

Inteligencia Artificial como recurso didáctico y rendimiento académico en estudiantes de Educación General Básica y Bachillerato.

Artificial Intelligence as a Teaching Resource and Academic Performance among General Basic Education and Baccalaureate Students.

Mayra Elizabeth Ramón Ordoñez ¹

Ministerio de Educación, Deporte y Cultura del Ecuador

(mayra.ramono@educacion.gob.ec) <https://orcid.org/0009-0002-8195-2240>

Norma Piedad Castillo Castillo ²

Ministerio de Educación, Deporte y Cultura del Ecuador

(normap.castillo@educacion.gob.ec) <https://orcid.org/0009-0008-3103-9748>

Grecia Greolandia Lucas Muñoz ³

Ministerio de Educación, Deporte y Cultura del Ecuador

(grecia.lucas@educacion.gob.ec) <https://orcid.org/0009-0003-2110-188X>

Johanna Rosa Barrera Salazar ⁴

Ministerio de Educación, Deporte y Cultura del Ecuador

(johannar.barrera@educacion.gob.ec) <https://orcid.org/0009-0000-4255-451X>

Bryan Ricardo Morales Naranjo ⁵

Ministerio de Educación, Deporte y Cultura del Ecuador

(bryan.morales@docentes.educacion.edu.ec) <https://orcid.org/0009-0009-3154-6456>

Cómo citar

Ramón Ordoñez, M. E., Castillo Castillo, N. P., Lucas Muñoz, G. G., Barrera Salazar, J. R., & Morales Naranjo, B. R. (2026). Inteligencia Artificial como recurso didáctico y rendimiento académico en estudiantes de Educación General Básica y Bachillerato. *Saber Estratégico Internacional*, 4(3), 768–785. <https://doi.org/10.70577/s.e.v4i3.83>

RESUMEN

La inteligencia artificial (IA) abre nuevos horizontes como recurso didáctico, pero incluirla en el aula conlleva evidencias sobre su relación con el aprendizaje. Se persiguió conocer las diferencias supeditadas al rendimiento académico en una intervención de integración didáctica de IA en estudiantes de Educación General Básica y Bachillerato. Se llevó a cabo un estudio de corte cuantitativo, cuasiexperimental, con grupos naturales no equivalentes y mediciones de pretest-posttest. Tomaron parte en él 420 estudiantes, 242 en condición de intervención y 178 en condición de comparación. La intervención se realizó durante las 12 semanas a partir de actividades académicas apoyadas por IA, mientras que la comparación experimentó la enseñanza habitual y los procedimientos de este tipo de herramientas de forma no estructurada. El rendimiento aumentó significativamente entre pretest y posttest en ambas condiciones. Sin embargo, la ganancia fue prácticamente idéntica entre los grupos y no se observaron diferencias significativas en el rendimiento posttest después de ajustar por rendimiento basal. Los análisis con covariables y considerando el agrupamiento por aula mantuvieron esta conclusión. Se concluye que, bajo las condiciones evaluadas, la integración didáctica estructurada de IA no estuvo asociada con un rendimiento académico adicional respecto de la enseñanza habitual. Los hallazgos respaldan una incorporación pedagógicamente planificada de la IA y una evaluación rigurosa de sus resultados educativos.

Palabras clave: inteligencia artificial; tecnología educativa; rendimiento académico; enseñanza; aprendizaje.

ABSTRACT

Artificial intelligence (AI) is opening new possibilities as an instructional resource, yet its integration into the classroom requires empirical evidence regarding its relationship with student learning. This study examined differences in academic performance associated with a classroom intervention involving the pedagogical integration of AI among students in General Basic Education and Upper Secondary Education. A quantitative, quasi-experimental design was employed, using nonequivalent naturally occurring groups and pretest–posttest measurements. The study included 420 students, of whom 242 participated in the intervention condition and 178 in the comparison condition. Over a 12-week period, students in the intervention group engaged in structured academic activities supported by AI, whereas those in the comparison group received regular instruction and used AI-based tools without a structured pedagogical framework. Academic performance improved significantly from pretest to posttest in both conditions. However, gains were nearly identical across groups, and no significant differences in posttest performance were found after adjusting for baseline achievement. Analyses incorporating covariates and accounting for classroom-level clustering yielded the same overall conclusion. Under the conditions examined, the structured pedagogical integration of AI was not associated with additional gains in academic performance compared with regular instruction. These findings support the need for pedagogically grounded AI integration and rigorous evaluation of its educational outcomes.

Keywords: artificial intelligence; educational technology; academic achievement; teaching; learning.

INTRODUCCION

La aplicación de la inteligencia artificial (IA) en los entornos educativos representa uno de los cambios tecnológicos más destacados de los últimos años. El desarrollo de sistemas de IA generativa que son capaces de producir explicaciones, responder preguntas, generar ejemplos, revisar textos o proporcionar retroalimentación ha ampliado las opciones de interacción entre estudiantes, docentes y recursos digitales. En el contexto escolar, este tipo de herramientas puede ayudar con la búsqueda y utilización de información, la resolución de problemas, la revisión de o los productos y la autoevaluación. Sin embargo, la disponibilidad de tales herramientas no significa que obtendremos beneficios de su uso, dado que la forma en que se utilizan dependerá de la finalidad pedagógica, del seguimiento por parte del docente y de las estrategias utilizadas. Un análisis encaminado a la educación K–12 identifica oportunidades vinculadas a la personalización, al apoyo del aprendizaje y a desafíos referidos a la calidad de las respuestas, la integridad académica, y la implementación pedagógica (Zhang & Tur, 2024). La IA sería una de las partes que viene a completar un proceso más amplio de renovación de las prácticas educativas donde la tecnología viene a tener razón de ser porque se une a decisiones didácticas y a necesidades del alumnado. En educación básica han ido apareciendo propuestas inspiradas en el Diseño Universal para el Aprendizaje y en metodologías inclusivas centradas en la diversificación de oportunidades de participación (Aguilar Tinoco et al., 2024; Bernal Párraga, Medina Marino, et al., 2024), mientras que la revisión de recursos digitales en Ciencias Naturales muestra cómo la tecnología va poco a poco incorporándose a los espacios de aula (Bernal Parraga, Orozco Maldonado, et al., 2024); en consecuencia, las herramientas deben elegirse en función de los objetivos didácticos y no a partir de su propia disponibilidad.

Desde esta visión, se da un valor añadido al interés por observar cómo la IA se implementa didácticamente y qué conexiones puede tener con aquellos resultados relevantes. El rendimiento académico permite evaluar el rendimiento académico antes y después de un proceso educativo cuando hay mediciones comparables en el mismo contexto curricular. Un metaanálisis sobre estudios, los cuales fueron publicados entre 2022 y 2025, reveló un efecto global positivo moderado de ChatGPT sobre el rendimiento académico, aunque con variaciones en función de la disciplina, su duración, el tipo de conocimiento y el estilo pedagógico (Liu et al., 2025), lo que sugiere que el potencial educativo de la IA depende, en buena medida, de las condiciones concretas de su integración.

Esta idea está de acuerdo con algunas aproximaciones producidas en el marco de las diversas áreas curriculares. De este modo, hemos tratado la resolución de problemas como una estrategia para favorecer el pensamiento lógico en Matemáticas (Alvarez Piza et al., 2024), mientras que el role-playing y otras metodologías activas han sido estudiadas en experiencias relacionadas con la creatividad, el pensamiento crítico y el aprendizaje en Lengua (Bernal Párraga, Toapanta Guanoquiza, et al., 2024; Guerrero Carrera et al., 2024). Asimismo, han sido puestas a punto unas propuestas de innovación específica en lo relativo a la didáctica de Lengua y Literatura (Mora Villamar et al., 2024). Estos antecedentes permiten conceptualizar la IA como un recurso integrable en el seno de estrategias didácticas construidas, no como una metodología aislada y de calidad autónoma.

