



INTELIGENCIA ARTIFICIAL Y NEUROEDUCACIÓN: IMPACTO EN LOS PROCESOS NEUROCOGNITIVOS DEL APRENDIZAJE ÉTICO

Anderson Fabian Algarra Pulido¹
ORCID: 0009-0006-6759-620X

¹ Doctorando en Educación, Universidad Americana de Europa (UNAE)

[Revista de Investigación Multidisciplinaria Iberoamericana, RIMI](#) © 2023 by Elizabeth Sánchez Vázquez is licensed under

RESUMEN

La integración de la inteligencia artificial (IA) en entornos educativos plantea interrogantes fundamentales sobre su incidencia en los procesos neurocognitivos. Este estudio busca equilibrar el entusiasmo tecnológico con un análisis riguroso de los riesgos y potencialidades para el desarrollo cerebral, dado que la literatura previa predomina en el plano funcional y técnico, dejando de lado el impacto ético y neurobiológico de la mediación algorítmica. El artículo analiza el impacto de la IA en las dimensiones afectivas, cognitivas y éticas del aprendizaje, articulando la neuroeducación contemporánea con desarrollos tecnológicos recientes. Se emplea una metodología de revisión teórica con enfoque hermenéutico, analizando una muestra de 25 fuentes principales —seleccionadas tras un proceso de filtrado de 31 documentos preseleccionados— publicadas preferentemente entre 2020 y 2025, consultadas en bases de datos como PubMed, Scopus, Dialnet y ERIC. Los resultados sugieren que, si bien la IA potencia la personalización del aprendizaje, existe un riesgo interpretativo de debilitamiento de los procesos de memoria, fragmentación atencional y una posible dependencia de sistemas externos que compromete la independencia deliberativa del estudiante. Se concluye que la mediación algorítmica requiere un marco de neuroética educativa que preserve la dimensión humana y el juicio crítico en el encuentro pedagógico.

Palabras clave: aprendizaje ético, inteligencia artificial, neuroeducación, neuroética, mediación algorítmica.

ABSTRACT

The integration of artificial intelligence (AI) in educational environments raises fundamental questions about its impact on neurocognitive processes. This study seeks to balance technological enthusiasm with a rigorous analysis of risks and potentialities for brain development, as previous literature predominantly focuses on functional and technical aspects, neglecting the ethical and neurobiological impact of algorithmic mediation. This article analyzes the impact of AI on the affective, cognitive, and ethical dimensions of learning, articulating contemporary neuroeducation with recent technological developments. A theoretical review with a hermeneutical approach is employed, analyzing a sample of 25 main sources —selected after filtering 31 preselected documents— published preferably between 2020 and 2025, retrieved from databases such as PubMed, Scopus, Dialnet, and ERIC. Results suggest that while AI enhances learning personalization, an interpretative risk of weakening memory processes, attentional fragmentation, and a possible dependence on external systems that could compromise student autonomy is identified. It is concluded that algorithmic mediation requires an educational neuroethics framework to preserve the human dimension and critical judgment in the pedagogical encounter.

Keywords: artificial intelligence, ethical learning, neuroeducation, neuroethics, algorithmic mediation.

INTRODUCCIÓN

La neuroeducación, como campo interdisciplinario que examina los mecanismos cerebrales para optimizar la pedagogía, enfrenta una transformación profunda ante la irrupción de la inteligencia artificial (IA) (Macías Acosta, Tomalá de la Cruz & Marqués Molias, 2025). Tradicionalmente, esta disciplina se centra en comprender la plasticidad sináptica y la modulación emocional como bases del aprendizaje significativo (Márquez, 2016); sin embargo, la incorporación de tutores inteligentes y algoritmos de computación afectiva introduce una complejidad técnica que demanda un reexamen crítico de la consolidación de valores en la era digital (Algarra Pulido, 2021).

Los estudios previos avanzan significativamente en la comprensión de la eficiencia de la IA para personalizar rutas de aprendizaje, adaptar contenidos al ritmo cognitivo del estudiante y potenciar la retención de información mediante sistemas de retroalimentación inmediata (Ouyang et al., 2023; UNESCO, 2024a). Asimismo, la literatura reciente destaca el potencial de la computación afectiva para regular estados emocionales y optimizar la atención en entornos digitales (Macías Acosta et al., 2025; Strielkowski et al., 2024). No obstante, estas investigaciones predominan en el plano funcional y técnico, dejando de lado una interrogante crucial: ¿qué impacto tiene la inteligencia artificial en los procesos neurocognitivos de la formación moral?

A pesar de la proliferación de herramientas tecnológicas, persiste un vacío en la comprensión sistémica de cómo la interacción con algoritmos afecta la educación en valores y el juicio moral desde una perspectiva neurobiológica. La literatura actual destaca avances en eficiencia y personalización, pero no profundiza suficientemente en la relación entre la intermediación tecnológica y la maduración de los procesos ejecutivos necesarios para la deliberación ética.

En consecuencia, aunque la literatura ha documentado ampliamente los beneficios funcionales de la IA en educación, permanece una insuficiente comprensión sobre sus implicaciones neuroéticas y cognitivas en el desarrollo del juicio moral. Esta brecha resulta particularmente relevante porque la maleabilidad cortical responde de manera bifurcada a las experiencias digitales: pueden fortalecer las conexiones mediante el compromiso activo o, por el contrario, generar déficits atencionales cuando la interacción es pasiva (Sethi & Jain, 2024). Por tanto, resulta imperativo equilibrar el entusiasmo tecnológico con un análisis riguroso de los riesgos y potencialidades para el desarrollo del cerebro humano.

