Artículo de revisión

# Del tutor mecánico al tutor generativo: inteligencia artificial en la educación y sus efectos en la salud cognitiva, mental y emocional de los estudiantes. Revisión integrativa (1955-2026)

*From the mechanical tutor to the generative tutor: artificial intelligence in education and its effects on students' cognitive, mental and emotional health. An integrative review (1955-2026)*

Fabián Armando Garzón León[1]

Recibido: 28 de septiembre de 2026
Aceptado: 12 de noviembre de 2026

## Resumen

**Objetivo.** Sintetizar la evidencia sobre la inteligencia artificial (IA) en la educación, desde el origen del término (1955) hasta 2026, y analizar sus efectos en la salud cognitiva, mental y emocional de estudiantes y docentes, por nivel formativo.

**Método.** Revisión integrativa según Whittemore y Knafl. Búsqueda sistematizada en PubMed el 27 de septiembre de 2026 (934 registros; 93 revisiones), complementada con búsqueda dirigida en arXiv, ACM, Nature, PNAS y literatura institucional, y rastreo de citas. Se incluyeron 23 fuentes empíricas o de síntesis y 14 documentos normativos e institucionales, jerarquizados por nivel de evidencia y analizados mediante síntesis temática y cuantitativa descriptiva.

**Resultados.** El efecto de la IA depende del diseño pedagógico. Tutores con andamiaje produjeron ganancias de 0,23 a 0,31 desviaciones estándar (DE) en Nigeria y mayor aprendizaje en física universitaria. El acceso libre a GPT-4 redujo el desempeño 17 % al retirarlo. En lo cognitivo, la descarga cognitiva se asocia con menor pensamiento crítico y menor conectividad neural en escritura asistida (evidencia preliminar). En lo mental, la dependencia de IA generativa se asocia con ansiedad, depresión y agotamiento académico, mediados por menor autoeficacia. El uso diario intensivo de chatbots se asocia con soledad y dependencia emocional. Los chatbots basados en terapia cognitivo-conductual y las simulaciones clínicas muestran beneficios. La evidencia es escasa en formación técnica, tecnológica y de posgrado.

**Conclusiones.** La IA funciona como andamio o como muleta según el diseño que la rodea. Se propone el Modelo de Andamiaje Protector y recomendaciones para la educación abierta y a distancia.

**Palabras clave:** Inteligencia artificial; Educación a distancia; Salud mental; Cognición; Estudiantes; Dependencia psicológica.

## Abstract

**Objective.** To synthesize evidence on artificial intelligence (AI) in education, from the coining of the term (1955) to 2026, and to analyze its effects on the cognitive, mental and emotional health of students and teachers across educational levels.

**Method.** Integrative review following Whittemore and Knafl. A systematized PubMed search was run on 27 September 2026 (934 records; 93 reviews), complemented by targeted searches in arXiv, ACM, Nature, PNAS and institutional literature, plus citation tracking. Twenty-three empirical or synthesis sources and fourteen normative and institutional documents were included, graded by level of evidence, and analyzed through thematic and descriptive quantitative synthesis.

**Results.** AI's effect depends on pedagogical design. Scaffolded tutors yielded gains of 0.23–0.31 standard deviations in Nigeria and greater learning in university physics. Unrestricted GPT-4 access lowered performance by 17% once removed. Cognitively, offloading is associated with weaker critical thinking and lower neural connectivity during assisted writing (preliminary evidence). Mentally, generative-AI dependence is associated with anxiety, depression and academic burnout, mediated by lower self-efficacy. Heavy daily chatbot use is associated with loneliness and emotional dependence. CBT-based chatbots and clinical simulations show benefits. Evidence is scarce for technical, technological and graduate education.

**Conclusions.** AI acts as a scaffold or a crutch depending on the design around it. We propose the Protective Scaffolding Model and recommendations for open and distance education.


**Keywords:** Artificial Intelligence; Education, Distance; Mental Health; Cognition; Students; Dependency, Psychological.

1. Estudiante del programa de Administración en Salud, Universidad Nacional Abierta y a Distancia (UNAD). Bogotá, Colombia.
Correo electrónico: fagarzonle@unadvirtual.edu.co
ORCID: https://orcid.org/0009-0001-0043-6387

## Introducción

La pregunta ya no es si los estudiantes usan inteligencia artificial, sino qué le hace ese uso a su mente. Desde el lanzamiento de ChatGPT en noviembre de 2022, la IA generativa pasó de curiosidad a rutina en todos los niveles educativos. En Estados Unidos, el 72 % de los adolescentes de 13 a 17 años ha usado al menos una vez un "compañero" de IA, y el 52 % lo hace con regularidad (1). Seis de cada diez docentes de educación básica y media usan herramientas de IA en su trabajo (2).

La investigación inicial midió sobre todo rendimiento académico. Esa agenda dejó una brecha: la salud. Aprender exige esfuerzo cognitivo sostenido, atención y tolerancia a la dificultad; si una herramienta sustituye ese esfuerzo, el costo puede aparecer en la memoria, el pensamiento crítico, la autoeficacia y el bienestar emocional. Zawacki-Richter et al. (3) ya advertían que la literatura sobre IA en educación superior carecía de reflexión crítica y de conexión pedagógica.

Siete años después, la dimensión sanitaria sigue dispersa entre la psicología, la educación médica y la interacción humano-computador.

Este artículo continúa una línea de trabajo del programa de Administración en Salud de la UNAD. En una revisión previa se analizó cómo la IA y el machine learning optimizan la gestión administrativa de las instituciones de salud y reducen cargas que se asocian con el agotamiento del personal (4). Aquí la mirada se desplaza del hospital al aula: del trabajador de la salud al estudiante como sujeto de salud pública.

**Pregunta de investigación.** ¿Qué evidencia existe, entre 1955 y 2026, sobre los efectos de la IA en la educación en la salud cognitiva (atención, memoria, pensamiento crítico), mental y emocional (ansiedad, depresión, agotamiento, soledad, dependencia) de estudiantes y docentes, y cómo varían esos efectos según el nivel formativo y el diseño pedagógico?

**Objetivo.** Sintetizar e interpretar esa evidencia desde una perspectiva de salud pública, situarla en su trayectoria histórica y normativa, y derivar un modelo explicativo y recomendaciones para la educación abierta y a distancia en Colombia.

## Marco teórico

### 1.1. El origen del término: el aprendizaje en el centro

La expresión artificial intelligence aparece impresa por primera vez en la propuesta del proyecto de verano de Dartmouth, firmada el 31 de agosto de 1955 por McCarthy, Minsky, Rochester y Shannon (5). Su conjetura fundacional afirmaba que todo aspecto del aprendizaje puede describirse con tal precisión que una máquina lo simule. La IA nació, por tanto, definida desde el aprendizaje. Setenta años después, la pregunta se invierte: ya no es si la máquina aprende, sino si el humano sigue aprendiendo cuando la máquina piensa por él.

### 1.2. El transcurso: del tutor mecánico al tutor generativo

Este trabajo denomina “Del tutor mecánico al tutor generativo” al recorrido histórico de la IA educativa, organizado en cinco olas. Cada ola cambió quién hace el esfuerzo cognitivo: la máquina, el estudiante o ambos.

**Tabla 1.** *Las cinco olas de la IA educativa, del tutor mecánico al tutor generativo*

| ***Ola*** | ***Periodo*** | ***Hito representativo*** | ***Lógica pedagógica*** | ***Quién hace el esfuerzo*** |
|---|---|---|---|---|
| *5. IA generativa conversacional* | 2022 - 2026 | ChatGPT y modelos de lenguaje; primeros ECA de tutoría generativa (6,7) | Diálogo abierto; produce respuestas completas | Riesgo de delegación total al sistema |
| *4. Analítica y aprendizaje adaptativo* | 2010 - 2022 | Machine learning para predecir deserción y personalizar rutas (3) | Predicción y recomendación | Compartido; el sistema decide la ruta |
| *3. Tutores inteligentes* | 1970 - 2010 | SCHOLAR, primer tutor con red semántica (8); problema de las dos sigmas (9) | Modelar al estudiante y guiarlo paso a paso | Estudiante, con pistas del sistema |
| *2. Instrucción asistida por computador* | 1958 - 1970 | Instrucción programada; sistemas como PLATO | Secuencias ramificadas y retroalimentación | Estudiante |
| *1. Máquinas de enseñar* | 1926 - 1958 | Aparato de Pressey (10); máquinas de Skinner (11) | Retroalimentación inmediata y respuesta construida | Estudiante |

Bloom (9) mostró que la tutoría individual humana supera en dos desviaciones estándar a la enseñanza grupal. Ese "problema de las dos sigmas" es la promesa que persigue cada ola. La quinta es la primera capaz de escalar la tutoría, pero también la primera capaz de entregar la respuesta completa sin exigir razonamiento.