La evidencia proveniente de los contextos más escolares es de gran relevancia. Estudiantes de siete curso han trabajado la integración de IA generativa en actividades de historia, considerando logros, motivación y carga cognitiva (Chen et al., 2025). En educación científica de secundaria, ChatGPT ha sido integrado como

apoyo al aprendizaje autorregulado en el marco de actividades didácticas constituidas (Ng et al., 2024). Asimismo, ha sido hallada una asociación entre el uso de ChatGPT como apoyo académico, motivación y rendimiento en estudiantes de secundaria (Caratiquit & Caratiquit, 2023). Estos trabajos evidencian la pertinencia de estudiar la IA en un sentido doble: resultados y procesos en la experiencia de aprender. Entre estos procesos son destacables, la motivación, la autonomía y la personalización. En el caso de la adolescencia, el uso de IA generativa para tareas escolares y su utilidad percibida dan cuenta de características tanto individuales como académicas, y ello hace desaconsejable asumir que sus resultados son uniformes (Klarin et al., 2024). En el caso de la educación superior, como evidencia complementaria no extrapolable a Educación Básica, el uso de IA generativa ha sido explorado para ayudar en el aprendizaje autodirigido y la autoevaluación (Ali et al., 2023). Esa misma evidencia complementaria fue también explorada en el caso de estudiantes universitarios, en donde se han estudiado tanto las percepciones y formas de uso de la IA generativa, así como los factores que se relacionan con su incorporación al aprendizaje (Han et al., 2025; Shin et al., 2025; Zhao et al., 2025). En lenguas, también se han descrito plataformas adaptativas basadas en IA (Chicaiza, 2025), así como aplicaciones de chatbots y asistentes virtuales para el aprendizaje personalizado (Jara Chiriboga et al., 2025). En su conjunto, estas posibilidades deberían ser consideradas dependientes del diseño pedagógico y no como resultados que se deriven automáticamente del acceso a la tecnología.

El pensamiento crítico está formado por otra dimensión esencial. Como la información que las IA generan puede contener errores, derivar en simplificaciones o estar insuficientemente sustentada, la utilización de las IA en entornos educativos requiere poder comparar la información, verificar fuentes y cuestionar las respuestas. La literatura invita a que la utilización de IA generativa vaya de la mano con el desarrollo intencionado del pensamiento crítico (Lee & Low, 2024), en tanto que la experiencia estudiantil demuestra que puede haber comportamientos reflexivos y dependientes poco críticos a la vez de utilizar ChatGPT (Šedlbauer et al., 2024). Así, el valor pedagógico aumenta pensando que en determinados trabajos las actividades exigen ciertas formas de analizar y contrastar con las respuestas generadas en vez de aceptarlas como respuestas acabadas.

Esta cuestión está vinculada con los usos éticos y el riesgo de dependencia. Entre la población estudiantil secundaria se han identificado no solamente oportunidades educativas sino también riesgos de dependencia y procesos cognitivos, lo cual refuerza la necesidad de alfabetización en IA y de acompañamiento (Adisa & Adefisayo, 2025). Asimismo, el análisis de los factores psicológicos y contextuales vinculados a la adopción de IA en adolescentes muestra que el uso que hacen de dicha tecnología está mediado por características personales y escolares (Ballestra Caffaratti et al., 2025). El balance entre las oportunidades y los riesgos a tener en cuenta en la evaluación educativa de dichas tecnologías ha de ser contemporáneo.

La función del profesorado y el diseño didáctico son centrales. En las revisiones estudiadas se encuentran oportunidades que refuercen la enseñanza y el aprendizaje, pero también cuestiones relacionadas con la exactitud, el sesgo, la autenticidad y la integridad académica (Ali et al., 2024). Las revisiones abordadas en educación superior y en la formación médica coinciden en la existencia de oportunidades educativas estimuladas por la IA, sino que también advierten de los desafíos que esta puede generar en su implementación, evaluación y uso responsable (Mortlock & Lucas, 2024; Simoni et al., 2025; Varma et al., 2023; Xu et al., 2024). Los datos complementarios universitarios también apuntan a que se puede incluir la

IA mediante soluciones por tareas, retroalimentación y aprendizaje activo independiente; aunque los resultados de la Educación Superior no permiten extraer la información que puede ser útil para la Educación Básica (Gan et al., 2024; Zeng et al., 2025). También se han evaluado estrategias de enseñanza universitaria que incorporaban chatbots generativos con actividades de carácter pedagógico (Shi et al., 2025). En este mismo sentido han sido articuladas experiencias de juego mediado por IA (Troya Santilán et al., 2024), modelos de acción en el que se cruzan las metodologías activas e IA en Ciencias Naturales (Bernal Párraga, Constante Olmedo et al., 2026) y propuestas que integran indagación y laboratorios virtuales (Bernal Párraga, Alvarado Cantos et al., 2026). En conjunto, estas aproximaciones ponen de relieve que la tecnología tiene sentido educativo en el marco de actividades y objetivos curriculares definidos.

Esta diferencia es especialmente importante dado que los adolescentes pueden afrontar herramientas generativas sin que la escuela haga nada. El uso informático espontáneo a la hora de realizar tareas o recibir explicaciones no equivale a participar en tareas scriptadas y planificadas. Los patrones del uso adolescente y el rendimiento académico son heterogéneos (Klarin et al., 2024), mientras que la evidencia K-12 sigue consolidándose (Zhang & Tur, 2024). Por eso, desde el punto de vista metodológico, resulta pertinente distinguir entre una integración didáctica estructurada de la inteligencia artificial (IA) y una enseñanza no estructurada general, a pesar de que haya un nivel de utilización informal de este tipo de herramientas.

La heterogeneidad contextual también lleva a tener que matizar. Se han generado propuestas pedagógicas vinculadas a las cómo las tecnologías emergentes y las metodologías activas vinculadas a la comprensión lectora y la escritura creativa (Bernal Párraga, Rodríguez Cajas, et al., 2026), propuestas pedagógicas innovadoras para la comprensión lectora y la producción textual (Cochancela Pérez et al., 2026), propuestas pedagógicas para Estudios Sociales (León Ruíz et al., 2024) y aplicaciones que vinculan IA con necesidades específicas del aprendizaje de la lectura (Mena Castillo et al., 2026). Estas propuestas demuestran que los resultados de una innovación en tecnología deben ser evaluados contextualizándolos con el currículum y con el resultado educativo específico considerado.

Aún así, a pesar de la expansión de la literatura, existe todavía un vacío relevante: de un lado, gran parte de la evidencia proviene de educación superior o se centra en percepciones de los usuarios o en patrones de adopción del estudiante (en general red) con lo que sigue siendo necesaria la evidencia contextualizada en el alumnado escolar que combine pretest-posttest, grupos naturales de comparación e integración didáctica estructurada de IA. Esta necesidad justifica tratar la tecnología como recurso que se integra a prácticas pedagógicas concretas y también cautela a la hora de inferir causalidades a partir de grupos no aleatorizados. Por consiguiente, el objetivo del presente estudio es dar cuenta de las diferencias en el rendimiento académico informadas por esta integración didáctica de la IA mediante pretest-posttest y grupos naturales no equivalentes.

MATERIALES Y MÉTODOS

Se llevó a cabo el presente estudio que fue cuantitativo, prospectivo y longitudinal para poder evaluar el rendimiento académico, con un diseño tipo de cuasiexperimental pretest-posttest con grupos naturales no equivalentes. Se agregó además un componente transversal postintervención con el propósito de caracterizar las percepciones y experiencias de los estudiantes al respecto del uso educativo de inteligencia artificial (IA).

La falta de asignación aleatoria constituyó una de las características centrales del diseño. Los estudiantes permanecieron en sus aulas o paralelos pre-establecidos y no fueron redistribuidos individualmente. Los grupos naturales se incorporaron a las condiciones de intervención o comparación de acuerdo a la disponibilidad institucional aceptable y a la factibilidad de implementación. Por lo tanto, los análisis se orientaron a considerar la no equivalencia de los grupos y no se asumió intercambiabilidad basal.

