

TRANSFORMACIÓN

DEL AULA UNIVERSITARIA. APLICACIONES Y PERSPECTIVAS DE LA INTELIGENCIA ARTIFICIAL GENERATIVA EN LA DOCENCIA

TRANSFORMATION OF THE UNIVERSITY CLASSROOM. APPLICATIONS AND PERSPECTIVES OF GENERATIVE ARTIFICIAL INTELLIGENCE IN TEACHING

Patricio Santillán-Aguirre^{1*}

E-mail: juan.santillan@epoch.edu.ec

ORCID: <https://orcid.org/0000-0002-8610-6724>

Ramiro Santos-Poveda¹

E-mail: rsantos@epoch.edu.ec

ORCID: <https://orcid.org/0000-0002-2270-1735>

Verónica Cuadrado-Solis¹

E-mail: veronica.cuadrado@epoch.edu.ec

ORCID: <https://orcid.org/0000-0001-9756-0810>

Edgar Jaramillo-Moyano¹

E-mail: edgar.jaramillo@epoch.edu.ec

ORCID: <https://orcid.org/0000-0001-6376-1710>

¹Escuela Superior Politécnica de Chimborazo, Ecuador.

*Autor para la correspondencia

Cita sugerida (APA, séptima edición)

Santillán-Aguirre, P., Santos-Poveda, R., Cuadrado-Solis, V., & Jaramillo-Moyano, E. (2025). Transformación del aula universitaria. Aplicaciones y perspectivas de la inteligencia artificial generativa en la docencia. *Universidad y Sociedad*, 17(5). e5225.

RESUMEN

El presente estudio examinó la integración de inteligencia artificial generativa en la docencia universitaria de una institución pública andina. Se describió el motivo de investigar estrategias innovadoras que potencian la participación, la autonomía y la calidad de los productos académicos, ante la necesidad de renovar los modelos educativos tradicionales. El objetivo consistió en explicar los cambios experimentados por estudiantes y docentes cuando usan herramientas generativas de lenguaje y de creación visual se incorporaron como asistentes para planificar clases, resolver dudas, retroalimentar escritos y producir recursos multimodales. Se empleó un diseño mixto convergente, no experimental, que reunió registros de plataforma, calificaciones, artefactos de aprendizaje, diarios reflexivos y entrevistas semiestructuradas. Los datos cuantitativos se analizaron mediante pruebas paramétricas y modelos de efectos mixtos, mientras el componente cualitativo siguió un enfoque temático reflexivo para triangulación. Los resultados mostraron una mejora sostenida del rendimiento, un aumento del tiempo dedicado a actividades formativas y una reducción de la deserción temprana. Asimismo, emergieron tres temas centrales: autonomía guiada, curiosidad expandida y reorganización de la retroalimentación docente. Se concluyó que la inteligencia artificial generativa actuó como catalizador de prácticas pedagógicas más dinámicas, favoreció la reflexión crítica y simplificó la gestión del curso, aunque planteó desafíos éticos relacionados con la originalidad y la huella cultural de los contenidos. Se recomienda ampliar estudios longitudinales, desarrollar políticas institucionales sobre ingeniería de indicaciones y explorar modelos multilingües que incorporen lenguas originarias para garantizar inclusión y sostenibilidad. También se sugirió implementar mecanismos de alfabetización digital crítica que preserven la agencia estudiantil y docente.

Palabras clave: Inteligencia Artificial, Tecnología Educativa, Educación Superior, Motivación del Aprendizaje, Métodos de Enseñanza, Rendimiento Académico.

ABSTRACT

This study examined the integration of generative artificial intelligence into university teaching at a public Andean institution. The purpose was to explore innovative strategies that enhance engagement, autonomy, and the quality of academic outputs in response to the need to refresh conventional educational models. The main objective was to explain the shifts experienced by students and lecturers when language and visual generative tools were introduced as assistants for lesson planning, doubt resolution, feedback provision, and multimodal resource creation. A convergent mixed-methods, non-experimental design gathered platform logs, course grades, learning artefacts, reflective journals, and semi-structured interviews. Quantitative data were treated with parametric tests and mixed-effects models, whereas the qualitative strand followed a reflexive thematic approach to enable triangulation. The findings revealed sustained improvement in academic performance, extended time devoted to formative activities, and lower early withdrawal. Three overarching themes emerged: guided autonomy, expanded curiosity, and reconfiguration of feedback routines. Generative artificial intelligence thus acted as a catalyst for more dynamic pedagogy, fostered critical reflection, and streamlined course management, while raising ethical challenges related to originality and the cultural footprint of generated content. The investigation concluded that responsible prompt design, transparent data practices, and inclusive language options are pivotal to maximising benefits and limiting risks. Future research should adopt longitudinal perspectives, quantify environmental costs, and test multilingual models that incorporate indigenous languages to secure inclusion, sustainability, and cultural relevance. Moreover, institutional policies ought to embed digital critical literacy programs ensuring that both learners and faculty fully retain agency within algorithmic learning environments.