### 1.3. Marco normativo

**Tabla 2.** *Instrumentos normativos y de política sobre IA en educación con implicaciones para la salud*

| *Instrumento* | *Año* | *Disposición relevante para salud y cognición* |
|---|---|---|
| *Ley 2626, Colombia (educación digital)* | 2026 | Incluye IA, pensamiento computacional y ciudadanía digital en básica y media (12) |
| *Digital Education Outlook, OCDE* | 2026 | Advierte "pereza metacognitiva" si la IA reemplaza el esfuerzo productivo (13) |
| *Aviso de salud sobre IA y adolescentes, APA* | 2025 | Pide proteger las relaciones reales y la privacidad de los adolescentes (14) |
| *CONPES 4144, Colombia* | 2025 | Política nacional de IA; 106 acciones a 2030 (15) |
| *Reglamento de IA, Unión Europea* | 2024 | Prohíbe el reconocimiento de emociones en instituciones educativas (16) |
| *Marcos de competencias en IA, UNESCO* | 2024 | 12 competencias para estudiantes y 15 para docentes (17,18) |
| *Guía sobre IA generativa, UNESCO* | 2023 | Edad mínima de 13 años para uso en aula; agencia humana (19) |
| *Recomendación sobre la ética de la IA, UNESCO* | 2021 | Primer instrumento normativo global sobre ética de la IA (20) |

### 1.4. Modelo conceptual: el esfuerzo como variable de salud

Cuatro teorías explican por qué la IA puede proteger o dañar la salud cognitiva y mental.

- **Carga cognitiva** (21): la memoria de trabajo es limitada. La IA puede reducir la carga extrínseca (lo accesorio) o, mal usada, eliminar la carga germinal que construye esquemas duraderos.
- **Descarga cognitiva** (22): delegar en herramientas externas libera recursos, pero reduce la práctica de la función delegada.
- **Autoeficacia** (23): la experiencia de dominio es su fuente principal. Si la IA resuelve la tarea, el estudiante pierde la experiencia que fortalece su confianza.
- **Conservación de recursos** (24): bajo estrés, las personas protegen sus recursos. Recurrir a la IA puede ser un afrontamiento que alivia a corto plazo y agota a largo plazo.

En este artículo, salud cognitiva designa la preservación de la atención, la memoria, el razonamiento y la metacognición para el aprendizaje autónomo. Salud mental y emocional designa la ausencia de síntomas de ansiedad, depresión y agotamiento, junto con vínculos sociales y autoeficacia preservados.

### 1.5. Ecosistema institucional 2024 - 2026: adopción masiva, gobernanza rezagada

Entre 2024 y 2026 la adopción de la IA generativa en la educación avanzó más rápido que la capacidad institucional para orientarla. En el Reino Unido, el 95 % de los estudiantes de pregrado de tiempo completo usó IA en al menos una forma y el 94 % la empleó en trabajos evaluados; en 2024 la proporción de usuarios era del 66 % (25). En contraste, solo el 48 % de esos estudiantes percibe que sus docentes les ayudan a desarrollar competencias en IA.

Del lado docente, la Encuesta Internacional de Enseñanza y Aprendizaje (TALIS) de 2024, con 280.000 educadores de 55 sistemas educativos, mostró que cerca de uno de cada tres docentes de la OCDE usa IA en su trabajo, que alrededor del 38 % recibió formación en el tema durante el

último año y que siete de cada diez temen que facilite el plagio (26). Del lado institucional, una encuesta de UNESCO a 400 cátedras y redes de 90 países halló que apenas el 19 % de las instituciones tenía política formal de IA y el 42 % la estaba elaborando; en América Latina y el Caribe solo el 45 % tenía orientaciones vigentes o en desarrollo, frente a cerca del 70 % en Europa y Norteamérica (27). Una de cada cuatro instituciones reportó problemas éticos, entre ellos la sobredependencia estudiantil.

**Tabla 3.** *Indicadores institucionales de adopción y gobernanza de la IA en educación, 2024 - 2026*

| *Fuente* | *Población y año* | *Indicador* | *Valor* |
|---|---|---|---|
| *Stephenson y Armstrong (25), HEPI* | 1.054 estudiantes de pregrado, Reino Unido, dic. 2025 | Uso de IA en al menos una forma | 95 % |
| *Stephenson y Armstrong (25), HEPI* | Misma muestra | Uso de IA generativa en trabajos evaluados | 94 % |
| *Stephenson y Armstrong (25), HEPI* | Misma muestra | La IA los hace sentir más solos / menos solos | 20 % / 21 % |
| *Common Sense Media (1)* | 1.060 adolescentes de 13 a 17 años, EE. UU., 2025 | Uso alguna vez de compañeros de IA | 72 % |
| *Gallup y Walton Family Foundation (2)* | 2.232 docentes de básica y media, EE. UU., 2025 | Uso de IA en el trabajo docente | 60 % |
| *OCDE (26), TALIS* | 280.000 educadores, 55 sistemas, 2024 | Docentes que usan IA en su trabajo (promedio OCDE) | ~33 % |
| *OCDE (26), TALIS* | Misma muestra | Docentes que temen que la IA facilite el plagio | ~70 % |
| *UNESCO (27)* | 400 cátedras UNESCO y redes UNITWIN, 90 países | Instituciones con política formal de IA | 19 % |
| *UNESCO (27)* | Misma muestra, América Latina y el Caribe | Instituciones con orientaciones vigentes o en desarrollo | 45 % |
| *Ministerio de Educación Nacional (28)* | Educación superior, Colombia | Programas activos en IA | 865 en 190 instituciones |

Nota. Las poblaciones, los países y las preguntas difieren; los valores describen el contexto y no deben compararse directamente.

En Colombia, la Universidad de los Andes publicó en octubre de 2024 los primeros lineamientos institucionales para el uso de IA generativa dirigidos a estudiantes, profesores y personal administrativo (29). El Concejo de Bogotá ordenó en 2025 expedir lineamientos distritales de uso responsable de la IA, incluidos talleres sobre IA generativa en colegios (30). En la UNAD, el Sello Editorial fijó en septiembre de 2024 lineamientos que prohíben usar IA para generar o modificar significativamente manuscritos, salvo casos justificados y descritos en la sección de métodos, y que exigen declarar su uso (31). El rezago de gobernanza no es solo administrativo: sin orientación institucional, la decisión de usar la IA como andamio o como muleta recae por completo en el estudiante, y con frecuencia la toma bajo estrés.

### 1.6. La salud del estudiante como asunto de salud pública

Este artículo adopta una perspectiva de salud pública porque la educación es un determinante social de la salud. El marco conceptual de la Organización Mundial de la Salud ubica la educación entre los determinantes estructurales que, a través de condiciones materiales, psicosociales y conductuales, producen gradientes de salud en la población (32). Si la IA transforma la experiencia educativa, transforma también uno de los canales por los que se produce la salud.

Tres constructos complementan el modelo de la sección 1.4:

- **Autodeterminación** (33). El bienestar depende de satisfacer tres necesidades psicológicas básicas: autonomía, competencia y relación. La IA puede apoyar la

competencia cuando andamia, erosionarla cuando sustituye, y competir con la relación humana cuando actúa como compañía.

- **Dificultades deseables** (34). Ciertas dificultades durante el aprendizaje, como la recuperación activa y el espaciamiento, reducen el desempeño inmediato pero mejoran la retención. Una herramienta que elimina toda dificultad maximiza la fluidez aparente y minimiza el aprendizaje duradero.
- **Salud mental positiva**. La ausencia de síntomas no basta: la autoeficacia, la agencia y los vínculos sociales son recursos protectores, y son precisamente los que la IA puede fortalecer o debilitar.

Desde esta perspectiva, las preguntas pertinentes para la gestión en salud no son tecnológicas sino poblacionales: quiénes están expuestos, con qué intensidad, con qué factores de riesgo y protección, y con qué desenlaces medibles.

## Método

### 2.1. Diseño

Se realizó una revisión integrativa siguiendo las cinco etapas de Whittemore y Knafl (35): formulación del problema, búsqueda, evaluación, análisis y presentación. Este diseño permite combinar ensayos, estudios observacionales, preprints y documentos normativos, necesarios en un campo que cambia cada mes. El reporte sigue los elementos aplicables de PRISMA 2020 (36). El protocolo no se registró en PROSPERO; esto se declara como limitación.

### 2.2. Estrategia de búsqueda

La búsqueda en PubMed se ejecutó el 27 de septiembre de 2026, con ventana del 1 de noviembre de 2022 al 30 de septiembre de 2026.