El desarrollo de la intervención se realizó en el período de abril a junio del 2026, y tuvo una duración aproximada de 12 semanas. La muestra analítica estuvo constituida por 420 sujetos de estudio de 9 a 19 años de edad procedentes de 64 instituciones educativas: 242 alumnos formaron la intervención didáctica con IA y 178 la condición de comparación. Los alumnos pertenecían a la Educación General Básica y Bachillerato de diferentes grados, materias y provincias. Se llevó a cabo un muestreo no probabilístico por conveniencia. Las instituciones educativas se seleccionaron en función de su accesibilidad, el consentimiento institucional, la disponibilidad de profesores y la viabilidad de la intervención. Para formar la muestra analítica era necesario pertenecer a alguno de los grupos de una de las direcciones previas, tener identidad institucional y condición de estudio, tener mediciones pretest y posttest de rendimiento académico y tener información suficiente sobre las principales variables sociodemográficas y académicas. Los 420 registros cumplían estos criterios. No hay documentación verificable que permita concluir una muestra inicial mayor ni poder estimar retrospectivamente las pérdidas en los registros que podrían haberse dado o por tasa de abandono. La exposición principal lo fue la condición de estudio, que fue clasificada como intervención didáctica estructurada mediante IA o enseñanza habitual sin estructuración didáctica mediante IA. El resultado longitudinal principal lo fue el rendimiento académico. 'Rendimiento_Pre' y 'Rendimiento_Post' correspondieron a las calificaciones obtenidas con la realización de evaluaciones llevadas a cabo por la parte docente en función de los contenidos curriculares trabajados antes del periodo de intervención o del momento posterior a este, respectivamente. No se utilizaron pruebas estandarizadas ni instrumentos de rendimiento diseñados a medida para la investigación. Ambas puntuaciones fueron expresadas en una escala de 0-100 puntos e hicieron referencia al contexto curricular, asignatura o área de aprendizaje de cada estudiante. En el contexto de la intervención entraron en juego distintos grados, asignaturas e instituciones, por lo que adoptar la escala común como evidencia de equivalencia psicométrica absoluta entre contextos no fue pertinente. Por lo tanto, las puntuaciones fueron consideradas indicadores del rendimiento académico registrado en cada contexto curricular y no en relación con el resultado de una prueba estandarizada común. Como resultado secundario, se realizó el cálculo de la ganancia individual de la siguiente manera:

$$\Delta \text{Rendimiento} = \text{Rendimiento_Post} - \text{Rendimiento_Pre}.$$

Hicimos que esta diferencia se calculase directamente a partir de las dos mediciones con el objetivo de evitar la discrepancia que puede darse cuando la variable que se está usando ha sido redondeada.

Se dispone también de edad, sexo, provincia, institución, nivel educativo, grado, asignatura, acceso digital, frecuencia de uso de IA previa (antes de la intervención) y herramienta IA principal utilizada. Las variables preintervención se utilizaron para la caracterización de la muestra, para analizar el equilibrio basal y para analizar resultados ajustados cuando considerábamos que existía una justificación conceptual.

Tras la intervención, se utilizó un cuestionario elaborado ad hoc de 40 ítems, estructurado en ocho dimensiones de cinco ítems: uso académico de IA, valor didáctico percibido, motivación y participación,

autonomía en el aprendizaje, pensamiento crítico ante contenidos generados por IA, aprendizaje percibido, uso ético y responsable y riesgo de dependencia o uso acrítico. Las respuestas fueron las obtenidas en una escala Likert 1=Totalmente en desacuerdo hasta 5=Totalmente de acuerdo; todos ellos fueron ítems de puntuación del tipo directo y el cálculo de cada dimensión fue a partir de los ítems promediados. En riesgo de dependencia, los mayores valores informan de mayor riesgo.

El cuestionario no se corresponde con una escala validada publicada en su totalidad anterior a su uso y no fue sometido en su momento a validación psicométrica propiamente dicha, V de Aiken, análisis factorial, pilotaje formal o validación externa, por lo que dicha validación no puede ser atribuido al estudio en retrospectiva; sus dimensiones fueron interpretadas como medidas exploratorias de percepciones y experiencias declaradas. Obtenerlo únicamente realizado al final de la intervención permite caracterizar el momento postintervención, pero no permiten estimar cambios intraindividuales ni atribuir las percepciones observadas a efectos de la intervención. La intervención tuvo una duración playera de 12 semanas. La intervención consistía en aproximadamente 12 sesiones semanales de 45–60 minutos durante las actividades académicas habituales y bajo la dirección de los profesores. Las actividades más comunes fueron la explicación de conceptos y preguntas guiadas; la resolución y la creación de ejemplos; la retroalimentación formativa; la búsqueda y el contraste de la información; el repaso de los textos y la autorreflexión; y la resolución de problemas contextualizados y de análisis crítico. La exposición pudo variar como consecuencia de la asistencia, actividades institucionales y planificación. Al no disponer de una medida individual suficientemente precisa de dosis o fidelidad acumulada, se asumió exposición heterogénea. La condición de comparación mantuvo la enseñanza habitual de los estudiantes durante el mismo periodo sin participar en el programa estructurado de la integración didáctica de IA. No se impidió la utilización espontánea o particular de ChatGPT, Gemini o cualquier otra herramienta que se empleara para las actividades no estructuradas; por tanto, esta condición se refiere a enseñanza habitual sin la intervención estructurada de IA, no a separación absoluta con IA.

Se tomaron el rendimiento basal y distintas características de acceso a la estructura de los estudiantes antes de la intervención. Tras las 12 semanas se definió el rendimiento del posttest y se pasó el cuestionario. Los registros fueron consolidados de acuerdo con identificadores anonimizados. 'ID_Institucion' identificó la institución educativa e 'ID_Aula' el aula o paralelo real, de forma que se alcanza la estructura jerárquica de los datos y es posible estudiar la dependencia de las observaciones bajo ciertas aulas.

Primero, se comprobaron duplicados, datos faltantes, rangos, coherencias entre variables y de datos potencialmente influyentes. Las variables continuas se resumieron mediante media y desviación estándar y categóricas a partir de frecuencias y porcentajes. Porque el diseño no era aleatorizado, las igualdades entre condiciones a nivel basal, se evaluaron mediante diferencias estandarizadas, además de descriptivas; un valor $*p*$ no significativo se interpretó como no evidencia de la igualdad basal.

El cambio en el rendimiento en cada condición se examinó mediante diferencias entre pretest-posttest, proporcionando la diferencia media, IC95 %, contrastes para medidas relacionadas, y el tamaño del efecto estandarizado intragrupo ($*d*z$). Estos resultados se interpretaron únicamente como cambios temporales intragrupo.

El análisis comparativo secundario pero también intra- entre condiciones, se realizó contrastando las medias de ganancias y proporcionando IC95% y el tamaño del efecto estandarizado. La estimación principal fue

realizada mediante un modelo lineal equivalente a ANCOVA, estableciendo 'Rendimiento_Post' como variable dependiente, el grupo como predictor principal y 'Rendimiento_Pre' como covariable normalizadora de la diferencia posttest. La homogeneidad de las pendientes fue evaluada a partir de la interacción de Condición $\times$ 'Rendimiento_Pre' y se evaluaron residuos, homocedasticidad (valores de heterocedasticidad importante) y desviaciones relevantes de normalidad, especificación funcional y observaciones influyentes. Se informaron R^2 , R^2 ajustado, estadístico F^* , coeficientes, errores estándar, IC95 % y valores p^* , así como los coeficientes estandarizados cuando los hubo.

El análisis de sensibilidad estimó un modelo ajustado por las variables preintervención sociodemográficas, tecnológicas y académicas que resultaron conceptualmente relevantes. Se consideró la multicolinealidad a partir del VIF. Las variables conceptualmente muy unidas — edad, grado, nivel educativo. — no se incluyeron de una sola vez de manera automática, y el modelo principal no se ajustó por las variables posteriores a la exposición, como la asistencia en la intervención.

Al considerar el agrupamiento de los estudiantes en aulas naturales, se estimó el coeficiente de correlación intraclase (ICC) y se ajustaron modelos con intercepto aleatorio por aula. Asimismo, como análisis complementario de robustez, se estimó el modelo principal con errores estándar robustos agrupados por 'ID_Aula'. La interpretación tuvo en cuenta la distribución asimétrica de la cantidad de observaciones por conglomerado y la existencia de aulas con una escasa/o baja/o frecuencia de registro y recursos/canales. A suma a lo anterior, la estructura institucional fue empleada también para describir la organización de los registros.