En este contexto, resulta necesario precisar algunas distinciones conceptuales. El aprendizaje ético se define como el proceso deliberativo mediante el cual el estudiante desarrolla capacidades de juicio moral, distinguiéndose de la formación moral, que abarca una dimensión más amplia de socialización de valores, hábitos y prácticas comunitarias. Por su parte, la autonomía moral alude a la capacidad de deliberar y decidir con base en principios propios, mientras que la neuroética educativa designa el campo de estudio que examina las implicaciones éticas de la neurociencia y la tecnología en contextos pedagógicos. Esta distinción resulta pertinente porque la IA puede optimizar la adquisición de información normativa sin necesariamente promover la maduración de la corteza prefrontal asociada a la deliberación ética.

El aporte original de este trabajo consiste en articular de manera sistémica hallazgos neurocognitivos con imperativos éticos del encuentro pedagógico, proponiendo un marco analítico-interpretativo que oriente la integración de la IA sin sustituir la agencia del estudiante.

El objetivo general de este artículo es analizar el impacto de la inteligencia artificial en los procesos neurocognitivos del aprendizaje ético desde una perspectiva de neuroética educativa. Los objetivos específicos son: (1) identificar las transformaciones neurocognitivas asociadas al arbitraje digital en dimensiones afectiva, cognitiva y ética; (2) examinar la relación entre la personalización algorítmica y el autogobierno cognitivo del estudiante; y (3) proponer un marco analítico-interpretativo de neuroética

educativa que oriente la integración ética de la IA en contextos pedagógicos. La pregunta de investigación que orienta este estudio es: ¿qué impacto tiene la inteligencia artificial en los procesos neurocognitivos vinculados al aprendizaje ético y qué implicaciones neuroéticas emergen de dicha intermediación tecnológica?

ESTADO DEL ARTE

La literatura científica de los últimos cinco años documenta avances sustanciales en la integración de la IA y la neuroeducación, aunque con enfoques metodológicos y epistemológicos divergentes que revelan tensiones no resueltas en el campo.

Convergencias y tensiones en la personalización algorítmica. Macías Acosta et al. (2025) realizan una revisión sistemática sobre IA y neuroeducación en la enseñanza de las ciencias, reportando que los sistemas inteligentes potencian la adaptación curricular al ritmo cerebral del estudiante. Ouyang et al. (2023) integran modelos predictivos de IA con análisis de aprendizaje, demostrando que el enfoque combinado aumenta el compromiso estudiantil y mejora los resultados colaborativos. Strielkowski et al. (2024) analizan bibliométricamente 3.518 publicaciones sobre aprendizaje adaptativo e IA, concluyendo que estas tecnologías pueden reducir la duración del entrenamiento entre un 40% y un 60% mientras incrementan las tasas de retención hasta en un 50%. Sin embargo, estos estudios comparten una limitación metodológica común: priorizan variables de eficiencia y rendimiento (tiempo de respuesta, retención inmediata, satisfacción) sin incorporar indicadores de maduración neurocognitiva a largo plazo ni de desarrollo del juicio moral. Mientras Macías Acosta et al. (2025) y Ouyang et al. (2023) enfatizan el potencial adaptativo de la IA, Strielkowski et al. (2024) advierten que la generalización de estos hallazgos está condicionada por la heterogeneidad de los diseños experimentales revisados, lo cual introduce una tensión entre la promesa de eficiencia y la ausencia de evidencia longitudinal sobre efectos neurobiológicos.

Riesgos de dependencia tecnológica y brechas de acompañamiento. Fu y Weng (2024) evidencian que los estudiantes latinoamericanos adoptan la IA de forma autónoma ante la ausencia de acompañamiento docente, lo que incrementa los riesgos de dependencia tecnológica. Este hallazgo contrasta con los estudios anteriores en cuanto a enfoque: mientras Macías Acosta et al. (2025) y Ouyang et al. (2023) se centran en métricas de rendimiento académico, Fu y Weng (2024) incorporan dimensiones éticas y socioculturales. La contradicción central radica en que, si bien la IA puede optimizar resultados de aprendizaje en contextos controlados (Ouyang et al., 2023), su implementación desregulada en contextos con escasa mediación docente genera externalización de funciones cognitivas (Fu y Weng, 2024). Esta tensión sugiere que los beneficios neurocognitivos de la IA no son inherentes a la tecnología, sino dependientes de la calidad del acompañamiento pedagógico.

Neuroplasticidad y calidad del compromiso. Sethi y Jain (2024) exploran la intersección entre IA, aprendizaje socioemocional y desarrollo infantil, advirtiendo que la reconfiguración neuronal depende de la calidad del compromiso activo frente al pasivo. Este estudio resulta particularmente relevante porque introduce una distinción neurobiológica ausente en los análisis bibliométricos de Strielkowski et al. (2024): la plasticidad sináptica no responde uniformemente a la mediación tecnológica, sino que se bifurca según la naturaleza del compromiso del estudiante. No obstante, la limitación de Sethi y Jain (2024) radica en que sus inferencias sobre reconfiguración neuronal se basan principalmente en estudios de neuroimagen indirectos y correlatos conductuales, sin disponer de evidencia neurofisiológica directa derivada del uso de IA educativa.

Marcos normativos y vacíos neuroéticos. UNESCO (2024a, 2024b) establece marcos normativos de competencia docente y gobernanza ética para la IA en educación. Estos documentos institucionales aportan una base regulatoria indispensable, pero presentan una limitación conceptual: articulan principios éticos generales (equidad, privacidad, transparencia) sin vincularlos específicamente con los mecanismos neurocognitivos del aprendizaje ético. Por tanto, existe un vacío entre la gobernanza ética de la IA y la comprensión de cómo dicha gobernanza incide en la arquitectura cerebral del estudiante.