**Cadena 1 (amplia), 934 registros:**

*("generative artificial intelligence"[tiab] OR ChatGPT[tiab] OR "large language model*"[tiab] OR chatbot*[tiab])*

*AND (student*[tiab] OR adolescen*[tiab] OR learner*[tiab])*

*AND ("mental health"[tiab] OR anxiety[tiab] OR depression[tiab] OR stress[tiab] OR dependence[tiab]*

*OR loneliness[tiab] OR "critical thinking"[tiab] OR attention[tiab] OR cognitive[tiab])*

**Cadena 2 (síntesis), 93 registros:** la cadena 1, ampliada con "artificial intelligence"[tiab] y wellbeing/well-being[tiab], y restringida a systematic review[pt], meta-analysis[pt] o "scoping review"[ti].

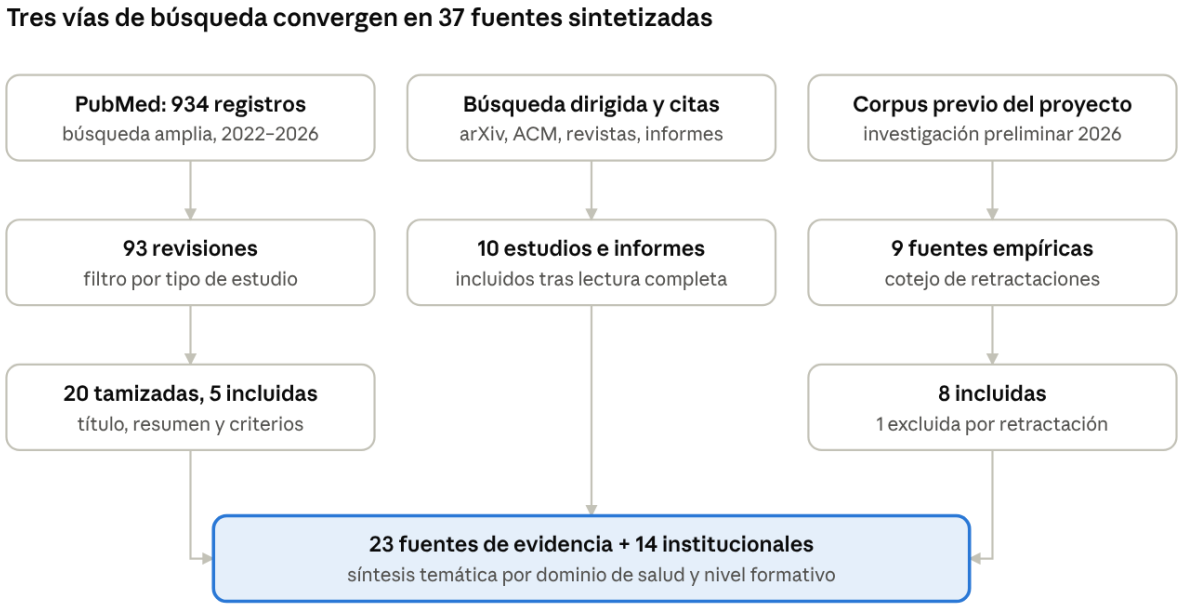
Se complementó con búsqueda dirigida en arXiv, ACM Digital Library, Nature, PNAS y sitios de UNESCO, OCDE, APA, Unión Europea, DNP y Ministerio de Educación Nacional, y con rastreo de citas hacia atrás. Se incorporó el corpus de la investigación preliminar del proyecto, previa verificación de su vigencia.

### 2.3. Criterios de elegibilidad

- **Inclusión:** estudios empíricos, revisiones o documentos oficiales sobre IA en contextos educativos formales o no formales, con desenlaces de aprendizaje, cognición, salud mental o bienestar; en español o inglés; con acceso al resumen estructurado o al texto completo.
- **Exclusión:** artículos de opinión sin datos; estudios de IA clínica sin componente educativo; publicaciones retractadas; fuentes sin autoría o fecha identificables.
- **Verificación de retractaciones:** se cotejó cada fuente contra los avisos del editor. El metaanálisis de Wang y Fan (37) fue excluido por su retractación en 2026.

## 2.4. Selección de las fuentes

**Figura 1.** *Flujo de identificación y selección de fuentes, adaptado de PRISMA 2020*




De las 93 revisiones se tamizaron por título y resumen las 20 de mayor relevancia; cinco cumplieron criterios. La búsqueda dirigida aportó diez estudios e informes, entre ellos las encuestas TALIS 2024 y HEPI 2026, y el corpus previo aportó ocho fuentes vigentes.

## 2.5. Valoración de la evidencia

Cada fuente se clasificó en un nivel de evidencia y cada hallazgo recibió una certeza global inspirada en GRADE: alta, moderada, baja o muy baja.

**Tabla 4.** *Niveles de evidencia aplicados a las fuentes*

| *Nivel* | *Tipo de fuente* | *Ejemplo en esta revisión* |
|---|---|---|
| *I* | Ensayo controlado aleatorizado o metaanálisis | Bastani et al. (6); Wu et al. (38) |
| *II* | Revisión sistemática o de alcance | Nasirun et al. (39) |
| *III* | Estudio cuasiexperimental u observacional | Chavez Sosa y Huancahuire-Vega (40) |
| *IV* | Preprint o encuesta sin revisión por pares | Kosmyna et al. (41); Fang et al. (42) |
| *V* | Documento normativo o consenso institucional | UNESCO (17); APA (14) |

## 2.6. Síntesis

Se aplicó síntesis temática en cuatro dominios: aprendizaje, salud cognitiva, salud mental y emocional, y ética. Cada hallazgo se cruzó con el nivel formativo: básica y media, técnica y

tecnológica, pregrado, posgrado, educación continua y docentes. Donde no hubo evidencia directa para un nivel se indica explícitamente, sin extrapolar como si fuera un hallazgo.

# Resultados

## 3.1. Caracterización de la evidencia

Se sintetizaron 23 fuentes de evidencia: cuatro de nivel I, seis de nivel II, seis de nivel III y siete de nivel IV. El 96 % (22 de 23) se publicó entre 2023 y 2026. Solo una proviene de América Latina (Perú) y ninguna de Colombia.

**Tabla 6.** *Fuentes de evidencia incluidas, ordenadas por nivel*

| *Fuente* | *Diseño* | *n* | *Nivel* | *Hallazgo principal* |
|---|---|---|---|---|
| *De Simone et al. (7)* | ECA, 6 semanas, Nigeria | 800 | I | Tutor GPT-4 supervisado: +0,31 DE global; +0,23 DE en inglés |
| *Bastani et al. (6)* | ECA, bachillerato, Turquía | ~1.000 | I | GPT libre: −17 % al retirarlo; GPT con pistas mitiga el daño |
| *Kestin et al. (43)* | ECA cruzado, física universitaria | 194 | I | Tutor con buenas prácticas supera la clase activa, en menos tiempo |
| *Wu et al. (38)* | Metaanálisis, 35 estudios | 4.193 | I | Efecto moderado sobre resultados de aprendizaje (g = 0,670) |
| *Jiang et al. (44)* | RS mixta, 19 estudios, enfermería | 1.253 | II | Simulación con IA: más autoeficacia y menos ansiedad al practicar |
| *Henttonen et al. (45)* | RS, 13 estudios, docentes de enfermería | 3.082 | II | Docentes temen erosión del pensamiento crítico e integridad |
| *Li et al. (46)* | RS, estudiantes de medicina | - | II | La sobredependencia produce inercia mental y debilita el juicio crítico |
| *Nasirun et al. (39)* | RS, 14 estudios (7 ECA), enfermería | 1.107 | II | Mejora procedimental; reflexión y retención comprometidas |
| *Nyakhar y Wang (47)* | RS rápida, 9 estudios, universitarios | 1.082 | II | 8 de 9 chatbots mejoraron ansiedad, depresión o bienestar |
| *Zawacki-Richter et al. (3)* | RS, educación superior | 146 | II | Escasa reflexión crítica, ética y pedagógica en la literatura |
| *Wang et al. (48)* | Transversal, universitarios, China | 1.623 | III | La dependencia media el paso del estrés al agotamiento y la ansiedad |
| *Chavez Sosa y Huancahuire-Vega (40)* | Transversal, medicina, Perú | 187 | III | Cada punto de dependencia: +5 % ansiedad y +4 % depresión |
| *Gerlich (49)* | Transversal, métodos mixtos | 666 | III | Uso de IA asociado a menor pensamiento crítico vía descarga |
| *Lee et al. (50)* | Encuesta a trabajadores del conocimiento | 319 | III | Más confianza en la IA, menos pensamiento crítico ejercido |
| *Zhang et al. (51)* | Transversal, universitarios | - | III | Estrés académico y baja autoeficacia predicen dependencia |
| *Liang et al. (52)* | Evaluación de detectores de IA | Ensayos | III | Más de la mitad de textos de no nativos marcados como IA |
| *Fang et al. (42)* | ECA, 4 semanas, preprint | 981 | IV | Mayor uso diario: más soledad y dependencia emocional |
| *Kosmyna et al. (41)* | Experimental con EEG, preprint | 54 | IV | Escritura con LLM: menor conectividad neural y autoría |
| *Common Sense Media (1)* | Encuesta nacional, 13 - 17 años | 1.060 | IV | 72 % usó un compañero de IA; 52 % con regularidad |
| *Gallup y Walton Family Foundation (2)* | Encuesta probabilística, docentes | 2.232 | IV | Uso semanal: 5,9 horas ahorradas por semana |
| *Bozkurt et al. (53)* | Ensayo colectivo, 35 autores | - | IV | Posibilidades y riesgos para la educación abierta y a distancia |
| *OCDE (26)* | Encuesta TALIS 2024, 55 sistemas educativos | 280.000 | IV | Un tercio de los docentes usa IA; siete de cada diez temen el plagio |

| | | | | |
|---|---|---|---|---|
| *Stephenson y Armstrong (25)* | Encuesta ponderada a universitarios, Reino Unido | 1.054 | IV | 95 % usa IA; efectos polarizados sobre la soledad |