No se planificaron comparaciones EGB-Bachillerato ni interacciones nivel educativo $\times$ IA como análisis confirmatorio, ya que no es necesario para responder al objetivo de investigación y se evitaron análisis comparativos seleccionados sólo por el hallazgo de significación estadística.

Los análisis del cuestionario se interpretaron solamente como postintervención. Se adoptó un nivel bilateral de significatividad $\alpha = .05$ y las inferencias se acompañaron de IC95 % y tamaños del efecto, priorizando la magnitud y la precisión de las estimaciones frente a la significación estadística aislada.

El estudio tuvo la autorización de las instituciones educativas en las que se llevó a cabo el estudio. Para el alumnado menor de edad se recogió el consentimiento informado por parte de los padres o representantes legales y el asentimiento de los propios estudiantes. La participación fue voluntaria y la información se trató con carácter confidencial. Se los registros fueron codificados y anonimizados sin ningún dato que evidenciara la identificación totales de los estudiantes o las instituciones y fueron creados exclusivamente con fines académicos y de investigación. No existe documentación que declare la aprobación formal de algún comité de ética por lo que esta aprobación no se considera válida por el estudio.

RESULTADOS

En nuestra muestra analítica podemos encontrar 420 estudiantes, de los cuales 242 (57.6 %) eran pertenecientes a la condición de intervención didáctica que se estructuró mediante la inteligencia artificial (IA) y 178 (42.4 %) a la condición de comparación. Procedían de 64 instituciones educativas que se encontraban distribuidas en 286 aulas o paralelos identificados.

Dada la no aleatorización de la asignación, examinamos el balance de las características preintervención a través de las diferencias estandarizadas. Considerando el rendimiento académico basal, constatamos que

éste era parecido entre condiciones: $M = 66,33$ ($DE = 8,03$) en comparación y $M = 66,69$ ($DE = 7,91$) en intervención, con $SMD = 0,045$. Pequeñas diferencias aproximadamente para edad, sexo, acceso digital, uso previo de IA, asignatura y nivel educativo. Las mayores diferencias se podían vislumbrar para algunas categorías de provincia, en especial Tungurahua ($SMD = -0,231$), Pichincha ($SMD = 0,161$) y Manabí ($SMD = -0,159$) (Tabla 1).

Tabla 1
Balance basal de las condiciones de estudio

Característica	Comparación (n = 178)	Intervención IA (n = 242)	SMD
Rendimiento pretest, M (DE)	66.33 (8.03)	66.69 (7.91)	0.045
Edad, M (DE)	14.48 (2.69)	14.55 (2.76)	0.026
Sexo, n (%)			
Femenino	103 (57.9 %)	132 (54.5 %)	-0.067
Masculino	75 (42.1 %)	108 (44.6 %)	0.050
Prefiere no indicar	0 (0.0 %)	2 (0.8 %)	0.129
Acceso digital, n (%)			
Alto	47 (26.4 %)	79 (32.6 %)	0.137
Medio	97 (54.5 %)	129 (53.3 %)	-0.024
Bajo	34 (19.1 %)	34 (14.0 %)	-0.136
Uso previo de IA, n (%)			
Frecuente	29 (16.3 %)	46 (19.0 %)	0.071
Semanal	61 (34.3 %)	70 (28.9 %)	-0.115
Mensual	32 (18.0 %)	46 (19.0 %)	0.027
Ocasional	40 (22.5 %)	46 (19.0 %)	-0.085
Nunca	16 (9.0 %)	34 (14.0 %)	0.159
Asignatura, n (%)			
Ciencias Naturales	36 (20.2 %)	49 (20.2 %)	0.001
Estudios Sociales	38 (21.3 %)	58 (24.0 %)	0.063
Inglés	29 (16.3 %)	43 (17.8 %)	0.039
Lengua y Literatura	35 (19.7 %)	42 (17.4 %)	-0.059
Matemática	40 (22.5 %)	50 (20.7 %)	-0.044
Nivel educativo, n (%)			
Bachillerato	74 (41.6 %)	109 (45.0 %)	0.070
EGB Media	47 (26.4 %)	65 (26.9 %)	0.010
EGB Superior	57 (32.0 %)	68 (28.1 %)	-0.086

Característica	Comparación (n = 178) Intervención IA (n = 242) SMD		
Provincia, n (%)			
Azuay	17 (9.6 %)	28 (11.6 %)	0.066
Esmeraldas	14 (7.9 %)	22 (9.1 %)	0.044
Guayas	37 (20.8 %)	48 (19.8 %)	-0.024
Loja	14 (7.9 %)	22 (9.1 %)	0.044
Los Ríos	17 (9.6 %)	26 (10.7 %)	0.040
Manabí	27 (15.2 %)	24 (9.9 %)	-0.159
Pichincha	33 (18.5 %)	61 (25.2 %)	0.161
Tungurahua	19 (10.7 %)	11 (4.5 %)	-0.231

Nota. M = media; DE = desviación estándar; SMD = diferencia estandarizada entre las condiciones de intervención y comparación. Para las variables categóricas, la SMD se calculó por categoría.

En la muestra total, el rendimiento aumentó de 66.54 puntos (DE = 7.95) en el pretest a 70.06 (DE = 9.19) en el posttest, con un incremento medio de 3.53 puntos (DE = 4.55), IC95 % [3.09, 3.96], $t(419) = 15.88$, $p < .001$, $dz = 0.77$.

El incremento también fue estadísticamente significativo dentro de ambas condiciones. En comparación fue de 3.65 puntos (DE = 4.60), IC95 % [2.97, 4.33], $t(177) = 10.59$, $p < .001$, $dz = 0.79$; en intervención fue de 3.44 puntos (DE = 4.53), IC95 % [2.87, 4.01], $t(241) = 11.82$, $p < .001$, $dz = 0.76$ (Tabla 2). Estos resultados representan cambios temporales intragrupo y no estiman por sí mismos una diferencia atribuible a la intervención.

Tabla 2

Rendimiento académico pretest, posttest y cambio según condición

Condición	n	Cambio M (DE)	IC95 %	t(gl)	p	dz
Comparación	178	3.65 (4.60)	[2.97, 4.33]	10.59 (177)	< .001	0.79
Intervención IA	242	3.44 (4.53)	[2.87, 4.01]	11.82 (241)	< .001	0.76
Total	420	3.53 (4.55)	[3.09, 3.96]	15.88 (419)	< .001	0.77

Nota. ΔM = Rendimiento_Post – Rendimiento_Pre. dz representa el tamaño del efecto estandarizado del cambio dentro de cada condición y no el efecto comparativo de la intervención.

La diferencia de las ganancias entre intervención y comparación fue muy pequeña y no estadísticamente significativa: -0.21, IC95 % [-1.10, 0.68], $t(378.01) = -0.47$, $p = .637$, Hedges $g = -0.046$. De este modo, aun reconociendo cambios entre las condiciones en los periodos pretest-posttest, no se dio evidencia de la presencia de una diferencia entre ellas en la magnitud de los cambios de rendimiento.

El modelo principal analizado tuvo como variable dependiente el rendimiento posttest, como predictor condición y como covariable rendimiento pretest, y resultó estadísticamente significativo, $F(2, 417) = 641.20$, $p < .001$, $R^2 = .755$ y R^2 ajustado = .753.

El rendimiento basal se asoció positivamente al rendimiento posttest, $B = 1.004$, $EE = 0.028$, $\beta = .869$, IC95 % [0.949, 1.059], $p < .001$. Una vez controlado el rendimiento pretest, la condición intervención no se asoció estadísticamente a una peor o mejor puntuación posttest, $B = -0.210$, $EE = 0.451$ poder estadístico = $-.011$, IC95 % [-1.096, 0.676], $p = .641$, $\eta^2 p = .001$ (Tabla 3).