En conjunto, estos estudios evidencian un avance significativo en la comprensión funcional de la IA en educación; sin embargo, permanece insuficientemente explorada su relación con los procesos neurocognitivos vinculados al desarrollo del juicio moral y la independencia deliberativa. Las investigaciones revisadas presentan tres limitaciones compartidas: (a) predominio de diseños cuantitativos de eficiencia que dejan de lado la dimensión neuroética; (b) ausencia de estudios longitudinales que examinen efectos cerebrales a largo plazo; y (c) escasa articulación entre los hallazgos neuroeducativos y los imperativos éticos del encuentro pedagógico. Esta investigación se sitúa precisamente en esa brecha, articulando los hallazgos neurocognitivos con los imperativos éticos del encuentro pedagógico.

MARCO TEÓRICO

Procesos emocionales

Las emociones constituyen ejes fundamentales del aprendizaje que condicionan la memoria y la atención. En ambientes de confianza, la secreción de oxitocina, dopamina y serotonina facilita la neurogénesis y la motivación intrínseca (Vega, 2019). Por el contrario, el estrés activa la amígdala y libera cortisol, interfiriendo con la consolidación de información (Waipan, 2017).

La computación afectiva intenta replicar estos procesos mediante el monitoreo de expresiones faciales y señales fisiológicas para adaptar el contenido educativo (Ocaña, 2015; Macías Acosta et al., 2025). No obstante, esta mediación técnica plantea dudas sobre si la regulación emocional externa mediante IA puede limitar el desarrollo de la autorregulación autónoma del estudiante, esencial para la madurez moral.

Desde una perspectiva de formación moral, la autorregulación emocional no es solo una habilidad técnica, sino un componente del carácter moral que permite postergar la gratificación y sostener la atención en dilemas éticos complejos.

Neuroplasticidad

El aprendizaje significativo altera la arquitectura cerebral a través de la adaptabilidad sináptica. La interacción con la tecnología sigue dos vías distintas. Por un lado, el compromiso activo y crítico con la IA promueve la potenciación a largo plazo (LTP), fortaleciendo las redes sinápticas. Por otro lado, la aceptación acrítica de sugerencias algorítmicas sugiere una tendencia al debilitamiento de las conexiones neuronales encargadas del esfuerzo cognitivo (Sethi & Jain, 2024).

Bajo la Teoría de la Carga Cognitiva, resulta vital que la IA actúe como un andamiaje que reduzca la carga extraña para liberar recursos hacia la carga pertinente (aprendizaje profundo). Sin este diseño, el aprendizaje autorregulado se ve comprometido, pues la tecnología sustituye el esfuerzo necesario para la codificación profunda en lugar de facilitarlo.

La relación entre la reconfiguración sináptica y el aprendizaje autorregulado es central: la autorregulación implica la activación intencional de redes prefrontales que, al ejercitarse repetidamente, se consolidan anatómicamente. Si la IA automatiza estas decisiones, el entrenamiento de la corteza prefrontal dorsolateral —implicada en el control ejecutivo y la resolución de dilemas morales— puede verse atenuado.

Aprendizaje social

El cerebro humano es intrínsecamente social. El vínculo interpersonal activa circuitos de recompensa y neuronas espejo fundamentales para la empatía y la unión grupal (Vega, 2019). El trabajo cooperativo genera oxitocina, lo que disminuye la reactividad de la amígdala y aumenta la capacidad de la corteza prefrontal para la toma de decisiones éticas (Guillén, 2017).

La integración de la IA debe considerar que la "empatía simulada" de una máquina no sustituye la retroalimentación socioafectiva necesaria para el desarrollo de comportamientos altruistas y la responsabilidad compartida. Desde la ética del discurso, la formación moral requiere la exposición a la alteridad: la capacidad de ser afectado por el otro y de revisar las propias posiciones en un espacio de intersubjetividad. La mediación digital, al personalizar contenidos en función de preferencias previas, puede aislar al estudiante de la confrontación dialogal que caracteriza a la educación en valores.

METODOLOGÍA

El presente estudio se desarrolló bajo un enfoque cualitativo, mediante una revisión teórica con orientación hermenéutica, cuyo propósito fue analizar críticamente la relación entre inteligencia artificial, neuroeducación y procesos neurocognitivos vinculados con el aprendizaje ético.

Justificación del enfoque hermenéutico

Se optó por un enfoque hermenéutico porque el objeto de estudio —el impacto de la IA en los procesos neurocognitivos del aprendizaje ético— constituye una problemática emergente y multidimensional, para la cual aún no existen teorías consolidadas que permitan un abordaje exclusivamente deductivo o experimental. La hermenéutica permite interpretar los significados, tensiones y convergencias presentes en la literatura científica, reconstruyendo comprensiones que articulan lo descriptivo (cómo funciona el cerebro ante la IA) con lo normativo (cómo debe mediar tecnológicamente la formación ética). Este enfoque asume que la comprensión es circular: el investigador parte de una precomprensión del fenómeno, que posteriormente es revisada y enriquecida en el diálogo iterativo con los textos (Gadamer, 1960/2013, como se citó en Grondin, 2003).

Procedimiento de interpretación

El análisis se realizó en tres fases. En la primera fase (precomprensión y selección), se identificaron y preseleccionaron 31 documentos mediante búsqueda sistematizada en bases de datos. En la segunda fase (lectura comprensiva y codificación axial), se organizaron los documentos en tres ejes de análisis: dimensión afectiva, dimensión cognitiva y dimensión ética, identificando unidades de significado relevantes. En la tercera fase (interpretación dialógica o hermenéutica), se efectuó una lectura interpretativa y comparativa para identificar convergencias, tensiones y vacíos en la literatura revisada, construyendo una comprensión integrada que da cuenta de la ambivalencia de la mediación algorítmica.