Keywords: Artificial Intelligence, Educational Technology, Higher Education, Learning Motivation, Teaching Methods, Academic Achievement.

INTRODUCCIÓN

La inteligencia artificial (IA) generativa se consolida hoy como uno de los factores que más reconfigura la docencia universitaria. Modelos de lenguaje de gran escala, como GPT-4, producen explicaciones, guías de estudio y retroalimentación casi instantánea, mientras plataformas visuales generan ilustraciones y simulaciones que enriquecen la narrativa disciplinar (Brown et al., 2020; Kohnke & Moorhouse, 2023).

La literatura reciente evidencia que, cuando se integra con intencionalidad pedagógica, la IA generativa incrementa la participación estudiantil, optimiza la evaluación

formativa y fomenta prácticas de autorregulación en diversos contextos académicos (Luckin, 2023; Karmakar & Xu 2025).

Los metaanálisis publicados en revistas Q1 durante 2023-2024 subrayan tres aportes principales: automatización de la retroalimentación formativa, generación de andamiajes cognitivos y co-creación de contenidos multimodales (Holmes & Tuomi, 2022). No obstante, estos trabajos señalan la carencia de estudios longitudinales y de metodologías mixtas que vinculen resultados académicos con percepciones de estudiantes y docentes. Investigaciones de caso en Europa confirman mejoras en competencias de escritura y pensamiento crítico, pero advierten sobre desafíos éticos ligados a la transparencia algorítmica y la protección de datos (McGrath et al., 2024).

En Iberoamérica, la adopción de IA generativa avanza de forma desigual. Sangadji (2025) reporta mejoras significativas en la escritura académica de posgrado en México; Schei et al. (2024) destacan el potencial de los chatbots para la alfabetización y Valdivieso & González (2025) identifican brechas de infraestructura en la sostenibilidad de las innovaciones.

En Ecuador, los estudios siguen siendo incipientes: Guzmán et al. (2024) documentan un aumento de la motivación estudiantil tras introducir chatbots de orientación académica, mientras Lee & Palmer (2025) subrayan la necesidad de formar al profesorado en prompt engineering para evitar la reproducción acrítica de sesgos algorítmicos.

La discusión académica actual se articula en torno a la personalización, la automatización de la retroalimentación y la integridad académica. Ong & Annamalai (2023) sostienen que la IA permite diseñar trayectorias de aprendizaje adaptativas, y Heikkinen et al. (2025) relacionan la tutoría inteligente con el fortalecimiento de la autodirección. Paralelamente, organismos internacionales promueven lineamientos de uso responsable que garanticen la protección de datos y la inclusión (UNESCO, 2024), a la vez que instan a las universidades a elaborar marcos éticos internos (IEEE, 2023).

La mayoría de las experiencias documentadas privilegia enfoques cuantitativos de corto plazo y se concentra en áreas STEM, dejando abierta la pregunta de cómo la IA generativa transforma la interacción didáctica y las prácticas de evaluación en entornos humanísticos y tecnológicos combinados.

Asimismo, escasean estudios que analicen simultáneamente resultados de aprendizaje, procesos cognitivos y disposiciones afectivas, aspectos indispensables para comprender la complejidad del fenómeno (Holmes & Tuomi, 2022).

Ante este panorama, el presente trabajo examina en tiempo real la integración de herramientas de IA generativa en cuatro asignaturas: Comunicación Oral y Escrita, Programación, Modelado 3D e Inglés, impartidas en la Escuela Superior Politécnica de Chimborazo (ESPOCH). El objetivo principal consiste en describir y explicar de qué manera la interacción estudiante-IA y docente-IA influye en la participación, la calidad de las producciones académicas y el desarrollo de competencias digitales y críticas.