Las medias posttest ajustadas al rendimiento basal medio fueron 70.18 puntos (IC95 % [69.51, 70.86]) para comparación y 69.97 (IC95 % [69.40, 70.55]) para intervención. La interacción Condición $\times$ Rendimiento_Pre tampoco fue estadísticamente significativa, $B = 0.015$, $EE = 0.057$, $p = .791$, sin evidencia de pendientes diferentes entre condiciones.

Tabla 3

Modelo principal de rendimiento académico posttest ajustado por pretest

Predictor	B	EE	β	IC95 %	t	p
Intercepto	3.404	1.890	—	[-0.312, 7.120]	1.80	.072
Intervención IA	-0.210	0.451	-.011	[-1.096, 0.676]	-0.47	.641
Rendimiento pretest	1.004	0.028	.869	[0.949, 1.059]	35.81	< .001

Nota. Comparación = categoría de referencia; B = coeficiente no estandarizado; EE = error estándar; β = coeficiente estandarizado; IC95 % = intervalo de confianza del 95 %. $N = 420$; $R^2 = .755$; R^2 ajustado = .753; $\eta^2 p$ para la condición = .001.

La sensibilidad del análisis fue el rendimiento basal, así como la edad, el sexo, el acceso digital, el uso anterior de IA o la asignatura, logrando un modelo con $R^2 = .773$ y R^2 ajustado = .765, manteniéndose una estimación pequeña y no estadísticamente significativa en la condición de intervención, $B = -0.273$, $EE = 0.446$, IC95 % [-1.150, 0.603], $p = .540$. El VIF máximo fue 1.92, sin evidenciar multicolinealidad problemáticas. Por lo tanto, estos resultados llevaron a mantener la conclusión extraída del modelo principal.

Los 420 estudiantes se encontraban distribuidos en 286 aulas o paralelos poniendo a disposición el modelo nulo con intercepto aleatorio por aula. Este modelo estimó un ICC $\approx .092$, evidenciando variabilidad del rendimiento entre aulas antes de ajuste. Una vez se incorporaron el rendimiento basal y la condición, la varianza atribuible al intercepto aleatorio del aula apenas permanecía en los límites inferiores de la estimación.

El modelo principal utilizando errores agrupados en aulas robustos generó una estimación también similar en la condición de intervención, $B = -0.210$, EE robusto = 0.440, IC95 % [-1.075, 0.656], $p = .633$, así, que considerar la dependencia intra-aula en nada modificó la conclusión principal. La estimación del componente aleatorio ha de tomarse con cautela, y con la reducción de n° de alumnos anotando en mule conglomerados.

El componente perceptivo complementa los resultados de rendimiento académico, aunque requiere una interpretación prudente. Las ocho dimensiones del cuestionario presentaron puntuaciones medias cercanas al punto central de la escala y una consistencia interna moderada y variable, con coeficientes alfa de Cronbach entre .601 y .720. Estos resultados aportan una caracterización exploratoria de las percepciones y experiencias declaradas por los estudiantes respecto al uso educativo de la IA, pero no deben interpretarse

como evidencia procedente de una escala previamente validada. Además, al haberse administrado el cuestionario únicamente después de la intervención, no existe una medición basal que permita determinar cambios en motivación, autonomía, pensamiento crítico, aprendizaje percibido, valor didáctico, uso ético o riesgo de dependencia. Por consiguiente, estos resultados complementan descriptivamente el análisis del rendimiento académico, pero no modifican ni constituyen evidencia independiente de la efectividad de la intervención.

DISCUSIÓN

La principal meta fue examinar las diferencias en el rendimiento académico relacionadas con una intervención de integración didáctica de inteligencia artificial (IA), mediante mediciones pretest-postest y grupos naturales no equivalentes. El resultado significativo fue coincidente entre los análisis: el rendimiento aumentó entre ambas mediciones, tanto en la intervención como en la comparación, pero el tamaño de este cambio fue prácticamente equivalente. En el análisis de la intervención controlando por rendimiento basal, la condición de intervención tampoco mostró una asociación estadísticamente significativa con el rendimiento postest y el tamaño del efecto fue prácticamente nulo. Por lo tanto, no se encontró indicios de una ventaja académica adicional detectable de la intervención planificada mediante IA en comparación con la enseñanza habitual durante el periodo de estudio.

Esta distinción entre cambio temporal y diferencia entre condiciones es vital. Los aumentos intragrupo estadísticamente significativos describen la evolución dentro de cada condición experimentada pero no estiman el efecto diferencial de la intervención. Los resultados de la comparación de las ganancias resultaron en un tamaño del efecto prácticamente nulo, razón por la cual podríamos afirmar que atribuir esos resultados únicamente a la IA sería incompatible con lo que finalmente está recogiendo la evidencia, puesto que el grupo de estudiantes que continuaron con la metodología tradicional lograron una evolución igualmente adecuada.

Este dato es parcialmente discordante con las investigaciones que sí han llegado a encontrar ventajas asociadas a la intervención con IA generativa. Por ejemplo, en estudiantes de séptimo grado, la intervención de IA generativa y el aprendizaje basado en artes en el contenido de la historia se traducían en un aprendizaje y una motivación mejores respecto de la enseñanza tradicional (Chen et al., 2025). De la misma manera, un metaanálisis de un total de 37 estudios refleja un efecto positivo moderado de ChatGPT sobre el logro académico de los alumnos, aunque modificándose según la disciplina y su duración, el tipo de conocimiento a trabajar y el tipo de metodología utilizada (Liu et al., 2025). La discrepancia no implica necesariamente contradicción, puesto que la evidencia indica que los resultados de la IA educativa son heterogéneos y que están sometidos a las condiciones de la implementación. Una interpretación plausible remite, en parte, a la naturaleza de la intervención. Las actividades con IA se implementaron aproximadamente una vez por semana durante 12 semanas e incluyeron la explicación de conceptos, preguntas guiadas, verificación de que el contenido del estudiante era correcto, la generación de ejemplos, la retroalimentación entre todos, el contraste, entre otros procesos pedagógicos. Sin embargo, la implementación no fue individual uniforme y tampoco se dispuso de una medida concreta de fidelidad o dosis administrada. Aunado a ello, existía la posibilidad que el alumnado del grupo de comparación utilizara espontáneamente ChatGPT, Gemini o cualquier otra herramienta de IA ajena al programa. La comparación fue, por tanto, entre integración

pedagógica estructurada e integración pedagógica habitual, no entre notable presencia/ausencia de IA. Esta posible contaminación podría haber contribuido a la reducción en la separación entre las condiciones.

La importancia de la implementación también puede observarse desde experiencias en la que se combinan diferentes componentes pedagógicos donde el aprendizaje cooperativo ha sido tratado en relación con las habilidades y motivación dentro de experiencias estructuradas con alumnado secundario (Aráuz Cerezo et al., 2026). Estas aproximaciones demuestran que la tecnología es tan solo un componente de intervenciones educativas más extensas y que conviene distinguir disponibilidad tecnológica de uso de IA e integración didáctica.

Los resultados evidencian el carácter significativo del rendimiento basal del alumnado. El pretest presenta una asociación fuerte con el postest y explica una gran proporción de la variabilidad del modelo. No obstante lo anterior, la elevado poder explicativo total debe diferenciarse del ofrecido por la intervención: el rendimiento previo fue un fuerte predictor mientras que el coeficiente de la condición fue prácticamente nulo. Es por esta razón que la significación total del modelo, acompañada de un elevado R^2 , no constituyen en sí mismas evidencia de la efectividad de la IA.

La estabilidad del resultado tras el ajuste por las características preintervención apoya esta interpretación. No obstante el ajuste por edad, sexo, acceso digital, uso previo de IA y asignatura específica, la estimación de la condición continúa estando cercana a cero. Esto es relevante porque los grupos no eran aleatorios y ya había algunos desequilibrios basales, pero el ajuste estadístico no convierte un diseño cuasiexperimental en aleatorio ni elimina la confusión residual. Los análisis ajustados mejoran la comparación, pero no permiten hacer inferencias causales.