## 3.2. Síntesis cuantitativa descriptiva

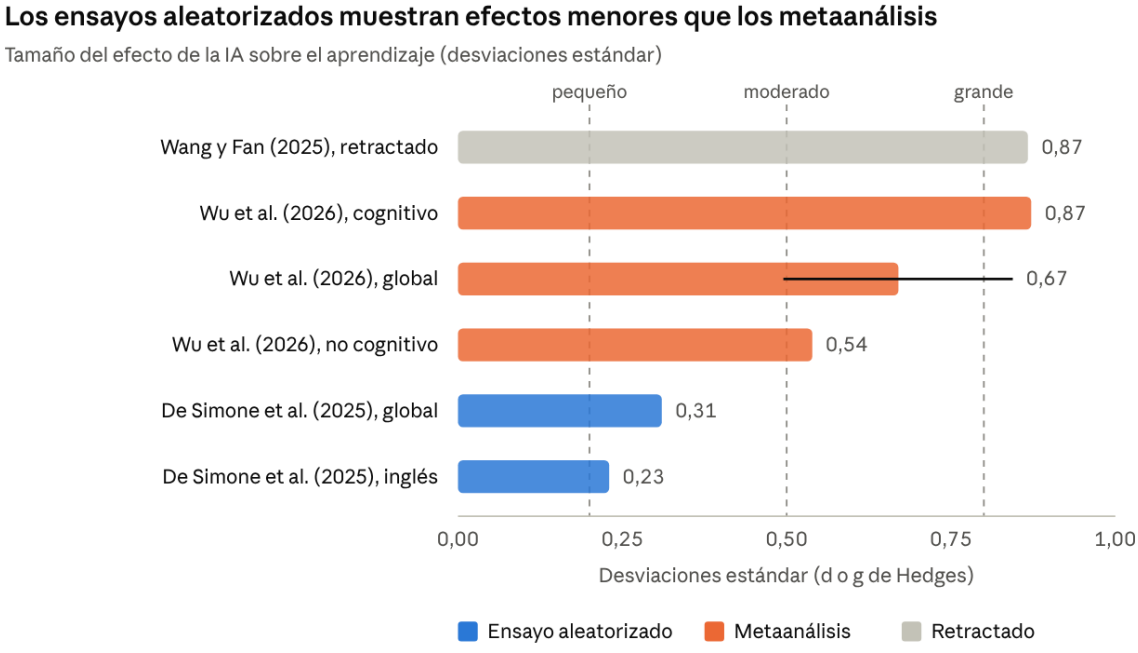
La heterogeneidad de diseños, poblaciones y desenlaces impide un metaanálisis propio; se presenta una síntesis cuantitativa descriptiva que reúne los estimadores reportados por las fuentes. La Tabla 7 resume la distribución de la evidencia por diseño.

**Tabla 7.** *Distribución de las 23 fuentes de evidencia por diseño*

| ***Diseño*** | ***n*** | ***%*** | ***Fuente*** |
|---|---|---|---|
| *Revisión sistemática o de alcance* | 6 | 26,1 | Henttonen, Jiang, Li, Nasirun, Nyakhar, Zawacki-Richter |
| *Transversal analítico* | 5 | 21,7 | Chavez Sosa, Gerlich, Lee, Wang, Zhang |
| *Encuesta institucional o poblacional* | 4 | 17,4 | Common Sense Media, Gallup, HEPI, OCDE-TALIS |
| *Ensayo controlado aleatorizado* | 4 | 17,4 | Bastani, De Simone, Fang (preprint), Kestin |
| *Metaanálisis* | 1 | 4,3 | Wu |
| *Experimental con EEG* | 1 | 4,3 | Kosmyna (preprint) |
| *Evaluación de desempeño de detectores* | 1 | 4,3 | Liang |
| *Ensayo colectivo de expertos* | 1 | 4,3 | Bozkurt |
| **Total** | **23** | **100** | |

Solo el 17,4 % de la evidencia proviene de ensayos aleatorizados, y uno de ellos es un preprint. Los desenlaces de salud mental se apoyan casi por completo en diseños transversales.

**Figura 2.** *Tamaños de efecto sobre el aprendizaje reportados en las fuentes incluidas; la línea horizontal indica el IC 95 %*




Los dos ensayos con andamiaje reportan efectos pequeños a moderados (0,23 a 0,31 DE), mientras el metaanálisis vigente estima un efecto moderado con heterogeneidad muy alta ($I^2$ = 91,4 %; 38). El metaanálisis de Wang y Fan (37), que reportaba un efecto grande (g = 0,867), fue retractado en abril de 2026; se muestra en gris solo como advertencia metodológica. La brecha entre ensayos y síntesis sugiere que los metaanálisis tempranos agrupan estudios pequeños, cortos y heterogéneos que inflan el efecto.

**Tabla 8.** *Estimadores de efecto no expresados en desviaciones estándar*

| *Fuente* | *Desenlace* | *Estimador* |
|---|---|---|
| *Bastani et al. (6)* | Desempeño en práctica con GPT libre / con GPT tutor | +48 % / +127 % |
| *Bastani et al. (6)* | Desempeño en examen sin IA tras usar GPT libre | −17 % |
| *Chavez Sosa y Huancahuire-Vega (40)* | Ansiedad por punto de dependencia (ajustado) | RR 1,05 (IC 95 %: 1,01–1,09) |
| *Chavez Sosa y Huancahuire-Vega (40)* | Depresión por punto de dependencia (ajustado) | RR 1,04 (IC 95 %: 1,01–1,08) |
| *Chavez Sosa y Huancahuire-Vega (40)* | Estrés por punto de dependencia (ajustado) | RR 1,03 (IC 95 %: 1,00–1,07), no significativo |
| *Wang et al. (48)* | Estrés → dependencia → autoeficacia → agotamiento y ansiedad | Mediación secuencial significativa |
| *Nyakhar y Wang (47)* | Chatbots con mejora en ansiedad, depresión o bienestar | 8 de 9 estudios; deserción hasta 61 % |
| *Gallup y Walton Family Foundation (2)* | Tiempo ahorrado por docentes que usan IA semanalmente | 5,9 horas por semana |
| *Stephenson y Armstrong (25)* | Estudiantes que se sienten más solos / menos solos por la IA | 20 % / 21 % |
| *Wu et al. (38)* | Heterogeneidad del metaanálisis | Q = 409,1; $I^2$ = 91,4 % |

### 3.3. Aprendizaje: el diseño decide el signo del efecto

**Certeza: moderada.** La misma tecnología produjo efectos opuestos según el andamiaje. En Nigeria, 800 estudiantes que usaron GPT-4 en parejas, con supervisión docente y consignas que exigían razonar, ganaron 0,31 DE en seis semanas; en inglés, 0,23 DE equivalen a 1,5 a 2 años de escolaridad convencional (7). En física universitaria, un tutor diseñado con los mismos principios de la clase activa produjo mayor aprendizaje en menos tiempo (43).

En Turquía, el acceso libre a GPT-4 mejoró 48 % la práctica, pero redujo 17 % el desempeño cuando se retiró la herramienta. La versión con pistas diseñadas por docentes mejoró 127 % la práctica y mitigó esa pérdida (6). El metaanálisis vigente estima un efecto moderado, g = 0,670, moderado por asignatura, duración y modalidad (38).

En Nigeria, los mayores beneficios fueron para mujeres y para quienes ya rendían mejor. La IA puede ampliar brechas si no se diseña para quienes parten con desventaja.