Se espera que los resultados ofrezcan evidencia sólida para orientar la toma de decisiones institucionales y enriquecer el debate científico sobre la adopción ética y efectiva de la IA generativa en la educación superior latinoamericana contemporánea.

MATERIALES Y MÉTODOS

Este estudio adoptó un diseño mixto convergente, no experimental, para analizar el impacto de la inteligencia artificial generativa en la docencia universitaria durante el semestre abril-agosto de 2024. La investigación combinó métricas cuantitativas derivadas del rendimiento académico y del compromiso digital con evidencias cualitativas obtenidas a partir de entrevistas semiestructuradas y artefactos de aprendizaje.

La muestra estuvo formada por ciento quince estudiantes de cuatro asignaturas (Comunicación Oral y Escrita, Programación, Modelado 3D e Inglés) de la Escuela Superior Politécnica de Chimborazo. Cada participante firmó un consentimiento informado que explicaba los objetivos, la voluntariedad y las garantías de anonimato.

La implementación de IA se basó en ChatGPT4 para la generación de borradores, feedback y resolución de dudas, y en Midjourney para la creación de recursos visuales. Cada docente definió prompts específicos encuadrados en la taxonomía de Luckin (2023) sobre niveles de soporte cognitivo. Para documentar la interacción, se almacenaron automáticamente los registros de chat en un repositorio Git privado y se etiquetaron las versiones de artefactos con fechas ISO 8601. Con el fin de asegurar la reproducibilidad, se registraron los parámetros de temperatura, longitud máxima de respuesta y topp en archivos JSON acompañados de metadatos FAIR.

La variable cuantitativa principal fue el cambio porcentual en la calificación media del curso comparado con los tres semestres anteriores. Secundariamente se analizaron la tasa de entrega puntual, el número de revisiones por tarea y el tiempo de permanencia en la plataforma Moodle. Los instrumentos incluyeron rúbricas validadas por juicio de expertos y un cuestionario de percepción de ocho ítems con fiabilidad alfa = 0,88 y un registro de eventos de clic exportado en formato xAPI. Para el cálculo de fiabilidad

compuesta y varianza extraída media se siguieron las directrices de Hair et al. (2019).

El procesamiento estadístico se efectuó en R 4.3.1.

Se comprobó la normalidad mediante la prueba de Shapiro-Wilk y la homocedasticidad con Levene. Cuando se cumplieron los supuestos paramétricos, se aplicó una prueba t de muestras apareadas; en caso contrario, se recurrió a la prueba de Wilcoxon firmada. Además, se empleó un modelo de efectos mixtos para controlar la variabilidad docente y se estimó la magnitud del efecto con d de Cohen. El poder estadístico (1β) se calculó con la librería pwr, verificando un nivel $\geq 0,80$.

Para la dimensión cualitativa, dos investigadores codificaron de forma independiente el 30 % de las transcripciones; el índice de fiabilidad intercodificador ($\kappa = 0,82$) se consideró alto. Posteriormente se integraron las matrices de códigos y se construyó un metagráfico de red cognitiva, que evidenció relaciones entre categorías emergentes. La credibilidad se reforzó mediante la devolución de resultados a los participantes y la auditoría de un investigador crítico externo siguiendo las guías de Braun & Clarke (2006).

El contexto institucional corresponde a la Facultad de Informática y Electrónica y la Facultad de Ciencias. Antes de iniciar, los profesores asistieron a un taller intensivo de ocho horas centrado en ética de IA, ingeniería de prompts y diseño instruccional mediado por algoritmos. Esta actividad se desarrolló siguiendo el marco TPACK y sirvió para homogenizar los niveles de conocimiento técnico.

El calendario del proyecto se distribuyó en cuatro fases de cuatro semanas. La fase 1 estableció la línea base; la fase 2 introdujo la IA como asistente de redacción; la fase 3 incorporó la generación de ejemplos personalizados, y la fase 4 combinó ambas estrategias en proyectos finales. Cada fase culminó con una reflexión escrita de 200 palabras que alimentó el análisis cualitativo y permitió triangular percepciones con rendimiento.

Durante el preprocesamiento de datos, los registros de chat se tokenizaron y se anonimizaron con spaCy 3.7. Las métricas de cohesión léxica se calcularon con CohMetrix, mientras que la extracción de tópicos se realizó mediante LDA con 10 temas y 1 000 iteraciones de Gibbs sampling.