Que no se detecte una diferencia, que no se observe una diferencia detectable, no significa que la IA no tenga valor. En adolescentes, el uso y la utilidad percibida de herramientas generativas presentan relaciones heterogéneas con las características académicas e individuales (Klarin et al., 2024) y en contextos escolares se sugiere que la implicación del docente/a y la estructuración del uso de chatbots pueden tener que ver con la motivación y la experiencia de aprendizaje (Chiu et al., 2024). Por tanto, este resultado debe ser restringido a las condiciones, duración, actividades y población que se estudiaron, y no generalizarlo a toda la diversidad de modalidades de integración educativa de la IA. La innovación pedagógica no solo requiere tecnologías, sino también condiciones organizativas favorables al respecto. Consideramos el liderazgo entendido desde la perspectiva del acompañamiento y motivación del docente como una de las características de la transformación de las prácticas escolares (Troya et al., 2024), por lo que la incorporación de la IA presenta la necesidad de que su herramienta y la forma de organizar su implementación, así como el profesor se contemplen entrelazadamente en la propuesta.

En cuanto a las limitaciones, el muestreo no probabilístico por conveniencia del que se parte y los grupos naturales no aleatorizados restringen la poblacionalidad y la inferencia causal asociada. Además, a pesar de que expresamos el rendimiento en una escala común que va de 0 a 100, las evaluaciones se correspondían con grados, materias e instituciones diferentes; y a pesar de que la escala es útil para el análisis del cambio a nivel individual no aporta equidad psicométrica entre contextos. Así como tampoco se obtuvo una medida a nivel individual suficientemente fidedigna o intensa respecto de la exposición a la intervención. La estructura de agrupamiento también fue analizada tomando como base el aula o paralelo real de pertenencia. El modelo nulo encontró un ICC del aula menor, del orden de .092. Al incorporar el rendimiento basal y

aplicar la condición de estudio, la variabilidad residual que puede ser atribuida al aula se aproximó al límite inferior de estimación. Los errores estándar robustos agrupados por aula tampoco produjeron un cambio en la conclusión asociada a la condición de intervención. Estos resultados evidencian que explicitando esta dependencia intraaula no encontramos la evidencia de una ventaja académica adicional a la que da cuenta la intervención. No obstante, la necesidad de considerar el gran número de conglomerados con pocos estudiantes registrados nos lleva a matizar los resultados de la estimación del componente aleatorio.

El componente perceptivo complementa los resultados de rendimiento académico, aunque requiere una interpretación prudente. Las ocho dimensiones del cuestionario presentaron puntuaciones medias cercanas al punto central de la escala y una consistencia interna moderada y variable, con coeficientes alfa de Cronbach entre .601 y .720. Estos resultados aportan una caracterización exploratoria de las percepciones y experiencias declaradas por los estudiantes respecto al uso educativo de la IA, pero no deben interpretarse como evidencia procedente de una escala previamente validada. Además, al haberse administrado el cuestionario únicamente después de la intervención, no existe una medición basal que permita determinar cambios en motivación, autonomía, pensamiento crítico, aprendizaje percibido, valor didáctico, uso ético o riesgo de dependencia. Por consiguiente, estos resultados complementan descriptivamente el análisis del rendimiento académico, pero no modifican ni constituyen evidencia independiente de la efectividad de la intervención.

La utilización planificada de IA también merece una reflexión. La analítica del aprendizaje y la IA responsable han sido planteadas como materiales para apoyar la toma de decisiones pedagógicas fundamentadas en la información (Bernal Parraga, 2026), y el acompañamiento educativo es una parte de las respuestas institucionales que emergen ante las diferentes problemáticas del alumnado (Tello Mayorga et al., 2025); estos antecedentes son contextuales pero no son prueba de un efecto de la intervención analizada, cuyos resultados no mostraron una ventaja diferencial asociada a la condición de IA.

A la luz de los aspectos pedagógicos, los hallazgos evidencian que la inclusión de IA no es per se un suficiente aumento del rendimiento de la enseñanza habitual; debe responderse a unos objetivos curriculares y hay de incluir la orientación del profesorado y promover la verificación, el contraste de las respuestas generadas utilizando la tecnología para apoyar y no para sustituir el razonamiento del alumnado. La evidencia escolar muestra que los resultados pueden depender de cómo se organiza dicha inclusión (Chen et al., 2025; Chiu et al., 2024). Futuras investigaciones deberán estructurar las actividades, registrar la intensidad real en la que se produjo la exposición en el uso del exterior de la IA y utilizar diseños que logren discriminar el desarrollo académico habitual del añadido adicional que pueda tener la intervención tecnológica.

CONCLUSIONES

El estudio permitió comprobar las diferencias de la intervención de integración didáctica de inteligencia artificial estructurada sobre el rendimiento académico a partir de los resultados de grupos naturales no equivalentes a partir de pretest-posttest. Los resultados correspondientes al incremento del rendimiento académico entre ambas mediciones en condición intervención y en condición comparación resultaron ser estadísticamente significativos pero los resultados intragrupo tampoco suponen evidencias de la intervención dada la similitud de la magnitud de la ganancia entre ambas condiciones. La comparación de

las ganancias nos muestra resultados con ausencia de diferencia significativa entre intervención y comparación y un tamaño del efecto muy reducido. De la misma forma, el modelo principal ajustado para rendimiento basal mostró ausencia de asociación entre ser de la condición intervención y el rendimiento académico de posttest. En consecuencia, dentro de las condiciones específicas sometidas a evaluación el modelo de didáctica estructurada de IA no se relacionó con un rendimiento académico superior con la práctica docente habitual. El rendimiento académico propio estuvo en la base de la predicción del rendimiento académico posterior. El modelo tiene un elevado poder explicativo global, no obstante, no se debe a la intervención: es fruto de que el rendimiento basal y el rendimiento posttest estaban fuertemente correlacionados. En el análisis de sensibilidad, las covariables recogidas antes de la intervención no alteraron la estimación asociada a la condición de estudio, manteniéndose así aquella consistente y estable tras los ajustes realizados. El resultado tiene que ser entendido en los límites del diseño. El uso de los grupos naturales sin asignación aleatoria y hacer un muestreo no probabilístico por conveniencia limitan la inferencia causal hasta la generalización de los resultados. Por otro lado, aunque las evaluaciones académicas se expresaron utilizando una misma escala, las de las diversas pruebas llevaron diferentes contextos aislados (curriculares) de ahí que no debiera asumirse una equivalencia psicométrica absoluta entre ellos. La forma de dependencia intraaula o paralelo pudo ser examinada a partir de los identificadores con los que se dispone en este caso. La consideración del agrupamiento intraaula ya sea a partir del ICC, ya sea a partir de errores estándar robustos por conglomerado, no alteraban la conclusión general principal. De todos modos, la distribución desigual del número de estudiantes que quedaban asignados por aula sí era un aspecto a considerar a la hora de interpretar la estimación correspondiente a esa estructura jerárquica. Las dimensiones postintervención del cuestionario permitieron complementar el análisis mediante una caracterización exploratoria de las percepciones y experiencias declaradas por los estudiantes respecto al uso educativo de la IA. En términos generales, las puntuaciones se situaron en niveles medios y la consistencia interna fue moderada y variable entre dimensiones. No obstante, debido al carácter ad hoc del instrumento, la ausencia de validación psicométrica previa formal y la existencia de una única medición postintervención, estos resultados no permiten establecer cambios respecto de una situación basal ni atribuir las percepciones observadas a la intervención. Por ello, deben considerarse evidencia descriptiva complementaria y no una demostración de efectividad de la integración didáctica de IA.

Desde una perspectiva pedagógica, los resultados indican que la incorporación estructurada de la IA no se debe considerar como una verdad absoluta para la obtención de resultados académicos superiores ante la enseñanza habitual. En futuros trabajos, se debería conservar el identificador al aula, tomar el registro al alumnado de una forma más encubierta al trabajo académico, aumentar la comparabilidad de las mediciones académicas, realizar una buena documentación de la codificación y de las propiedades del instrumento perceptivo y, en fin, la implementación de diseños experimentales que permitieran distinguir el cambio temporal de la diferencia que se encuentra asociada específicamente a la integración didáctica de la IA.