### 3.4. Salud cognitiva: pensamiento crítico, memoria y atención

**Certeza: baja.** La señal es consistente, pero la evidencia causal es escasa.

- **Pensamiento crítico.** En 666 participantes, el uso de herramientas de IA se asoció negativamente con el pensamiento crítico, mediado por la descarga cognitiva; la asociación fue más fuerte en jóvenes y menor en quienes tenían más educación (49). En 319 trabajadores del conocimiento, mayor confianza en la IA predijo menos pensamiento crítico ejercido, y mayor confianza en sí mismo predijo más (50).
- **Memoria y autoría.** En un experimento con electroencefalografía (n = 54), quienes escribieron con un modelo de lenguaje mostraron la conectividad cerebral más débil, menor sentido de autoría y dificultad para citar su propio texto (41). El estudio es un preprint con muestra pequeña y críticas metodológicas publicadas sobre su análisis EEG (54).
- **Retención.** En enfermería, la IA mejoró destrezas procedimentales, pero en ocasiones comprometió la reflexión crítica y la retención a largo plazo frente a la instrucción humana

(39). La caída del 17 % al retirar GPT-4 en Turquía es la señal causal más sólida de pérdida de competencia autónoma (6).

- **Atención.** No se hallaron ensayos que midan la atención sostenida como desenlace. La OCDE (13) describe el riesgo de "pereza metacognitiva": la reducción del esfuerzo productivo que el aprendizaje exige. Esta es una brecha de investigación, no un hallazgo.

### 3.5. Salud mental y emocional: dependencia, ansiedad, agotamiento y soledad

**Certeza: baja para daños; baja a moderada para beneficios terapéuticos.** Los estudios son mayoritariamente transversales, pero convergen en un mecanismo común.

- **Dependencia, ansiedad y depresión.** En 187 estudiantes de medicina peruanos, la dependencia de IA generativa se correlacionó con ansiedad (ρ = 0,336) y depresión (ρ = 0,316). En modelos ajustados, cada punto adicional de dependencia se asoció con 5 % más de ansiedad (RR 1,05; IC 95 %: 1,01–1,09) y 4 % más de depresión (RR 1,04; IC 95 %: 1,01–1,08); la asociación con estrés no fue significativa (40).
- **El mecanismo: autoeficacia.** En 1.623 universitarios, el estrés académico se asoció con mayor dependencia de la IA; esta, con menor autoeficacia; y la menor autoeficacia, con más agotamiento académico y ansiedad (48). El estrés y la baja autoeficacia también predicen el uso problemático (51). Se configura un círculo: el estrés empuja a delegar, delegar quita experiencias de dominio, y la falta de dominio aumenta el estrés.
- **Soledad y dependencia emocional.** En un ensayo de cuatro semanas con 981 participantes y más de 300.000 mensajes, un mayor uso diario de chatbots se asoció con más soledad, dependencia emocional y uso problemático, y con menos socialización con personas reales, en todas las modalidades (42). Las ventajas iniciales de la voz se diluyeron con el uso intensivo.
- **Adolescentes.** El 72 % de los adolescentes estadounidenses ha usado compañeros de IA y el 52 % lo hace con regularidad (1). La APA (14) pide diseños que protejan las relaciones reales, la privacidad y el acceso a información de salud confiable.
- **Beneficios.** En una revisión rápida, ocho de nueve chatbots para universitarios mejoraron ansiedad, depresión o bienestar, sobre todo los basados en terapia cognitivo-conductual; la deserción alcanzó el 61 % (47). En enfermería, las simulaciones con IA ofrecieron práctica segura y sin juicio que redujo la ansiedad y aumentó la autoeficacia (44).

La IA actúa en ambas direcciones sobre la salud mental. El mismo rasgo, disponibilidad permanente sin juicio, alivia la ansiedad en una simulación diseñada y la alimenta cuando reemplaza el esfuerzo o la compañía humana.

### 3.6. Efectos por nivel formativo

La evidencia se concentra en pregrado y en la educación básica y media. En formación técnica, tecnológica y de posgrado no se hallaron estudios con desenlaces de salud.

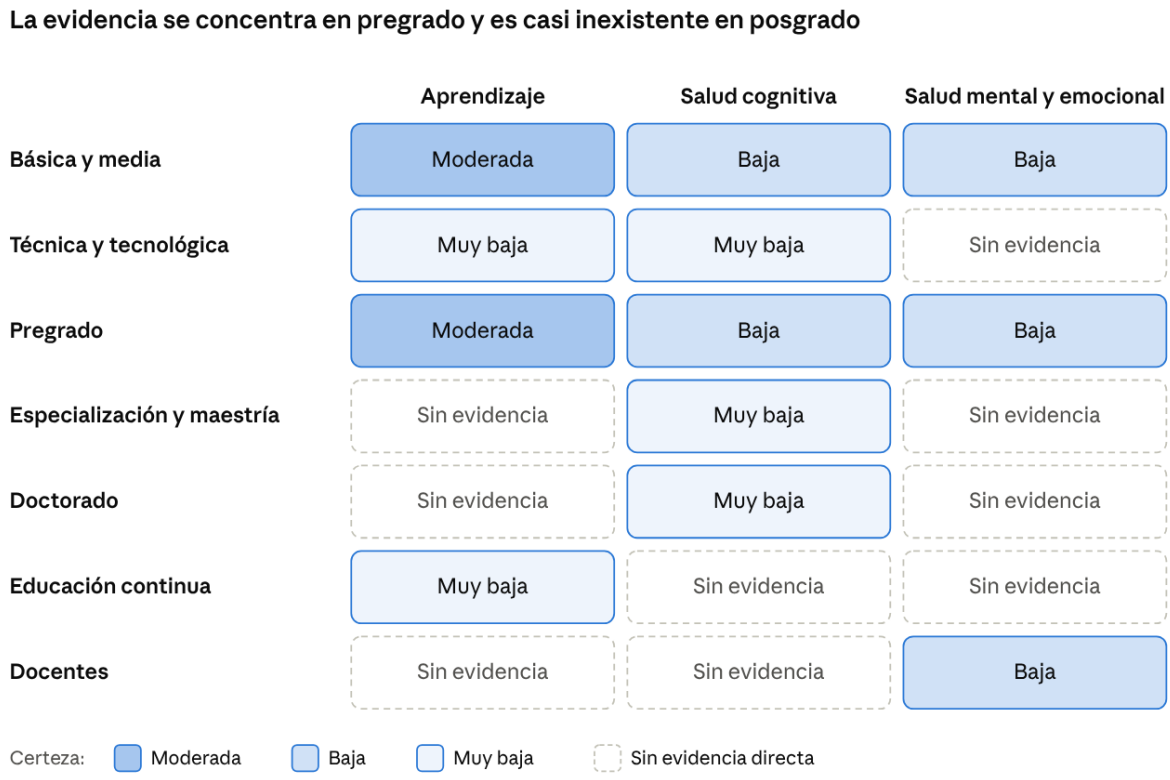
**Tabla 9.** *Síntesis por nivel formativo*

| *Nivel formativo* | *Evidencia directa* | *Riesgos para la salud identificados* | *Factores protectores* | *Certeza* |
|---|---|---|---|---|
| *Básica y media (6-17 años)* | Bastani et al. (6); De Simone et al. (7); Common Sense Media (1); APA (14) | Pérdida de competencia autónoma; sustitución de vínculos por | Supervisión docente; pistas en lugar de respuestas; edad mínima de 13 años (19) | Moderada (aprendizaje); baja (salud) |

| | | | | |
|---|---|---|---|---|
| | | compañeros de IA; exposición de datos | | |
| *Técnica laboral y tecnológica* | Ninguna específica; indirecta en simulación de enfermería (39,44) | Destreza procedimental sin reflexión; retención menor | Simulación diseñada como complemento de la práctica real | Muy baja |
| *Pregrado (salud, ingenierías, ciencias sociales y humanas)* | Kestin et al. (43); Gerlich (49); Wang et al. (48); Chavez Sosa y Huancahuire-Vega (40); Li et al. (46) | Dependencia asociada a ansiedad, depresión y agotamiento; menor pensamiento crítico | Alfabetización en IA; autoeficacia; chatbots terapéuticos con base en TCC | Baja a moderada |
| *Especialización y maestría* | Ninguna específica; indirecta en profesionales en ejercicio (50) | Delegación del juicio profesional cuando se confía más en la IA que en sí mismo | Confianza en el propio criterio; verificación sistemática | Muy baja |
| *Doctorado* | Ninguna específica; indirecta en escritura asistida (41) y detectores (52) | Menor autoría percibida; estrés por acusaciones de fraude con detectores sesgados | Declaración transparente de uso; evaluación oral de la tesis | Muy baja |
| *Educación continua (cursos, diplomados, MOOC)* | Indirecta (38,53) | Aprendizaje autónomo sin tutor: escenario de uso libre sin salvaguardas | Tutores configurados para guiar; actividades síncronas breves | Muy baja |
| *Docentes* | Gallup y Walton Family Foundation (2); Henttonen et al. (45) | Temor a la pérdida del rol profesional; presión por detectar fraude | 5,9 horas semanales ahorradas en tareas administrativas | Baja |

El ahorro de tiempo docente reproduce en el aula lo que la revisión previa observó en hospitales: automatizar tareas repetitivas libera tiempo y puede aliviar el agotamiento (4). La diferencia es que en el aula la tarea delegable no siempre es accesoria: a veces es el aprendizaje mismo.