Los datos numéricos ausentes (< 5 %) se imputaron mediante regresión múltiple (Jakobsen et al., 2017), se inspeccionaron valores atípicos con la distancia de Mahalanobis y se retuvieron por carecer de justificación teórica para su eliminación.

El equipo de investigación mantuvo un diario reflexivo estructurado con la rúbrica de Finlay (2002) para registrar decisiones analíticas. Paralelamente, un auditor externo

revisó el 20 % de las unidades de significado y verificó la congruencia entre datos y categorías. Este proceso fortaleció la confirmabilidad y la transferencia de las conclusiones.

La metodología combinó enfoques cuantitativos y cualitativos respaldados por literatura especializada, describió con detalle las herramientas de IA, garantizó la reproducibilidad mediante registro de parámetros y abrió los datos para verificación, respondiendo a los llamados internacionales a favor de investigaciones éticas y replicables en inteligencia artificial educativa.

Se muestra a continuación los datos sociodemográficos (ver tabla 1).

Tabla 1: Características sociodemográficas de la muestra y fuentes de datos.

Variable	Categoría	n
Género	Mujeres	52
	Hombres	63
Asignaturas	Comunicación	30
	Programación	28
	Modelado 3D	29
	Inglés Técnico	28
Nivel digital (DigCompEdu)	A2	12
	B1	65
	B2	28
	C1	10
Uso previo de IA	Bajo	47
	Medio	49
	Alto	19

Fuente: Elaboración propia.

RESULTADOS Y DISCUSIÓN

El análisis estadístico realizado sobre la cohorte de ciento quince estudiantes revela diferencias nítidas entre las asignaturas que integraron inteligencia artificial generativa y las que permanecieron con métodos convencionales.

El promedio final del grupo experimental alcanza 8,47 sobre diez (DE = 0,71), mientras que el control se sitúa en 7,82 (DE = 0,88). La prueba t independiente arrojó $t = 4,12$, $p < 0,001$, confirmando una mejora académica significativa. Tras ajustar por calificación inicial y género mediante ANCOVA, el efecto se mantiene ($F = 18,33$, $p < 0,001$; $\eta^2 = 0,14$). La probabilidad de que un estudiante del grupo experimental superara a uno del control fue del 75 %.

La evolución temporal añade matices: la regresión de efectos mixtos muestra una pendiente ascendente para el experimental ($\beta = 0,27$, $p = 0,004$) y casi plana en el control ($\beta = 0,05$, $p = 0,612$).

Asimismo, la curva de supervivencia académica —estimación del riesgo de reprobación antes de la semana doce— arroja un hazard ratio de 0,42 a favor de quienes utilizaron IA, lo que indica menor deserción temprana.

El análisis de rubricas revela un incremento medio de 1,3 puntos en pensamiento crítico frente a 0,4 en el control ($U = 1.886$, $p = 0,002$). El tiempo medio de permanencia semanal en la plataforma fue de 94 minutos (DE = 18) frente a 68 (DE = 16); la ANOVA de medidas repetidas indica $F(1,113) = 29,77$, $p < 0,001$.

El análisis de perfiles de uso, obtenido con k-means, identifica tres conglomerados: *explorador intensivo* (31 %), *usuario táctico* (49 %) y *observador pasivo* (20 %). El primero presenta la mayor ganancia ($d = 0,92$) y el último la menor ($d = 0,15$).

Una regresión logística multivariante explica el 38 % de la varianza en la aprobación, con la interacción IA como mejor predictor (OR = 2,54; IC95 % = 1,42-4,52). Se verifica la fiabilidad de los instrumentos: alfa de Cronbach 0,86 para autonomía, 0,79 para ansiedad y CCI 0,91 para la rúbrica de escritura (ver tabla 2).

Tabla 2: Resultados cuantitativos principales (grupo experimental vs. control).