En la fase de codificación axial, las unidades de significado fueron agrupadas según su pertenencia a las tres dimensiones analíticas preestablecidas (afectiva, cognitiva, ética). Este procedimiento permitió organizar la información de manera no lineal, identificando relaciones entre las categorías emergentes. La interpretación dialógica implicó un movimiento circular entre las precomprensiones del investigador y los sentidos emergentes de los textos, sin que ello supusiera una neutralidad epistemológica, sino una explicitación consciente de los horizontes de comprensión desde los cuales se abordó la literatura.

Estrategia de búsqueda y selección de fuentes

Para la selección de fuentes se revisaron bases de datos académicas como PubMed, Scopus, Dialnet, Web of Science y ERIC. Se emplearon cadenas de búsqueda booleana combinando términos como

("artificial intelligence" OR "machine learning" OR "intelligent tutoring systems") AND ("neuroeducation" OR "neurocognitive processes" OR "brain-based learning") AND ("ethics" OR "moral education" OR "neuroethics"). Se consideraron publicaciones comprendidas preferentemente entre 2020 y 2025, con excepciones teóricas fundacionales previas (2015-2019) debido a su relevancia conceptual indispensable para el marco argumentativo. Se priorizaron artículos científicos, revisiones sistemáticas, documentos institucionales y textos académicos de alto impacto.

Criterios de inclusión

- Publicaciones académicas con respaldo científico (revistas indexadas, editoriales universitarias o instituciones reconocidas) vinculadas con inteligencia artificial y educación.
- Estudios que abordaron procesos neurocognitivos como atención, memoria, emoción o autorregulación en contextos de aprendizaje.
- Documentos con aportes teóricos o empíricos sobre ética, formación moral, aprendizaje ético o neuroética.
- Textos publicados preferentemente entre 2020 y 2025, con excepciones fundacionales justificadas.
- Fuentes disponibles en español o inglés con acceso al texto completo.

Criterios de exclusión

- Documentos sin respaldo académico ni revisión por pares (blogs, artículos de opinión no fundados).
- Textos de opinión sin fundamentación teórica o empírica explícita.
- Publicaciones duplicadas o redundantes.
- Trabajos que abordaron la inteligencia artificial solo desde una perspectiva puramente técnica o de ingeniería, sin relación con procesos educativos, neurocognitivos o éticos.
- Capítulos de libro con desarrollo argumental insuficiente o sin datos identificables de edición y publicación.

Del universo inicial de 31 documentos preseleccionados, 25 cumplieron estrictamente con los criterios de elegibilidad y constituyeron el corpus principal de análisis. Las 6 fuentes excluidas correspondieron a textos de divulgación técnica sin implicaciones pedagógicas (3), capítulos de libro sin desarrollo propio (2) y un documento duplicado (1).

Consideraciones sobre la confiabilidad del proceso interpretativo

Dado que el análisis hermenéutico implica inevitablemente la subjetividad del investigador en la categorización e interpretación de los textos, se adoptaron las siguientes medidas de rigor: (a) la explicitación de los criterios de elegibilidad y de los ejes analíticos preestablecidos, que delimitaron el horizonte interpretativo; (b) la lectura iterativa de los documentos, que permitió contrastar las primeras comprensiones con relecturas posteriores; y (c) la vinculación explícita de cada interpretación con las fuentes primarias, de modo que el lector pueda reconstruir el trayecto argumentativo desde la evidencia textual hasta la inferencia del autor. No obstante, se reconoce que la hermenéutica no aspira a la replicabilidad estadística, sino a la comprensión situada y argumentativamente fundada.

RESULTADOS

Los hallazgos de la revisión se presentan diferenciando entre lo que reporta directamente la literatura consultada y las interpretaciones hermenéuticas derivadas del análisis del investigador. Esta distinción responde a la necesidad de transparentar el nivel de evidencia disponible y evitar que las inferencias interpretativas se confundan con datos empíricos reportados por los estudios primarios.

Detección emocional y autorregulación

Lo que reporta la literatura: La IA facilita la monitorización de estados afectivos, lo cual evidencia una capacidad para personalizar el ritmo de aprendizaje (Macías Acosta et al., 2025; Ocaña, 2015). Los estudios analizados señalan una tendencia a generar dependencia de estímulos externos para la gestión del estrés. Sethi y Jain (2024) advierten que la reconfiguración neuronal depende de la calidad del compromiso (activo versus pasivo), aunque no ofrecen datos neuroimagenológicos directos sobre este proceso en contextos específicos de IA educativa.

Interpretación del autor: La regulación emocional externalizada mediante computación afectiva podría interferir con el desarrollo de la autorregulación autónoma, componente esencial del carácter moral que permite postergar la gratificación y sostener la atención en dilemas éticos. Esta inferencia se deriva de la articulación entre los hallazgos sobre neuroquímicos del estrés (Waipan, 2017; Vega, 2019) y las advertencias sobre dependencia tecnológica (Fu y Weng, 2024), pero no constituye una evidencia empírica directamente demostrada por los estudios revisados.

Optimización de la atención versus fragmentación

Lo que reporta la literatura: La analítica atencional ayuda a identificar momentos de distracción (Ouyang et al., 2023). Los estudios revisados indican que la exposición prolongada a interfaces digitales fragmentadas afecta la atención sostenida necesaria para la reflexión ética profunda, aunque esta observación requiere validación empírica con diseños longitudinales.