**Figura 3.** *Mapa de evidencia: certeza de los hallazgos por nivel formativo y dominio de salud*




El mapa hace visible la desigualdad de la evidencia. Las celdas vacías no significan ausencia de efecto, sino ausencia de estudio: la formación técnica, tecnológica y de posgrado carece de datos sobre salud, pese a ser componente sustancial de la oferta de la educación abierta y a distancia. En el caso docente, la columna de salud mental se refiere al bienestar laboral y a la carga de trabajo.

### 3.7. Dimensión ética con efectos en salud

- **Detectores sesgados.** Más de la mitad de los textos de hablantes no nativos de inglés fueron clasificados como generados por IA, con precisión casi perfecta para nativos (52). Una acusación falsa de fraude es un estresor académico evitable.
- **Vigilancia emocional.** La Unión Europea prohibió desde febrero de 2025 el reconocimiento de emociones en instituciones educativas y clasificó como de alto riesgo la admisión, la calificación y la vigilancia de exámenes (16).
- **Datos de menores.** La APA (14) pide que los sistemas prioricen la privacidad y el bienestar de los adolescentes sobre el lucro comercial.

### 3.8. Colombia

Colombia tiene marco de política, pero no evidencia local sobre salud. El CONPES 4144 establece la Política Nacional de IA con 106 acciones a 2030 (15). En agosto de 2025 existían 865 programas activos en IA en 190 instituciones de educación superior (28). La Ley 2626 de 2026 incorpora la IA y la ciudadanía digital en básica y media, y su artículo 13 permite alianzas público-privadas cuyo acceso a datos de estudiantes depende del decreto reglamentario (12). En los documentos

revisados no se identificó una línea explícita sobre salud mental estudiantil asociada al uso de IA, ni estudios colombianos con desenlaces cognitivos o de salud mental.

### 3.9. Instrumentos de medición empleados en la evidencia

La comparabilidad entre estudios depende de los instrumentos. La Tabla 10 reúne los empleados en las fuentes primarias incluidas; su heterogeneidad explica parte de la dificultad para agregar resultados.

**Tabla 10.** *Instrumentos de medición utilizados en las fuentes primarias*

| *Constructo* | *Instrumento* | *Fuente que lo emplea* | *Observación* |
|---|---|---|---|
| *Dependencia de IA generativa* | Escala de Dependencia de la Inteligencia Artificial | Chavez Sosa y Huancahuire-Vega (40) | Autoinforme; puntaje continuo; aplicada en población peruana |
| *Dependencia de IA* | Escala de Dependencia de IA | Wang et al. (48) | Operacionaliza la dependencia como afrontamiento desadaptativo |
| *Uso problemático de IA* | Escala de uso problemático | Zhang et al. (51) | Enfocada en conducta, no en síntomas |
| *Ansiedad, depresión y estrés* | DASS-21 | Chavez Sosa y Huancahuire-Vega (40) | Tres subescalas de siete ítems |
| *Ansiedad generalizada* | GAD-7 | Wang et al. (48) | Tamizaje breve con puntos de corte establecidos |
| *Agotamiento académico* | Escala de Agotamiento Académico | Wang et al. (48) | Adaptación del síndrome de *burnout* al estudio |
| *Autoeficacia* | Escala de Autoeficacia General | Wang et al. (48) | Mediador central del modelo |
| *Pensamiento crítico* | Autorreporte de esfuerzo cognitivo por tarea | Lee et al. (50) | Mide pensamiento crítico ejercido, no capacidad |
| *Conectividad neural* | Electroencefalografía | Kosmyna et al. (41) | Conectividad funcional; interpretación discutida |
| *Soledad y dependencia emocional* | Escalas psicosociales longitudinales | Fang et al. (42) | Medición repetida durante cuatro semanas |
| *Aprendizaje* | Pruebas estandarizadas pretest y postest | De Simone et al. (7); Bastani et al. (6) | Desenlace objetivo, no autorreportado |

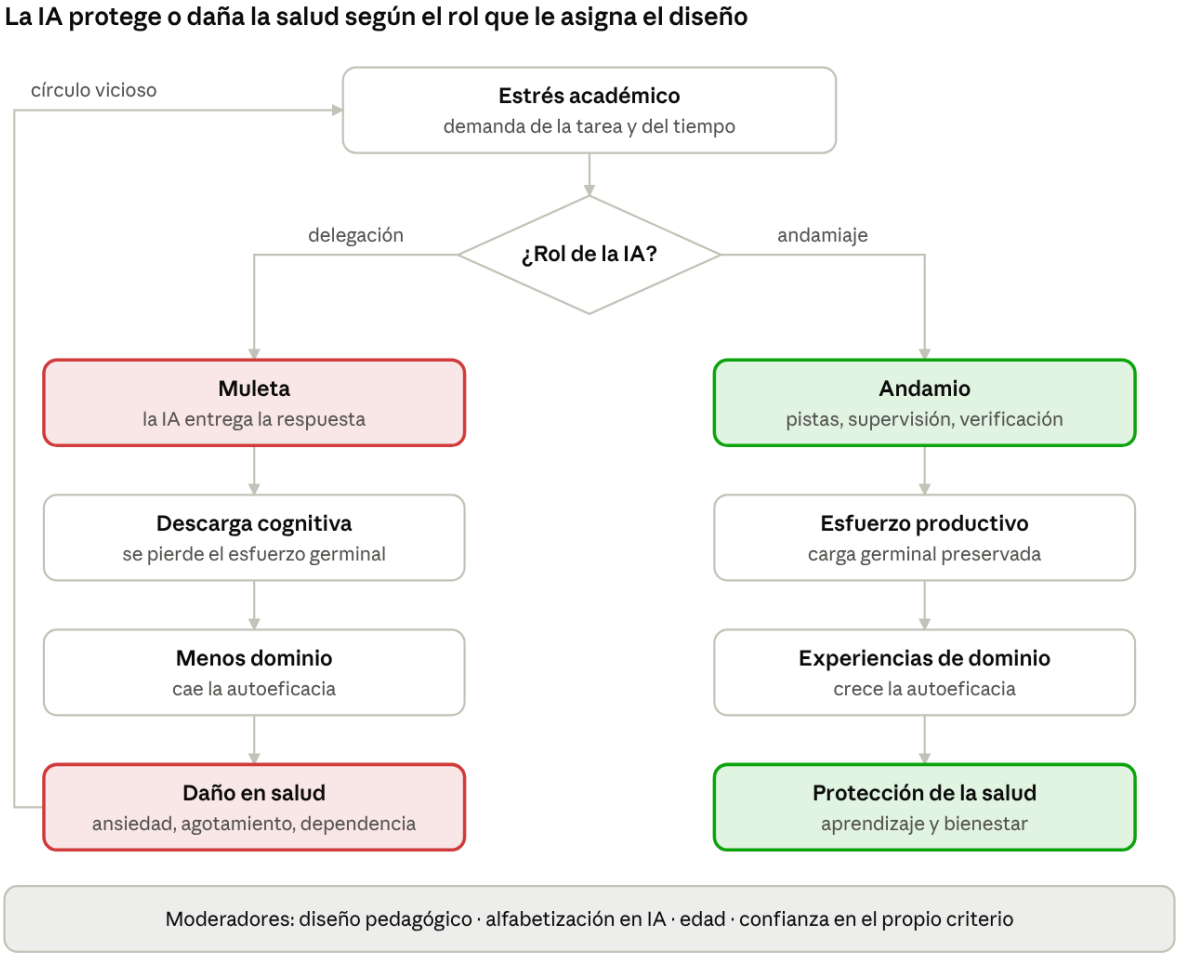
La mayoría de los desenlaces de salud son autorreportados. Faltan medidas objetivas de atención sostenida, pruebas de memoria diferida y marcadores de funcionamiento académico longitudinal.

## Discusión

### 4.1. Hallazgo central: la IA es andamio o muleta

La evidencia no permite afirmar que la IA mejore o dañe la salud de los estudiantes por sí misma. Lo que decide el efecto es el rol que el diseño le asigna. Se propone el Modelo de Andamiaje Protector para integrar los hallazgos.