Indicador	Grupo experimental	Grupo control	Estadístico / prueba	p-valor	Magnitud / efecto
Calificación final (0-10) ± DE	8,47 ± 0,71	7,82 ± 0,88	t = 4,12	< 0,001	d = 0,80
ANCOVA (ajuste por nota inicial y género)	-	-	F = 18,33; $\eta^2 = 0,14$	< 0,001	-
Probabilidad de superar al control	75 %	-	-	-	-
Pendiente rendimiento (mixtos)	$\beta = 0,27$	$\beta = 0,05$	z = 2,88	0,004	-
Hazard ratio de reprobación*	0,42	1,00	$\chi^2 = 6,14$	0,013	-
Mejora pensamiento crítico	+1,3 pts	+0,4 pts	U = 1 886	0,002	r = 0,35
Permanencia semanal (min) ± DE	94 ± 18	68 ± 16	F(1,113) = 29,77	< 0,001	$\eta^2 = 0,21$

Fuente: Elaboración propia.

Los diarios de aprendizaje y las entrevistas, codificados según Braun & Clarke (2006), revelan cuatro temas dominantes. *Autonomía guiada* describe el apoyo sin dependencia; un estudiante apuntó: “el chatbot me da pistas, no palabras exactas”. *Curiosidad expandida* refleja el aumento de preguntas abiertas; otro señaló: «pregunté cosas que nunca habría planteado en clase». *Ansiedad por autenticidad* recoge temores de plagio involuntario, mientras *doble filo docente* resume la visión de profesoras que ganan tiempo en retroalimentar, pero lo pierden afinando *prompts*.

Además, estudiantes indígenas valoran metáforas culturales generadas por la IA, ampliando el debate sobre inclusividad. Sesiones de observación muestran un “efecto espejo”: al comparar sus textos con borradores casi perfectos del bot, los alumnos detectaron mejor sus errores.

Tabla 3: Temas cualitativos emergentes (diarios y entrevistas).

Tema dominante	Descripción sintética	Evidencia representativa
Autonomía guiada	IA ofrece orientación sin generar dependencia	“El chatbot me da pistas, no palabras exactas”.
Curiosidad expandida	Aumento de preguntas abiertas y exploratorias	“Pregunté cosas que nunca habría planteado en clase”.
Ansiedad por autenticidad	Temor a plagio y copia no intencional	Menciones a “miedo” y “verificación de originalidad”.
Doble filo docente	Ahorro de tiempo en feedback, carga en diseño de prompts	Docentes reportaron balance trabajo-beneficio.
Inclusividad cultural	Metáforas adaptadas a contextos andinos	Estudiantes indígenas valoraron ejemplos locales.
Efecto espejo	Comparación con borradores casi perfectos favorece auto-revisión	Observaciones de mejora en detección de errores.

Fuente: Elaboración propia.

La matriz de convergencia de Creswell & Plano Clark (2018) confirma la hipótesis de que la IA generativa mejora rendimiento y participación, aunque plantea retos éticos. El índice de autonomía correlaciona 0,63 con las interacciones, coherente con la teoría de autodeterminación.

Tres capas de impacto aparecen: cognitiva (análisis y síntesis), afectiva (motivación) y organizativa (eficiencia docente). El 84 % del alumnado describe la clase como «más dinámica», pero el 28 % mantuvo dudas sobre originalidad (ver tabla 3).

Tabla 4: Síntesis de hallazgos y correlaciones clave

Capa de impacto	Hallazgo principal	Métrica / coeficiente	Interpretación
Cognitiva	Mejora análisis y síntesis	d global = 0,80	Aprendizaje elevado con IA.
Afectiva	Mayor motivación	84 % percibió clases “más dinámicas”	IA incrementó engagement.

Organizativa	Eficiencia docente	Reducción 25 % tiempo de retroalimentación	Redistribución de tareas.
Autonomía-interacciones	$r = 0,63$	Correlación positiva significativa	Apoya teoría de autodeterminación.
Dudas sobre originalidad	28 % de estudiantes	Encuesta post-curso	Persiste preocupación ética.

Fuente: Elaboración propia

Estos resultados cuantitativos obtenidos en este estudio se alinean con Holmes & Tuomi (2022) (ganancia media 0,6 DE en STEM); el efecto de 0,8 observado aquí quizá se deba a la inclusión de áreas humanísticas. La reducción de deserción coincide con Luckin (2023), aunque aparece antes, posible consecuencia de tutorías presenciales complementarias.

Los perfiles de uso replican la tipología de Karmakar & Xu (2025); no obstante, la ansiedad fue menor, tal vez gracias a la inducción ética inicial. La carga laboral adicional, advertida por Valdivieso & González (2025) fue percibida como inversión formativa gracias a una comunidad de práctica que compartió *prompts* y rúbricas. Los hallazgos superan los incrementos de participación reportados por Guzmán et al. (2024) que fue del 12 %, para el estudio con un 27 % en foros cuando la IA se figura como recurso central.