Interpretación del autor: La fragmentación atencional inducida por la mediación digital constituye un riesgo interpretativo para la deliberación moral, que demanda concentración sostenida y control inhibitorio mediado por la corteza prefrontal. Esta inferencia se sustenta en la teoría de la carga cognitiva y en los correlatos neurobiológicos de la atención profunda, pero debe entenderse como una hipótesis teórica pendiente de contrastación experimental.

Efectos en la memoria

Lo que reporta la literatura: Los documentos analizados reportan la persistencia de la externalización mnemónica a sistemas de IA (Strielkowski et al., 2024; Fu y Weng, 2024). Se evidencia un potencial debilitamiento de los circuitos neuronales encargados de la retención a largo plazo, aunque no se dispone de evidencia neuroimagenológica directa que establezca una causalidad biológica.

Interpretación del autor: La externalización sistemática de funciones mnemónicas a algoritmos podría atenuar el fortalecimiento de las redes sinápticas asociadas a la codificación profunda, esencial para la retención de principios éticos y la aplicación contextualizada del juicio moral. Esta inferencia se deriva de la articulación entre los hallazgos sobre plasticidad sináptica (Márquez, 2016; Blanco, 2019) y los patrones de uso de IA reportados, pero no constituye una demostración empírica.

Personalización algorítmica y andamiaje

Lo que reporta la literatura: El uso de perfiles predictivos facilita el acceso a contenidos adaptados, funcionando como un andamiaje técnico (Ouyang et al., 2023; Strielkowski et al., 2024). La literatura consultada reporta que una personalización excesiva limita la exposición a perspectivas diversas, restringiendo el diálogo con la alteridad necesario para el juicio moral (Fu y Weng, 2024; UNESCO, 2024b).

Interpretación del autor: La lógica de optimización individual de los algoritmos tensiona directamente con la lógica de la formación moral profunda, que depende de la intersubjetividad y del vínculo relacional (Vega, 2019). Esta inferencia articula los hallazgos sobre neuroquímicos del aprendizaje social (Guillén, 2017; Vega, 2019) con las advertencias sobre personalización algorítmica, pero debe entenderse como una reconstrucción hermenéutica del autor, no como una conclusión empírica de los

estudios

primarios.

DISCUSIÓN

El análisis crítico revela un riesgo de "reduccionismo conductual", donde la IA optimiza resultados medibles, pero descuida la convicción interior. Desde una perspectiva deontológica comunicativa, el diálogo requiere la capacidad de ser afectado por el otro (Domínguez, 2016), una dimensión que la personalización predictiva puede fragilizar al aislar al estudiante en contenidos que refuerzan sus sesgos previos (Ouyang et al., 2023). En contraste, Vega (2019) sostiene que la formación moral profunda depende de la intersubjetividad y del vínculo relacional, postura que tensiona directamente con la lógica de optimización individual de los algoritmos descrita por Strielkowski et al. (2024).

Tensiones entre personalización algorítmica y autonomía cognitiva desde una perspectiva neurocientífica

La personalización predictiva opera mediante la reducción de la variabilidad de estímulos y la anticipación de respuestas, disminuyendo la carga cognitiva extrínseca. Sin embargo, la independencia deliberativa y la capacidad de juicio moral dependen del ejercicio repetido de la corteza prefrontal dorsolateral y ventromedial, áreas implicadas en el control ejecutivo, la inhibición de respuestas impulsivas y la resolución de dilemas éticos.

Cuando el algoritmo automatiza la toma de decisiones de aprendizaje (qué contenido ver, en qué secuencia, a qué ritmo), el estudiante puede experimentar una reducción del esfuerzo deliberativo necesario para fortalecer estas redes sinápticas. Sethi y Jain (2024) advierten que la reconfiguración cerebral depende de la calidad del compromiso: activo y crítico frente a pasivo y receptivo. Por tanto, una adaptación excesiva no solo limita la agencia pedagógica, sino que atenúa el entrenamiento de los procesos ejecutivos que sustentan el juicio ético.

La libertad moral no es solo independencia de elección, sino capacidad de sostener la atención, postergar la gratificación y revisar las propias convicciones, funciones todas ellas mediadas por circuitos prefrontales que requieren práctica deliberada.

Es fundamental integrar la ética del cuidado en este diálogo. Mientras que la IA ofrece un procesamiento algorítmico "frío", el desarrollo moral humano requiere el cuidado relacional "cálido" que activa la liberación de oxitocina y fortalece el vínculo interpersonal (Vega, 2019). Bourret (2011) distingue ontológicamente entre la "empatía simulada" de los sistemas computacionales y la "empatía humana" basada en la intersubjetividad. Sin el componente humano, se pierden elementos esenciales de la formación moral. Esta distinción se ve reforzada por Sethi y Jain (2024), quienes advierten que la reconfiguración neuronal depende de la calidad del compromiso: activo y crítico frente a pasivo y receptivo.

Articulación entre neuroplasticidad, aprendizaje autorregulado y ética discursiva

Para evitar la fragmentación conceptual, resulta necesario explicitar cómo estos tres constructos se articulan en el marco propuesto. La reconfiguración sináptica constituye el substrato biológico: el cerebro se remodela según la calidad de las experiencias de aprendizaje. El aprendizaje autorregulado es el proceso metacognitivo mediante el cual el estudiante planifica, monitorea y evalúa su propio desempeño, ejercitando intencionalmente las redes prefrontales. La deontología comunicativa es la dimensión normativa que exige la exposición a la alteridad, la argumentación pública y la revisibilidad de las posiciones.

La relación entre estos tres niveles permite comprender que una adaptabilidad sináptica favorable al juicio ético requiere precisamente de un aprendizaje autorregulado que no sea sustituido por la IA, sino andamiado por ella, y que dicho aprendizaje se desarrolle en espacios de intersubjetividad genuina, no en burbujas algorítmicas. La tecnología, por tanto, solo es éticamente deseable cuando funciona como andamiaje que potencia estos tres niveles sin colonizar la agencia del estudiante.