**Figura 4.** Modelo de Andamiaje Protector: dos trayectorias de salud según el rol de la IA




En la trayectoria de muleta, el estrés empuja a delegar, la delegación elimina la carga germinal (21), la descarga reduce las experiencias de dominio que alimentan la autoeficacia (23), y la baja autoeficacia aumenta la ansiedad y el agotamiento, que a su vez elevan el estrés (48,51). En la trayectoria de andamio, la IA da pistas, exige razonar y se usa bajo supervisión; el esfuerzo se preserva y la salud se protege (6,7). El modelo es una hipótesis integradora: sus vínculos provienen de estudios distintos y requieren validación longitudinal.

### 4.2. Contraste con la literatura

El modelo es coherente con la teoría de conservación de recursos: recurrir a la IA bajo estrés protege recursos a corto plazo y los agota a largo plazo (24). También confirma el diagnóstico de Zawacki-Richter et al. (3): sin pedagogía explícita, la tecnología no produce aprendizaje. Y traduce a salud la distinción de la OCDE (13) entre reemplazo y complementariedad del juicio humano.

La quinta ola se acerca a la promesa de Bloom (9) solo cuando imita al buen tutor humano, que no entrega respuestas. La tutoría de Nigeria y la de física universitaria funcionaron porque trasladaron buenas prácticas pedagógicas al sistema, no por la potencia del modelo (43).

### 4.3. Implicaciones para la educación abierta y a distancia

La educación a distancia concentra tres rasgos del modelo. Primero, la menor presencia síncrona aumenta la exposición al uso libre sin salvaguardas. Segundo, la evaluación por textos escritos es la más vulnerable a la delegación y a los detectores sesgados (52). Tercero, la tutoría escalable con IA puede cubrir el acompañamiento que hoy es escaso (53). Para una universidad como la UNAD, el mismo recurso es el mayor riesgo y la mayor oportunidad.

### 4.4. La IA como determinante educativo de la salud

Los hallazgos permiten conceptualizar la IA generativa como un determinante intermedio de la salud estudiantil, en el sentido del marco de Solar e Irwin (32): no produce enfermedad por sí misma, pero modifica las condiciones psicosociales y conductuales del aprendizaje. Tres rasgos refuerzan esta lectura.

Primero, la relación parece graduada. En el ensayo de Fang et al. (42), los efectos psicosociales se asociaron con la intensidad del uso diario más que con la modalidad, un patrón compatible con una relación dosis-respuesta que la epidemiología reconoce como criterio de plausibilidad. Segundo, la exposición se distribuye de forma desigual: en el Reino Unido, los estudiantes de hogares con ocupaciones directivas o profesionales usan con más frecuencia algunas herramientas de IA (25), y en Nigeria los beneficios fueron mayores para quienes ya rendían mejor (7). Tercero, existen grupos de mayor vulnerabilidad: adolescentes, estudiantes con estrés académico elevado y estudiantes de ciencias de la salud en semestres clínicos, donde la ansiedad basal ya es alta (40).

De ahí se derivan tres niveles de prevención. La prevención primaria actúa sobre el diseño: asistentes en modo tutor, alfabetización en IA y lineamientos institucionales. La prevención secundaria detecta tempranamente la dependencia y los síntomas asociados mediante tamizaje en los servicios de bienestar. La prevención terciaria asegura la remisión oportuna a atención psicológica humana cuando la dependencia coexiste con ansiedad, depresión o aislamiento.

### 4.5. Propuesta de indicadores de vigilancia

Para que una institución gestione estos riesgos necesita medirlos. La Tabla 11 propone un conjunto mínimo de indicadores, construido con los instrumentos identificados en la sección 3.9. Los umbrales de alerta deben fijarse con líneas de base locales; no se proponen valores de referencia porque la evidencia no los respalda.

**Tabla 11.** *Indicadores propuestos para la vigilancia institucional de la salud estudiantil asociada al uso de IA*

| ***Dimensión*** | ***Indicador*** | ***Instrumento sugerido*** | ***Periodicidad*** |
|---|---|---|---|
| *Exposición* | Proporción de estudiantes que usan IA generativa en actividades evaluadas | Encuesta institucional análoga a Stephenson y Armstrong (25) | Semestral |
| *Exposición* | Proporción de usuarios diarios intensivos | Encuesta institucional | Semestral |
| *Riesgo* | Prevalencia de dependencia alta de IA | Escala de Dependencia de la Inteligencia Artificial | Semestral |
| *Salud mental* | Razón de prevalencias de ansiedad y depresión entre usuarios dependientes y no dependientes | GAD-7; DASS-21 | Semestral |
| *Protección* | Autoeficacia académica media | Escala de Autoeficacia General | Pretest y postest por curso |
| *Salud cognitiva* | Diferencia de desempeño entre evaluación con IA y evaluación síncrona sin IA | Evaluación oral o síncrona breve | Por periodo académico |

| | | | |
|---|---|---|---|
| *Capacidad institucional* | Proporción de docentes con formación en IA en el último año | Registro institucional, comparable con OCDE (26) | Anual |
| *Gobernanza* | Existencia, difusión y aplicación de lineamientos de IA | Auditoría documental | Anual |
| *Equidad y ética* | Proporción de procesos por presunto fraude que usaron un detector como prueba única | Registro de casos disciplinarios | Semestral |

### 4.6. De la evidencia a la práctica: formación en IA ética en la UNAD

El Modelo de Andamiaje Protector sugiere que la alfabetización en IA es la intervención de prevención primaria más factible para una universidad abierta. En esa línea, el autor facilita en la UNAD el programa Inteligencia Artificial Ética, Analítica y Aplicada en la Educación Abierta y a Distancia: ocho horas distribuidas en cuatro sesiones sincrónicas de dos horas por videoconferencia, con participantes de 15 a 90 años y evaluación mediante cuestionario pretest y postest.

El programa ofrece un escenario natural para probar el modelo. Incorporar al pretest y al postest la Escala de Dependencia de la Inteligencia Artificial, una escala de autoeficacia y el GAD-7 permitiría estimar si la formación en uso como andamio reduce la dependencia y protege la autoeficacia, con un diseño cuasiexperimental de bajo costo. La amplitud etaria del grupo aporta, además, una oportunidad poco frecuente para explorar diferencias por edad que la literatura apenas documenta. Sus resultados se reportarán en un estudio posterior.

### 4.7. Limitaciones

- La búsqueda sistematizada se hizo en una sola base (PubMed); las demás fuentes se obtuvieron por búsqueda autodirigida con modelos de IA avanzados.
- De 93 revisiones se tamizaron 20 por relevancia; no se alcanzó saturación y pudo omitirse evidencia.
- El tamizaje lo realizó con apoyo de tres asistentes de IA.
- Se incluyeron dos preprints (41,42) y predominan los estudios transversales, que no permiten inferir causalidad.
- Solo una fuente es latinoamericana y ninguna colombiana; la extrapolación al contexto local es limitada.
- El campo cambia con rapidez; los hallazgos reflejan la evidencia disponible al 27 de septiembre de 2026.

## Conclusiones

La IA nació en 1955 definida desde el aprendizaje y, setenta años después, su principal riesgo educativo es sustituirlo. La evidencia 2022-2026 muestra que el efecto de la IA sobre el aprendizaje y la salud depende del diseño: como andamio produce ganancias de aprendizaje y protege la autoeficacia; como muleta se asocia con pérdida de competencia autónoma, menor pensamiento crítico, ansiedad, depresión, agotamiento y soledad.

La certeza es moderada para el aprendizaje y baja para los desenlaces de salud. La evidencia se concentra en pregrado y en básica y media; la formación técnica, tecnológica y de posgrado carece de estudios, y Colombia no tiene evidencia propia pese a contar con marco de política.

## Recomendaciones

1. **Instituciones de educación superior, en especial las de modalidad a distancia.** Configurar asistentes institucionales en modo tutor (pistas, preguntas socráticas, verificación) en lugar de acceso libre; alinear la formación con los marcos de competencias de UNESCO (17,18); no usar detectores de IA como prueba única de fraude.
2. **Docentes.** Diseñar actividades donde la IA intervenga después del primer intento propio; incluir evaluación oral o síncrona breve; usar el tiempo ahorrado en tareas administrativas para acompañamiento.
3. **Estudiantes.** Usar la IA para pedir pistas y contrastar, no para entregar productos; reconocer el uso de IA bajo estrés como señal de alerta, no como solución.
4. **Bienestar universitario y salud pública.** Incorporar el tamizaje de dependencia de IA junto con escalas de ansiedad, depresión y agotamiento; articular chatbots basados en TCC con servicios humanos, nunca como reemplazo.
5. **Política pública colombiana.** Incluir en la reglamentación de la Ley 2626 de 2026 y en la ejecución del CONPES 4144 una línea de salud mental estudiantil y de protección de datos de menores.

