal of Tourism Research*, 34, 3414. <https://doi.org/10.54055/ejtr.v34i.3169>
- Jereb, E., & Urh, M. (2024). The use of artificial intelligence among students in higher education. *Organizacija*, 57(4), 333–345. <https://doi.org/10.2478/orga-2024-0024>
- Jin, F. J.-Y., Maheshi, B., Lai, W., Li, Y., Gasevic, D., Chen, G., Charwat, N., Chan, P. W. K., Martinez-Maldonado, R., & Tsai, Y.-S. (2025). Students' perceptions of generative AI-powered learning analytics in the feedback process: A feedback literacy perspective. *Journal of Learning Analytics*, 12(1), 152–168. <https://doi.org/q7t2>
- Kosmyna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X.-H., Beresnitzky, A. V., Braunstein, I., & Maes, P. (2025). Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing tasks (arXiv:2506.08872v2) [Preprint]. arXiv. <https://linkly.link/2jcvH>
- Labadze, L., Grigolia, M., & Machaidze, L. (2023). Role of AI chatbots in education: Systematic literature review. *International Journal of Educational Technology in Higher Education*, 20, 56. <https://doi.org/gtqrkm>
- Liew, T. W., Tan, S.-M., Pang, W. M., Gan, C. L., Chan, T. J., & Ahmad, F. (2025). Cringe, lit, or mid: Affective and cognitive effects of youth slang in an educational chatbot. *Acta Psychologica*, 256, 105036. <https://doi.org/hbgrs4>
- Mahapatra, S. (2024). Impact of ChatGPT on ESL students' academic writing skills: A mixed-methods intervention study. *Smart Learning Environments*, 11, 9. <https://doi.org/g9xwd4>
- McGuire, A., Qureshi, W., & Saad, M. (2024). A constructivist model for leveraging GenAI tools for individualized, peer-simulated feedback on student writing. *International Journal of Technology in Education*, 7(2), 326–352. <https://doi.org/q7t3>
- Meyer, A., Riese, J., & Streichert, T. (2024). Comparison of the performance of

- GPT-3.5 and GPT-4 with that of medical students on the written German medical licensing examination: Observational study. *JMIR Medical Education*, 10, e50965. <https://doi.org/q7t4>
- Morell-Mengual, V., Fernández-García, O., Berenguer, C., Ortega-Barón, J., Gil-Llario, M. D., & Estruch-García, V. (2025). *Characteristics, motivations and attitudes of students using ChatGPT and other language model-based chatbots in higher education*. Education and Information Technologies. Advance online publication. <https://doi.org/g95bnw>
- Mujtaba, B. (2024). Clarifying ethical dilemmas in sharpening students' artificial intelligence proficiency: Dispelling myths about using AI tools in higher education. *Business Ethics and Leadership*, 8(2), 107–127. <https://doi.org/q7z7>
- Naznin, K., Al Mahmud, A., Nguyen, M. T., & Chua, C. (2025). ChatGPT integration in higher education for personalized learning, academic writing, and coding tasks: A systematic review. *Computers*, 14, 53. <https://doi.org/q7t5>
- Ngo, T. T. A., An, G. K., Nguyen, P. T., & Tran, T. T. (2024). Unlocking educational potential: Exploring students' satisfaction and sustainable compromise with ChatGPT using the ECM model. *Journal of Information Technology Education: Research*, 23, Article 21. <https://doi.org/10.28945/5344>
- Oktarin, I. B., Saputri, M. E. E., Magdalena, B., Hastomo, T., & Maximilian, A. (2024). Leveraging ChatGPT to enhance students' writing skills, comprehension, and feedback literacy. *Edelweiss Applied Science and Technology*, 8(4), 2306–2319. <https://doi.org/q7t6>
- OpenAI. (2022, November 30). Introducing ChatGPT. <https://openai.com/blog/chatgpt/>
- OpenAI. (2023). GPT-4 technical report. <https://cdn.openai.com/papers/gpt-4.pdf>
- Oyelere, S. S., & Aruleba, K. (2025). A comparative study of student perceptions on generative AI in programming education across Sub-Saharan Africa. *Computers and Education Open*, 8, 100245. <https://doi.org/g8828j>
- Perifanou, M., & Economides, A. A. (2025). Collaborative uses of GenAI tools in project-based learning. *Education Sciences*, 15(3), 354. <https://doi.org/g9trzt>