Asimismo, destaca la paradoja identificada por Fu y Weng (2024): los estudiantes adoptan la IA de forma autónoma ante la falta de acompañamiento docente, lo que aumenta los riesgos de una pérdida de la capacidad de análisis crítico por dependencia tecnológica. Esta situación pone en tensión el principio de independencia deliberativa con el de protección cognitiva, obligando a repensar el rol del docente no como mero gestor tecnológico, sino como mediador ético indispensable (UNESCO, 2024a).

Limitaciones del estudio

Resulta necesario reconocer las limitaciones inherentes a esta investigación. En primer lugar, por su naturaleza teórica y no empírica, los hallazgos son interpretativos y no pueden generalizarse a poblaciones específicas sin contrastación experimental. No se dispone de datos neuroimagenológicos propios ni de mediciones de desempeño cognitivo que establezcan causalidades directas entre el uso de IA y modificaciones cerebrales. Las inferencias sobre debilitamiento sináptico o fragmentación atencional deben entenderse como hipótesis teóricas derivadas de la literatura revisada, no como evidencias empíricas demostradas.

En segundo lugar, la selección de fuentes dependió de la disponibilidad en bases de datos indexadas y del dominio lingüístico (español e inglés), lo que pudo excluir producciones relevantes en otros idiomas o formatos no indexados.

En tercer lugar, el análisis hermenéutico implicó inevitablemente la subjetividad del investigador en la categorización e interpretación de los textos, aunque se mitigó mediante la explicitación de los criterios de elegibilidad, la lectura iterativa y la vinculación argumentativa con las fuentes primarias. Finalmente, la delimitación conceptual entre neuroética, formación moral y aprendizaje ético, aunque operativa para este estudio, requiere desarrollos subsiguientes que profundicen en sus diferencias ontológicas y epistemológicas.

CONCLUSIONES

Los hallazgos permiten sostener, desde una perspectiva interpretativa, que la inteligencia artificial tiene un impacto ambivalente en el cerebro. Si bien optimiza la personalización y potencia la maleabilidad cortical mediante el compromiso activo, también presenta riesgos interpretativos de atenuar funciones cognitivas esenciales cuando la intermediación tecnológica es pasiva y acrítica.

La aportación central de este trabajo es la construcción de un Marco de Neuroética Educativa, concebido explícitamente como un marco analítico-interpretativo —no como un modelo teórico formal ni como una propuesta pedagógica aplicada directamente—, diseñado para orientar una integración de la IA que priorice el desarrollo humano. Dicho marco constituye una propuesta original frente a la literatura predominante, al articular de manera sistémica los hallazgos neurocognitivos con los imperativos éticos del encuentro pedagógico, integrando la reconfiguración sináptica, el aprendizaje autorregulado y la deontología comunicativa en una propuesta comprensiva.

Los hallazgos permiten sostener que la mediación algorítmica incide en los procesos neurocognitivos del aprendizaje ético de forma bifurcada: potenciadora cuando funciona como andamiaje activo, y riesgosa cuando sustituye la agencia del estudiante. Los objetivos específicos se cumplieron al (1) identificar las transformaciones en las dimensiones afectiva, cognitiva y ética; (2) evidenciar la tensión entre personalización y autogobierno cognitivo; y (3) construir el marco analítico-interpretativo propuesto.

La pregunta de investigación fue respondida al determinar, desde una perspectiva hermenéutica, que la IA afecta los procesos neurocognitivos éticos principalmente mediante la fragmentación atencional, la externalización mnemónica y la reducción de la intersubjetividad, implicaciones que demandan un marco de neuroética educativa para preservar la dimensión humana. Se enfatiza que estas conclusiones

constituyen una base conceptual integradora que deberá ser validada mediante investigaciones empíricas futuras con diseños longitudinales y neuroimagenológicos.

RECOMENDACIONES

A los docentes se les recomienda una formación continua en neuroética educativa que trascienda el manejo técnico de herramientas. El docente debe asumirse como mediador ético que diseña andamiajes progresivos: la IA puede usarse para liberar carga cognitiva extrínseca, pero debe reservarse espacio deliberativo sin intermediación técnica donde el estudiante ejercite la toma de decisiones morales. Se sugiere implementar "rúbricas de autonomía" que evalúen no solo el producto, sino también el proceso de autorregulación y la capacidad de cuestionar las sugerencias de la IA.

A las instituciones educativas corresponde establecer comités de bioética educativa que regulen el uso de IA, priorizando la privacidad de los datos afectivos y cognitivos de los estudiantes. Las políticas institucionales deben garantizar acceso equitativo, evitando que la brecha digital se traduzca en una brecha neuroética, y limitar la fragmentación de interfaces en entornos de formación moral.

En materia de diseño curricular, la neuroética debe incorporarse como eje transversal, no como un módulo aislado. El currículo debe incluir espacios de "deliberación moral desconectada" (sin IA) donde los estudiantes resuelvan dilemas éticos complejos mediante el diálogo cara a cara, fortaleciendo así los circuitos sociales y prefrontales. La evaluación debe integrar dimensiones de juicio crítico, empatía intersubjetiva y autorregulación, más allá de la eficiencia algorítmica.