REFERENCIAS BIBLIOGRÁFICAS

- Adisa, I. O., & Adefisayo, A. O. (2025). Middle school students' perspectives on adopting generative AI in K-12 education. *The Journal of Educational Research*, 118*(6), 724–735.
<https://doi.org/10.1080/00220671.2025.2510393>

- Aguilar Tinoco, R. J., Carvallo Lobato, M. F., Román Camacho, D. E., Liberio Anzules, A. M., Hernández Centeno, J. A., Duran Fajardo, T. B., & Bernal Parraga, A. P. (2024). El Impacto del Diseño Universal para el Aprendizaje (DUA) en la Enseñanza de Ciencias Naturales: Un Enfoque Inclusivo y Personalizado. *Ciencia Latina Revista Científica Multidisciplinar*, 8(5), 2162-2178. https://doi.org/10.37811/cl_rcm.v8i5.13682
- Ali, D., Fatemi, Y., Boskabadi, E., Nikfar, M., Ugwuoke, J., & Ali, H. (2024). ChatGPT in teaching and learning: A systematic review. **Education Sciences*, 14*(6), 643. <https://doi.org/10.3390/educsci14060643>
- Ali, F., Choy, D., Divaharan, S., Tay, H. Y., & Chen, W. (2023). Supporting self-directed learning and self-assessment using TeacherGAIA, a generative AI chatbot application: Learning approaches and prompt engineering. **Learning: Research and Practice*, 9*(2), 135–147. <https://doi.org/10.1080/23735082.2023.2258886>
- Alvarez Piza, R. A., Del Hierro Pérez, M. C., Vera Molina, R. M., Moran Piguave, G. D., Pareja Mancilla, S. S., Narváez Hoyos, J. J., & Bernal Parraga, A. P. (2024). Desarrollo del Pensamiento Lógico a través de la Resolución de Problemas en Matemáticas Estrategias Eficaces para la Educación Básica. *Ciencia Latina Revista Científica Multidisciplinar*, 8(5), 2212-2229. https://doi.org/10.37811/cl_rcm.v8i5.13686
- Aráuz Cerezo, E., Tuarez Cobeña, M., Bernal Parraga, A., Yépez Yépez, L., & Loo Quintero, J. (2026). Efecto del aprendizaje cooperativo en el desarrollo de habilidades técnico-tácticas y la motivación en estudiantes de Educación Física durante la enseñanza del baloncesto en secundaria. *Polo del Conocimiento*, 11(4), 2266-2283. doi:<https://doi.org/10.23857/pc.v11i4.11541>
- Ballestra Caffaratti, L., Longobardi, C., Badenes-Ribera, L., & Marengo, D. (2025). AI adoption among adolescents in education: Extending the UTAUT2 with psychological and contextual factors. **Frontiers in Artificial Intelligence*, 8*, 1614993. <https://doi.org/10.3389/frai.2025.1614993>
- Bernal Párraga, A. P., Constante Olmedo, D. F., López Sánchez, I. Y., Padilla Portocarrero, D. K., & Duarte Salinas, E. M. (2026). Ecosistema híbrido inteligente para la enseñanza de Ciencias Naturales: un modelo integrador de metodologías activas, IA y regulación emocional. *Tesla Revista Científica*, 6(1). <https://doi.org/10.55204/trc.v6i1.e588>
- Bernal Parraga, A. P., Orozco Maldonado, M. E., Salinas Rivera, I. K., Gaibor Davila, A. E., Gaibor Davila, V. M., Gaibor Davila, R. S., & Garcia Monar, K. R. (2024). Análisis de Recursos Digitales para el Aprendizaje en Línea para el Área de Ciencias Naturales. *Ciencia Latina Revista Científica Multidisciplinar*, 8(4), 9921-9938. https://doi.org/10.37811/cl_rcm.v8i4.13141
- Bernal Parraga, A. P., Alvarado Cantos, T. M., Lucas Farías, M. S., Palacios Palma, P. P., & Demera Moreira, M. V. (2026). Aprendizaje basado en la indagación y laboratorios virtuales para fortalecer las competencias científicas en Biología del Bachillerato. *Prometeo Conocimiento Científico*, 6(2), e147. <https://doi.org/10.55204/pcc.v6i2.e147>
- Bernal Párraga, A. P., Medina Marino, P. A., Cholango Tenemaza, E. G., Zamora Franco, A. F., Zamora Franco, C. G., & López Sánchez, I. Y. (2024). Educación especial en metodologías de discapacidad múltiple intelectual y física: Un enfoque inclusivo. *Ciencia Latina Revista Científica Multidisciplinar*, 8(3), 3229-3248. https://doi.org/10.37811/cl_rcm.v8i3.11544

- Bernal Parraga, A. P., Rodríguez Cajas, G. A., & Ramírez Campoverde, E. A. (2026). Impacto de las metodologías activas y tecnologías emergentes en el desarrollo de la comprensión lectora y la escritura creativa en Lengua y Literatura en educación básica. *Educaf5-Berit*, 1(01), 147–157. <https://doi.org/10.5281/zenodo.18210399>
- Bernal Párraga, A. P., Toapanta Guanoquiza, M. J., Martínez Oviedo, M. Y., Correa Pardo, J. A., Ortiz Rosillo, A., Guerra Altamirano, I. del C., & Molina Ayala, R. E. (2024). Aprendizaje Basado en Role-Playing: Fomentando la Creatividad y el Pensamiento Crítico desde Temprana Edad. *Ciencia Latina Revista Científica Multidisciplinar*, 8(4), 1437-1461. https://doi.org/10.37811/cl_rcm.v8i4.12389
- Caratiquit, K. D., & Caratiquit, L. J. C. (2023). ChatGPT as an academic support tool on the academic performance among students: The mediating role of learning motivation. *Journal of Social, Humanity, and Education*, 4(1), 21–33. <https://doi.org/10.35912/jshe.v4i1.1558>
- Chen, J., Mokmin, N. A. M., & Qi, S. (2025). Generative AI-powered arts-based learning in middle school history: Impact on achievement, motivation, and cognitive load. **The Journal of Educational Research*, 118*(6), 688–700. <https://doi.org/10.1080/00220671.2025.2510395>
- Chicaiza, V. (2025). Inteligencia artificial y aprendizaje de idiomas: Personalización del aula de inglés a través de plataformas adaptativas. **Revista Veritas de Difusão Científica*, 6*(2), 477–506. <https://revistaveritas.org/index.php/veritas/article/view/643>
- Chiu, T. K. F., Moorhouse, B. L., Chai, C. S., & Ismailov, M. (2024). Teacher support and student motivation to learn with artificial intelligence (AI) based chatbot. *Interactive Learning Environments*, 32(7), 3240–3256. <https://doi.org/10.1080/10494820.2023.2172044>
- Cochancela Pérez, N., Bernal Parraga, A., Romero Torres, T., & Morales Loayza, D. (2026). Impacto de las estrategias innovadoras sobre la comprensión lectora y la producción textual en estudiantes de educación básica elemental media y superior. *Polo del Conocimiento*, 11(7), 2447-2471. doi:<https://doi.org/10.23857/pc.v11i7.12154>
- Gan, W., Ouyang, J., Li, H., Xue, Z., Zhang, Y., Dong, Q., Huang, J., Zheng, X., & Zhang, Y. (2024). Integrating ChatGPT in orthopedic education for medical undergraduates: Randomized controlled trial. **Journal of Medical Internet Research*, 26*, e57037. <https://doi.org/10.2196/57037>
- Guerrero Carrera, L. M., Bernal Parraga, A. P., Ordóñez Quitizaca, N. K., Toapanta Guanoquiza, M. J., Cabrera Brown, M. N., Álvarez León, D. S., & Yanchapaxi Oña, K. G. (2024). Efectividad de Metodologías Activas Innovadoras de Aprendizaje en el Área de Lengua. *Ciencia Latina Revista Científica Multidisciplinar*, 8(3), 9213-9244. https://doi.org/10.37811/cl_rcm.v8i3.12073
- 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/10.3390/nursrep15020068>
- Jara Chiriboga, S. P., Troncoso Burgos, A. L., Ruiz Avila, M. M., Cosquillo Chida, J. L., Aldas Macias, K. J., Castro Morante, Y. E., & Bernal Párraga, A. P. (2025). Inteligencia Artificial y Aprendizaje Personalizado en Lenguas Extranjeras: Un Análisis de los Chatbots y los Asistentes Virtuales en Educación. *Revista Científica De Salud Y Desarrollo Humano*, 6(1), 882–905. <https://doi.org/10.61368/r.s.d.h.v6i1.515>