- Phua, J. T. K., Neo, H.-F., & Teo, C.-C. (2025). Evaluating the impact of artificial intelligence tools on enhancing student academic performance: Efficacy amidst security and privacy concerns. *Big Data and Cognitive Computing*, 9(5), 131. <https://doi.org/q7t7>
- Pikhart, M., Klimova, B., & Al-Obaydi, L. H. (2024). Exploring university students' preferences and satisfaction in utilizing digital tools for foreign language learning. *Frontiers in Education*, 9, 1412377. <https://doi.org/q7t8>
- Rezai, A., Namaziandost, E., & Hwang, G.-J. (2024). How can ChatGPT open promising avenues for L2 development? A phenomenological study involving EFL university students in Iran. *Computers in Human Behavior Reports*, 16, 100510. <https://doi.org/q7t9>
- Riedel, M., Kaefinger, K., Stuehrenberg, A., Ritter, V., Amann, N., Graf, A., Recker, F., Klein, E., Kiechle, M., Riedel, F., & Meyer, B. (2023). ChatGPT's performance in German OB/GYN exams: Paving the way for AI-enhanced medical education and clinical practice. *Frontiers in Medicine*, 10, 1296615. <https://doi.org/q7vb>
- Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences*, 20(9), 676–688. <https://linkly.link/2ja1m>
- Sallam, M., & Al-Salahat, K. (2023). Below-average ChatGPT performance in medical microbiology exam compared to university students. *Frontiers in Education*, 8, 1333415. <https://doi.org/q7vc>
- Santhosh, J., Dengel, A., & Ishimaru, S. (2024). Gaze-driven adaptive learning system with ChatGPT-generated summaries. *IEEE Access*, 12, 173715–173733. <https://doi.org/q7vd>
- Sayed, B. T., Bani Younes, Z. B., Alkhayyat, A., Adhamova, I., & Teferi, H. (2024). To be with artificial intelligence in oral test or not to be: A probe into the traces of success in speaking skill, psychological well-being, autonomy, and academic buoyancy. *Language Testing in Asia*, 14, 49. <https://doi.org/g9gzhq>
- Stadler, M., Bannert, M., & Sailer, M. (2024). Cognitive ease at a cost: LLMs reduce mental effort but compromise depth in student scientific inquiry. *Computers in Human Behavior*, 160, 108386. <https://doi.org/hb3bx2>
- Stojanov, A., Liu, Q., & Koh, J. H. L. (2024). University students' self-reported reliance on ChatGPT for learning: A latent profile analysis. *Computers and Education: Artificial Intelligence*, 6, 100243. <https://doi.org/g855xj>

- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds.). Harvard University Press.
- Wei, B., Yao, L., Hu, X., Hu, Y., Rao, J., Ji, Y., Dong, Z., Duan, Y., & Wu, X. (2025). Evaluating the effectiveness of large language models in providing patient education for Chinese patients with ocular myasthenia gravis: Mixed-methods study. *Journal of Medical Internet Research*, 27, e67883. <https://doi.org/q7vf>
- Xie, X., Zhang, L. J., & Wilson, A. J. (2025). Comparing ChatGPT feedback and peer feedback in shaping students' evaluative judgement of statistical analysis: A case study. *Behavioral Sciences*, 15(7), 884. <https://doi.org/q7vg>
- Yang, Y.-T. C., Newby, T. J., & Bill, R. L. (2005). Using Socratic questioning to promote critical thinking skills through asynchronous discussion forums in distance learning environments. *American Journal of Distance Education*, 19(3), 163–181. <https://doi.org/c8mtbn>
- Zekaj, R. (2023). AI language models as educational allies: Enhancing instructional support in higher education. *International Journal of Learning, Teaching and Educational Research*, 22(8), 120–134. <https://doi.org/q7vh>