Tabla 1

Marco de Neuroética Educativa

Principio	Fundamento teórico	Implicación práctica
Preservación de la agencia humana	Neuroplasticidad activa (Sethi & Jain, 2024)	El diseño de interacciones con IA debe evitar la aceptación automática de sugerencias para promover el compromiso activo, lo cual es esencial para potenciar la reconfiguración cerebral y evitar el debilitamiento interpretativo de los circuitos prefrontales.
Prioridad de la intersubjetividad	Ética del cuidado (Vega, 2019)	La IA se concibe como un complemento y nunca como un sustituto del encuentro pedagógico, preservando el vínculo humano que activa los circuitos de recompensa y la oxitocina necesarios para el aprendizaje ético.
Protección de la atención profunda y gestión de la carga cognitiva	Teoría de la Carga Cognitiva	Minimizar la fragmentación atencional de las interfaces digitales para permitir la concentración prolongada requerida en la deliberación de dilemas morales complejos.
Fomento de la autonomía y	Aprendizaje autorregulado	Promover la capacidad del

aprendizaje autorregulado

Equidad y responsabilidad contextual Neuroética educativa aplicada

estudiante para establecer sus propios criterios de evaluación, evitando la dependencia algorítmica y favoreciendo una ética comunicativa basada en el juicio crítico.

Asegurar que la implementación de la IA no profundice las brechas existentes, adaptando las herramientas a las necesidades específicas de las escuelas rurales y contextos de vulnerabilidad.

Nota. Elaboración propia, basada en la revisión teórica realizada.

Tabla 2
Síntesis de Fuentes Principales Analizadas

Autor(es) y año	Enfoque metodológico/teórico	Hallazgos principales	Contribución al estudio
Algarra Pulido (2021)	Investigación formativa (trabajo de máster)	Diseño de formación ética con técnicas de neuroeducación	Base para articular neuroeducación y competencia ética
Arango (2013)	Ensayo filosófico	Solidaridad, democracia y derechos como fundamentos del vínculo social	Fundamentación ético-política de la intersubjetividad
Blanco (2019)	Monografía teórica	Arquitectura neurológica del aprendizaje y sus implicaciones pedagógicas	Marco neurobiológico de la plasticidad sináptica
Bourret (2011)	Divulgación científica	Distinción entre empatía simulada y empatía humana	Alerta sobre los límites ontológicos de la computación afectiva
Colmenares (2012)	Metodología cualitativa	Investigación-acción participativa como estrategia de cambio social	Referencia metodológica para la dimensión comunitaria de la ética
Domínguez (2016)	Tesis doctoral (neuroética)	Bases neuroéticas para la educación moral y una neurorracionalidad dialógica	Fundamentación de la ética discursiva desde la neurociencia
Fu & Weng (2024)	Revisión sistemática mixta	Principios de IA responsable centrada en el humano: equidad, privacidad,	Marco ético para la gobernanza de la IA en educación

		no maleficencia, agencia y transparencia	
Gamo (2016)	Guía didáctica (MOOC)	Neurodidáctica como aproximación práctica al cerebro en el aula	Puente entre neurociencia y diseño instruccional
Gracia (2019)	Ensayo filosófico	La libertad incorporada como clave para la neuroeducación moral	Conexión entre corporalidad, autonomía y aprendizaje ético
Guillén (2017)	Artículo científico	Bases neurobiológicas del trabajo en equipo y la cooperación	Evidencia de los correlatos neuroquímicos (oxitocina) del aprendizaje social
López (2004)	Ensayo político	El diálogo como condición de posibilidad para la democracia	Fundamentación de la ética comunicativa en el espacio público
Macías Acosta et al. (2025)	Revisión sistemática	IA y neuroeducación en la enseñanza de las ciencias	Síntesis actualizada de la intersección IA- neuroeducación
Márquez (2016)	Artículo de divulgación científica	Neuroplasticidad y modulación emocional como bases del aprendizaje significativo	Referencia fundacional sobre los mecanismos cerebrales del aprendizaje
Ocaña (2015)	Monografía	Neuroeducación aplicada: cómo aprende el cerebro y cómo deberían enseñar los docentes	Marco teórico sobre computación afectiva y adaptación pedagógica
Ouyang et al. (2023)	Investigación cuasiexperimental	Integración de predicción de rendimiento de IA con análisis de aprendizaje para mejorar el compromiso y la satisfacción estudiantil	Evidencia empírica sobre personalización algorítmica y resultados de aprendizaje
Poblete (2007)	Monografía	Aprendizaje basado en competencias y evaluación de competencias genéricas	Marco para vincular autorregulación con evaluación formativa
Richart (2018)	Artículo científico	Fronteras entre neuroética y neuroeducación en las neurociencias sociales	Delimitación del campo de la neuroética educativa

Rodríguez (2020)	Artículo científico	Aprendizaje como proceso vinculado al amor y la motivación intrínseca	Dimensión afectiva y motivacional del aprendizaje profundo
Román Fabián y Victoria Poenitz (2018)	Artículo de revisión	Aportes, desafíos y oportunidades de la neurociencia aplicada a la educación en América Latina	Contextualización regional de la neuroeducación
Sethi & Jain (2024)	Revisión sistemática	Tecnologías de IA para el aprendizaje socioemocional: realidad virtual, chatbots, análisis de sentimientos, gamificación y wearables	Evidencia sobre la bifurcación de la neuroplasticidad ante el compromiso activo vs. pasivo
Strielkowski et al. (2024)	Análisis bibliométrico	Análisis de 3.518 publicaciones sobre aprendizaje adaptativo e IA; reducción del 40-60% en duración del entrenamiento e incremento del 50% en retención	Estado del arte sobre personalización algorítmica y eficiencia del aprendizaje adaptativo
UNESCO (2024a)	Documento institucional	Marco de competencia en IA para docentes	Referencia normativa para la formación del profesorado
UNESCO (2024b)	Documento institucional	Guía de recomendaciones sobre ética e IA en educación	Fundamentación institucional para la gobernanza ética de la IA
Vega (2019)	Monografía	El ágora de la neuroeducación: explicación y aplicación de los mecanismos cerebrales al aula	Eje teórico central sobre emociones, empatía y aprendizaje ético
Waipan (2017)	Monografía	Neuroeducación, adolescencia y escuela secundaria; efectos del estrés y la amígdala	Correlatos neuroquímicos del estrés y su impacto en la memoria

Nota. Elaboración propia a partir de la revisión sistemática de la literatura.