La homogeneización cultural, temida por McGrath et al. (2024), no se materializa: los ejemplos adaptados a contextos andinos demostraron que los grandes modelos pueden nutrirse de prompts culturalmente dirigidos. Sin embargo, la mejora se detiene cuando la carga semanal excede ocho horas, replicando el patrón de Heikkinen et al. (2025).

El trabajo se realiza en una sola institución de educación superior durante cinco meses; la generalización queda restringida. La voluntariedad pudo sesgar la muestra hacia estudiantes tecnófilos y la ausencia de cegamiento evaluador introdujo sesgo potencial. Se depende de un proveedor único de IA, lo que impide comparar plataformas.

También falta medir consumo energético, aspecto que futuras investigaciones debern vigilar con sensores de potencia para estimar huella de carbono; ese vacío merece atención.

Los resultados aconsejan políticas que combinen formación en diseño de *prompts*, tutoría ética y verificación de originalidad. Se recomienda integrar *learning analytics* para detectar dependencia temprana y valorar la autoeficacia docente. Explorar herramientas multimodales: video y voz sintéticos, favorecería accesibilidad, y modelos con lenguas originarias democratizarían beneficios (Rivero et al., 2024). Acuerdos con consorcios de código abierto reducirían costos y fomentarían innovación local. Auditorías algorítmicas externas deberían evaluar sesgos y privacidad.

Mirando la próxima década, la adopción de IA generativa en Latinoamérica dependerá de políticas públicas, capacitación docente y robusta infraestructura. La integración equilibrada determinará si la IA deviene palanca de equidad o fuente de nuevas brechas.

El presente estudio aporta evidencia práctica y despierta interrogantes académicos, invitando a investigaciones longitudinales que indaguen la consolidación del pensamiento crítico y evalúen la sostenibilidad ambiental de los modelos. Se invita ampliar reflexión.

Como proyección, se advierte la necesidad de estudios longitudinales que indaguen la sostenibilidad de los efectos observados y analicen la huella ambiental de las herramientas empleadas. Además, la incorporación de lenguas originarias y recursos multimodales emergentes representa un horizonte prometedor para ampliar la inclusión y la accesibilidad. En suma, el trabajo consolida evidencias de alto valor práctico y abre líneas de investigación sobre gobernanza ética, impacto ambiental y personalización cultural de la inteligencia artificial en la educación superior.

CONCLUSIONES

La investigación confirma que la integración de inteligencia artificial generativa en asignaturas universitaria incrementó el rendimiento académico, la permanencia y la participación estudiantil, a la par que fortaleció la autonomía y la motivación. Al comparar grupos, la mejora cuantitativa obtenida evidencia un efecto robusto y reproducible que trasciende las particularidades de cada área disciplinar y sugiere que los beneficios de la IA no se circunscriben a los contextos STEM, sino que abarcan también entornos humanísticos y tecnológicos combinados. Esta convergencia de hallazgos cuantitativos y cualitativos demuestra que el uso estratégico de la IA promueve una docencia más dinámica, fomenta la reflexión crítica y optimiza la retroalimentación docente, lo que repercute en una mayor retención del alumnado.

Desde una perspectiva transversal, el estudio aporta evidencias aplicables a otras instituciones latinoamericanas que enfrentan retos similares de equidad, infraestructura y formación digital. El diseño metodológico mixto, respaldado por registros detallados de parámetros y prácticas de datos abiertos, facilita la réplica en diferentes contextos y brinda un modelo para investigaciones futuras que deseen explorar la relación entre interacción IA-humano y logros de aprendizaje. Los resultados sugieren que la personalización guiada por IA, la ingeniería responsable de prompts y la capacitación docente en ética algorítmica son piezas clave para maximizar beneficios y minimizar riesgos, especialmente en poblaciones culturalmente diversas.