- Klarin, J., Hoff, E., Larsson, A., & Daukantaitė, D. (2024). Adolescents' use and perceived usefulness of generative AI for schoolwork: Exploring their relationships with executive functioning and academic achievement. *Frontiers in Artificial Intelligence, 7*, 1415782. <https://doi.org/10.3389/frai.2024.1415782>
- Lee, C.-C., & Low, M. Y. H. (2024). Using GenAI in education: The case for critical thinking. *Frontiers in Artificial Intelligence, 7*, 1452131. <https://doi.org/10.3389/frai.2024.1452131>
- León Ruíz, M. E., Bernal Párraga, A. P., Bustamante Peñaherrera, G. S., Yanza Rojas, C. J., Guzmán Quiña, M. de los A., Davila Amari, M. A., & López Villacis, D. E. (2024). Enfoques Pedagógicos para la Enseñanza de Estudios Sociales en Libros de Texto de Educación Básica Superior. *Ciencia Latina Revista Científica Multidisciplinar*, 8(4), 9132-9152. https://doi.org/10.37811/cl_rcm.v8i4.13060
- Liu, Z., Zuo, H., & Lu, Y. (2025). The impact of ChatGPT on students' academic achievement: A meta-analysis. *Journal of Computer Assisted Learning*, 41(4), e70096. <https://doi.org/10.1111/jcal.70096>
- Mena Castillo, M. L., Iza Maigua, A. C., Bernal Parraga, A. P., & Lalangui Sandoval, J. A. (2026). Neurociencia de la lectura e inteligencia artificial: comparación del método fonológico y global en estudiantes con dislexia temprana. *Tesla Revista Científica*, 6(2), e696. <https://doi.org/10.55204/trc.v6i2.e696>
- Mora Villamar, F. M., Bernal Párraga, A. P., Molina Ayala, E. T., Salazar Veliz, E. T., Padilla Chicaiza, V. A., & Zambrano Lamilla, L. M. (2024). Innovaciones en la didáctica de la lengua y literatura: estrategias del siglo XXI. *Ciencia Latina Revista Científica Multidisciplinar*, 8(3), 3852-3879. https://doi.org/10.37811/cl_rcm.v8i3.11595
- Mortlock, R., & Lucas, C. (2024). Generative artificial intelligence (Gen-AI) in pharmacy education: Utilization and implications for academic integrity: A scoping review. *Exploratory Research in Clinical and Social Pharmacy, 15*, 100481. <https://doi.org/10.1016/j.rcsop.2024.100481>
- Ng, D. T. K., Tan, C. W., & Leung, J. K. L. (2024). Empowering student self-regulated learning and science education through ChatGPT: A pioneering pilot study. *British Journal of Educational Technology, 55*(4), 1328–1353. <https://doi.org/10.1111/bjet.13454>
- Parraga, A. P. B. (2026). Learning analytics and responsible artificial intelligence for data-driven pedagogical decision-making in basic education. *Luz*, 25, e1601–e1601. <https://luz.uho.edu.cu/index.php/luz/article/view/1601>
- Šedlbauer, J., Činčera, J., Slavík, M., & Hartlová, A. (2024). Students' reflections on their experience with ChatGPT. *Journal of Computer Assisted Learning, 40*(4), 1526–1534. <https://doi.org/10.1111/jcal.12967>
- Shi, J., Li, X., Ning, Y., Kong, W., Guo, W., Ma, T., Yang, N., Lu, J., Guo, Y., Liu, L., & Yang, C. (2025). Application of generative artificial intelligence chatbots + project task driven teaching in undergraduate nursing students: A quasi-experimental study. *BMC Medical Education, 25*(1), 1754. <https://doi.org/10.1186/s12909-025-08324-y>
- Shin, S., Sang, J., & You, J.-S. (2025). Use and perception of generative artificial intelligence in traditional Korean medicine education: A cross-sectional survey of undergraduate students in Korea. *Integrative Medicine Research, 14*(4), 101216. <https://doi.org/10.1016/j.imr.2025.101216>

- Simoni, J., Urtubia-Fernandez, J., Mengual, E., Simoni, D. A., Royo, M., Egaña-Yin, D., Hertog, O. L. A., López-Ortiz, L., Muñoz-Tomás, A., Santiago-Martínez, P., Vahamaki, A., & Pereira, J. L. (2025). Artificial intelligence in undergraduate medical education: An updated scoping review. **BMC Medical Education*, 25*(1), 1609. <https://doi.org/10.1186/s12909-025-08188-2>
- Tello Mayorga, L. E., Bowen Castro, D. J., Rosero Ubidia, A. M., Rea Elizalde, C. A., & Bernal Parraga, A. P. (2025). El papel del orientador educativo en la prevención del acoso escolar modelos de intervención basados en evidencia. *ASCE*, 4(3), 1089–1115. <https://doi.org/10.70577/ASCE/1089.1115/2025>
- Troya Santilán, B. N., García Sosa, S. M., Medina Marino, P. A., Campoverde Duran, V. D. R., & Bernal Párraga, A. P. (2024). Diseño e Implementación del Gamming Impulsados por IA para Mejorar el Aprendizaje. *Ciencia Latina Revista Científica Multidisciplinar*, 8(3), 4051-4071. https://doi.org/10.37811/cl_rcm.v8i3.11611
- Troya, B. N., Arzube, M. C., Arzube, D. M., Troya, C. M., Martínez, M. Y., Zapata, Y. F. y Bernal, A. P. (2024). Liderazgo Educativo Transformacional: Estrategias para Inspirar y Motivar a los Docentes en el Contexto Escolar. *Ciencia Latina Revista Científica Multidisciplinar*, 8(5), 2230-2246. https://doi.org/10.37811/cl_rcm.v8i5.13687
- Varma, J. R., Fernando, S., Ting, B. Y., Aamir, S., & Sivaprakasam, R. (2023). The global use of artificial intelligence in the undergraduate medical curriculum: A systematic review. **Cureus*, 15*(5), e39701. <https://doi.org/10.7759/cureus.39701>
- Xu, X., Chen, Y., & Miao, J. (2024). Opportunities, challenges, and future directions of large language models, including ChatGPT in medical education: A systematic scoping review. **Journal of Educational Evaluation for Health Professions*, 21*, 6. <https://doi.org/10.3352/jeehp.2024.21.6>
- Zeng, J., Sun, K., Qin, P., & Liu, S. (2025). Enhancing ophthalmology students' awareness of retinitis pigmentosa: Assessing the efficacy of ChatGPT in AI-assisted teaching of rare diseases—A quasi-experimental study. **Frontiers in Medicine*, 12*, 1534294. <https://doi.org/10.3389/fmed.2025.1534294>
- Zhang, P., & Tur, G. (2024). A systematic review of ChatGPT use in K-12 education. **European Journal of Education*, 59*(2), e12599. <https://doi.org/10.1111/ejed.12599>
- Zhao, Y., Yuan, Y., Wen, Z., Leng, L., Shi, L., Hu, X., Wei, X., Zuo, M., Mou, J., Luo, Q., Chen, M., Hu, R., & Gao, H. (2025). The current status, knowledge, attitudes, and challenges of generative artificial intelligence use among undergraduate nursing students: A single-center cross-sectional survey of western China. **Frontiers in Public Health*, 13*, 1648416. <https://doi.org/10.3389/fpubh.2025.1648416>

Conflicto de intereses:

Los autores declaran que no existe conflicto de interés