AGRADECIMIENTOS

Dra. Nubia Cabrera Tovar.

[Revista de Investigación Multidisciplinaria Iberoamericana, RIMI](#) © 2023 by [Elizabeth Sánchez Vázquez](#) is licensed under

REFERENCIAS

- Algarra Pulido, A. F. (2021). *Diseño de formación para el fortalecimiento de la competencia ética con técnicas de neuroeducación* [Trabajo de fin de máster, Universidad Americana de Europa].
- Arango, R. (2013). Solidaridad, democracia y derechos. *Revista de Estudios Sociales*, (47), 12–23.
- Blanco, I. (2019). *Arquitectura neurológica del aprendizaje y sus implicaciones pedagógicas*. Ediciones Académicas.
- Bourret, M. (2011). El poder de la empatía: Límites de la computación afectiva. *Revista de Divulgación Científica*, 15(3), 45–58.
- Colmenares, A. (2012). Investigación-acción participativa: Una metodología para el cambio social. *Revista de Investigación Social*, 12(2), 112–129.
- Domínguez, D. (2016). *Bases neuroéticas para la educación moral: Una neurorracionalidad dialógica y práctica* [Tesis doctoral, Universitat Jaume I].
- Fu, Y., & Weng, Z. (2024). Navigating the ethical terrain of AI in education: A systematic review on framing responsible human-centered AI practices. *Computers and Education: Artificial Intelligence*, 7, 100306. <https://doi.org/10.1016/j.caeai.2024.100306>
- Gamo, J. R. (2016). *Neurodidáctica: Una aproximación práctica al cerebro en el aula* [Guía didáctica MOOC]. Universidad Nacional.
- Gracia, J. (2019). La libertad incorporada como clave para la neuroeducación moral. *Sophia. Colección de Filosofía de la Educación*, 27(2), 201–218.
- Grondin, J. (2003). *Introducción a Gadamer* (A. Martínez Riu, Trad.). Herder. (Obra original publicada en 1991).
- Guillén, J. C. (2017). Cooperación y cerebro: Bases neurobiológicas del trabajo en equipo. *Aula Abierta*, 46(1), 45–60.
- López, J. N. (2004). El diálogo ¿otro imposible? *Reflexión Política*, 6(11), 89–102.
- Macías Acosta, R. E., Tomalá de la Cruz, Á. del R., & Marqués Molias, L. (2025). Artificial intelligence and neuroeducation in science education: A systematic review. *TPM: Testing, Psychometrics, Methodology in Applied Psychology*, 32(S9), 45–62.
- Márquez, A. C. (2016). Neuroeducación: De la neurociencia al aula. *Revista Aula Infantil*, 85(3), 18–25.
- Ocaña, A. O. (2015). *Neuroeducación: ¿Cómo aprende el cerebro humano y cómo deberían enseñar los docentes?* Ediciones de la U.
- Ouyang, F., Wu, M., Zheng, L., Zhang, L., & Jiao, P. (2023). Integration of artificial intelligence performance prediction and learning analytics to improve student learning in online engineering course. *International Journal of Educational Technology in Higher Education*, 20(1), Article 4. <https://doi.org/10.1186/s41239-022-00372-4>
- Poblete, A. V. (2007). *Aprendizaje basado en competencias: Una propuesta para la evaluación de competencias genéricas*. Editorial Universitaria.
- Richart, D. P. (2018). Entre la neuroética y la neuroeducación: Las fronteras de las neurociencias sociales. *Revista de Pensament i Anàlisi*, 21(2), 145–162.
- Rodríguez, C. N. (2020). Neuroeducación: Solo se puede aprender aquello que se ama. *Educatio Siglo XXI*, 38(1), 23–40.
- Román Fabián, V. P., & Victoria Poenitz, M. (2018). La neurociencia aplicada a la educación: Aportes, desafíos y oportunidades en América Latina. *RELAdeI. Revista Latinoamericana de Aprendizaje y Desarrollo en Educación Infantil*, 5(2), 55–72.

- Sethi, S. S., & Jain, K. (2024). AI technologies for social emotional learning: Recent research and future directions. *Journal of Research in Innovative Teaching & Learning*, 17(2), 213–225. <https://doi.org/10.1108/JRIT-03-2024-0073>
- Strielkowski, W., Grebennikova, V., Lisovskiy, A., Rakhimova, G., & Vasileva, T. (2024). AI-driven adaptive learning for sustainable educational transformation. *Sustainable Development*, 33(2), 1921–1947. <https://doi.org/10.1002/sd.3221>
- UNESCO. (2024a). *Marco de competencia en materia de inteligencia artificial para docentes*. <https://doi.org/10.54675/ZJTE2084>
- UNESCO. (2024b). *Guía para el uso de IA generativa en educación e investigación*. <https://unesdoc.unesco.org/ark:/48223/pf0000389227>
- Vega, L. L. (2019). *El ágora de la neuroeducación: La neuroeducación explicada y aplicada*. Ediciones Académicas.
- Waipan, A. M. (2017). *El cerebro adolescente va al aula: Neuroeducación, adolescencia y escuela secundaria; efectos del estrés y la amígdala*. Editorial Educar.