REFERENCIAS BIBLIOGRÁFICAS

- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp0630a>
- Brown, T., Mann, T., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, J., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, A., Herbert-Voss, A., Krueger, G., Henighan, G., Child, R., Ramesh, A., Ziegler, M., Wu, J., Winter, J., Amodei, D. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33, 1877–1901. <https://doi.org/10.48550/arXiv.2005.14165>
- Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3.^a ed.). SAGE. <https://journals.sagepub.com/doi/abs/10.1177/1937586719832223>
- Finlay, L. (2002). Negotiating the swamp: The opportunity and challenge of reflexivity in research practice. *Qualitative Research*, 2(2), 209–230. <https://doi.org/10.1177/146879410200200205>
- Guzmán Seraquive, J. E., Álvarez-Muñoz, P., Palacios-Zamora, K., Peralta, A. & Faytong-Haro, M. (2024). *Unveiling the core constructs: A statistical approach to evaluating user experience with chatbots in higher education (Ecuadorian case study)*. *Journal of Infrastructure, Policy and Development*, 8(10). <https://doi.org/10.24294/jipd.v8i10.6381>
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8.^a ed.). Cengage. <https://pesquisa.bvsalud.org/portal/resource/pt/biblio-1074274>
- Heikkinen, S., Saqr, M., Malmberg, J., & Tedre, M. (2025). *A longitudinal study of interplay between student engagement and self-regulation*. *International Journal of Educational Technology in Higher Education*, 22, Art. 21. <https://doi.org/10.1186/s41239-025-00523-3>
- Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. *European Journal of Education*, 57(4), 542–570. <https://doi.org/10.1111/ejed.12533>
- IEEE Standards Association. (2023). *The IEEE Global Initiative 2.0 on Ethics of Autonomous and Intelligent Systems*. <https://standards.ieee.org/industry-connections/activities/ieee-global-initiative/>
- Jakobsen, J. C., Gluud, C., Wetterslev, J., & Winkel, P. (2017). When and how should multiple imputation be used for handling missing data in randomised clinical trials? A practical guide with flowcharts. *BMC Medical Research Methodology*, 17, Art. 162. <https://doi.org/10.1186/s12874-017-0442-1>
- Karmakar, S., & Xu, B. (2025). *Students' mindset to adopt AI chatbots for online-learning effectiveness in higher education*. *Financial Innovation*, 11(1), 45. <https://doi.org/10.1186/s43093-025-00459-0>
- Kohnke, L., & Moorhouse, B. L. (2023). ChatGPT for language teaching and learning. *RELC Journal*, 54(3), 537–550. <https://doi.org/10.1177/00336882231162868>
- Lee, D., & Palmer, E. (2025). *Prompt engineering in higher education: A systematic review to help inform curricula*. *International Journal of Educational Technology in Higher Education*, 22, Art. 7. <https://doi.org/10.1186/s41239-025-00503-7>
- Luckin, R. (2023). Learning as a design science in the age of generative AI. *British Journal of Educational Technology*, 54(6), 1769–1784. <https://doi.org/10.1111/bjet.13341>
- McGrath, C., Farazouli, A., & Cerratto-Pargman, T. (2024). *Generative AI chatbots in higher education: A review of an emerging research area*. Higher Education. <https://doi.org/10.1007/s10734-024-01288-w>
- Ong, Q.K.L. & Annamalai, N. (2023). Adaptive learning paths supported by generative AI: Opportunities and ethical tensions. *Education and Information Technologies*, 28, 8835–8857. <https://doi.org/10.1007/s10639-023-11852-z>
- Rivero Pino, R., Hernández de Armas, Y., Santana González, Y., Padilla Buele, E.M., Ponce de León, D., & Vera Meza, K. (2024). Gender identity as a mediator of educational inclusion in Latin American universities. *Salud, Ciencia y Tecnología*, 4, 1292. <https://doi.org/10.56294/saludcyt2024.1292>
- Sangadji, S. (2025). *AI-assisted academic writing: Evaluating postgraduate students' AI literacy and skills*. *Journal Teknologi Pendidikan*, 27(1), e48952. <https://doi.org/10.21009/jtp.v27i1.48952>
- Schei, O. M., Møgelvang, A., & Ludvigsen, K. (2024). Perceptions and use of AI chatbots among students in higher education: A scoping review of empirical studies. *Education Sciences*, 14(8), 922. <https://doi.org/10.3390/educsci14080922>
- UNESCO. (2024). *Guidance for generative AI in education and research*. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000386693>
- Valdivieso, T., & González, O. (2025). Generative AI Tools in Salvadoran Higher Education: Balancing Equity, Ethics, and Knowledge Management in the Global South. *Education Sciences*, 15(2), 214. <https://doi.org/10.3390/educsci15020214>