## Agenda de investigación

- Ensayos controlados en Colombia que comparen IA en modo tutor frente a uso libre, con desenlaces de aprendizaje y salud.
- Estudios longitudinales que prueben la secuencia estrés, dependencia, autoeficacia y agotamiento del modelo propuesto.
- Medición de la atención sostenida como desenlace, hoy ausente en la literatura.
- Estudios en formación técnica, tecnológica, especialización, maestría y doctorado.
- Evaluación pretest y postest de programas de alfabetización en IA con escalas validadas de dependencia, ansiedad (GAD-7) y autoeficacia.

## Referencias

1. Common Sense Media. Talk, trust, and trade-offs: how and why teens use AI companions [Internet]. 2025. Disponible en: https://www.commonsensemedia.org/research/talk-trust-and-trade-offs-how-and-why-teens-use-ai-companions

2. Gallup, Walton Family Foundation. Teaching for tomorrow: unlocking six weeks a year with AI [Internet]. 2025. Disponible en: https://news.gallup.com/poll/691967/three-teachers-weekly-saving-six-weeks-year.aspx

3. Zawacki-Richter O, Marín VI, Bond M, Gouverneur F. Systematic review of research on artificial intelligence applications in higher education – where are the educators? International Journal of Educational Technology in Higher Education. 2019;16:39. doi: 10.1186/s41239-019-0171-0

4. Dazza Pineda CA, Garzón León FA. De los datos a la decisión: revisión sistemática del rol de la inteligencia artificial y el machine learning en la optimización de la gestión administrativa de instituciones de salud. Biociencias. 2026;9(1). doi: 10.22490/26194759.11707

5. McCarthy J, Minsky ML, Rochester N, Shannon CE. A proposal for the Dartmouth summer research project on artificial intelligence, August 31, 1955. AI Magazine. 2006;27(4):12-4. (Obra original publicada en 1955).

6. Bastani H, Bastani O, Sungu A, Ge H, Kabakcı Ö, Mariman R. Generative AI without guardrails can harm learning: evidence from high school mathematics. Proceedings of the National Academy of Sciences. 2025;122(26):e2422633122. doi: 10.1073/pnas.2422633122

7. De Simone ME, Tiberti FH, Barron Rodriguez MR, Manolio FA, Mosuro W, Dikoru EJ. From chalkboards to chatbots: evaluating the impact of generative AI on learning outcomes in Nigeria. Policy Research Working Paper No. 11125. World Bank; 2025.

8. Carbonell JR. AI in CAI: An artificial-intelligence approach to computer-assisted instruction. IEEE Transactions on Man-Machine Systems. 1970;11(4):190-202. doi: 10.1109/TMMS.1970.299942

9. Bloom BS. The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. Educational Researcher. 1984;13(6):4-16. doi: 10.3102/0013189X013006004

10. Pressey SL. A simple apparatus which gives tests and scores—and teaches. School and Society. 1926;23(586):373-376.

11. Skinner BF. Teaching machines. Science. 1958;128(3330):969-977. doi: 10.1126/science.128.3330.969

12. Congreso de la República de Colombia. Ley 2626 de 2026. Diario Oficial; 2026.

13. OCDE. OECD Digital Education Outlook 2026. París: OECD Publishing; 2026. doi: 10.1787/062a7394-en

14. American Psychological Association. Artificial intelligence and adolescent well-being: an APA health advisory [Internet]. 2025. Disponible en: https://www.apa.org/topics/artificial-intelligence-machine-learning/health-advisory-ai-adolescent-well-being

15. Departamento Nacional de Planeación. Documento CONPES 4144: Política nacional de inteligencia artificial. Bogotá: DNP; 2025.

16. Unión Europea. Reglamento (UE) 2024/1689 del Parlamento Europeo y del Consejo, de 13 de junio de 2024, por el que se establecen normas armonizadas en materia de inteligencia artificial. Diario Oficial de la Unión Europea, L, 12 de julio de 2024.

17. UNESCO. AI competency framework for students. París: UNESCO; 2024. doi: 10.54675/JKJB9835

18. UNESCO. AI competency framework for teachers. París: UNESCO; 2024. doi: 10.54675/ZJTE2084

19. Miao F, Holmes W. Guidance for generative AI in education and research. París: UNESCO; 2023. doi: 10.54675/EWZM9535

20. UNESCO. Recomendación sobre la ética de la inteligencia artificial. París: UNESCO; 2021.

21. Sweller J. Cognitive load during problem solving: Effects on learning. Cognitive Science. 1988;12(2):257-285. doi: 10.1207/s15516709cog1202_4

22. Risko EF, Gilbert SJ. Cognitive offloading. Trends in Cognitive Sciences. 2016;20(9):676-688. doi: 10.1016/j.tics.2016.07.002

23. Bandura A. Self-efficacy: Toward a unifying theory of behavioral change. Psychological Review. 1977;84(2):191-215. doi: 10.1037/0033-295X.84.2.191

24. Hobfoll SE. Conservation of resources: A new attempt at conceptualizing stress. American Psychologist. 1989;44(3):513-524. doi: 10.1037/0003-066X.44.3.51

25. Stephenson R, Armstrong C. Student generative AI survey 2026. HEPI Report 199 [Internet]. Higher Education Policy Institute; 2026. Disponible en: https://www.hepi.ac.uk/reports/student-generative-ai-survey-2026/

26. OCDE. Results from TALIS 2024: the state of teaching. París: OECD Publishing; 2025. doi: 10.1787/90df6235-en

27. UNESCO. UNESCO survey: two-thirds of higher education institutions have or are developing guidance on AI use [Internet]. 2025. Disponible en: https://www.unesco.org/en/articles/unesco-survey-two-thirds-higher-education-institutions-have-or-are-developing-guidance-ai-use

28. Ministerio de Educación Nacional. Oferta de programas de educación superior en inteligencia artificial [comunicado]. Bogotá: MEN; 2025.

29. Universidad de los Andes. Lineamientos para el uso de inteligencia artificial generativa (IAG) en la Universidad de los Andes [Internet]. 2024. Disponible en: https://www.uniandes.edu.co/es/noticias/inteligencia-artificial/un-documento-pionero-en-colombia-para-usar-la-ia-generativa

30. Concejo de Bogotá D. C. Acuerdo 1017 de 2025, por el cual se ordena la expedición de lineamientos para el uso responsable y ético de la inteligencia artificial en el Distrito Capital y se dictan otras disposiciones. Registro Distrital 8467; 2025.

31. Gómez-Orduz M, Hernández-Pérez JE, Pineda-Ángel JP, Mogollón-Benavides AL, compiladores. Lineamientos para el uso de la I.A. en publicaciones con el Sello Editorial UNAD [Internet]. Bogotá: Sello Editorial UNAD; 2024. Disponible en: https://selloeditorial.unad.edu.co/images/2024/02/19/LineamientosIAr.pdf

32. Solar O, Irwin A. A conceptual framework for action on the social determinants of health. Social Determinants of Health Discussion Paper 2. Ginebra: Organización Mundial de la Salud; 2010.

33. Deci EL, Ryan RM. The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. Psychological Inquiry. 2000;11(4):227-268. doi: 10.1207/S15327965PLI1104_01

34. Bjork RA. Memory and metamemory considerations in the training of human beings. En: Metcalfe J, Shimamura AP, editores. Metacognition: knowing about knowing. Cambridge (MA): MIT Press; 1994. p. 185-205.

35. Whittemore R, Knafl K. The integrative review: Updated methodology. Journal of Advanced Nursing. 2005;52(5):546-553. doi: 10.1111/j.1365-2648.2005.03621.x

36. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. BMJ. 2021;372:n71. doi: 10.1136/bmj.n71

37. Wang J, Fan W. Retraction note: The effect of ChatGPT on students' learning performance, learning perception, and higher-order thinking: Insights from a meta-analysis. Humanities and Social Sciences Communications. 2026;13:528. doi: 10.1057/s41599-026-07310-z

38. Wu X, Zhu P, Zhang J, Yin M, et al. ChatGPT's impact on student learning outcomes: A meta-analysis of 35 experimental studies. Humanities and Social Sciences Communications. 2026;13:684. doi: 10.1057/s41599-026-07019-z

39. Nasirun H, Mulyono S, Hapsari AWD, Wiradinata HH, Ratu M. Generative AI and intelligent tutoring systems in nursing education: A systematic review of impacts on higher-order thinking skills and clinical competency. Nurse Education in Practice. 2026;93:104847. doi: 10.1016/j.nepr.2026.104847

40. Chavez Sosa JV, Huancahuire-Vega S. Anxiety and depression associated with the dependent use of generative AI in medical students: Cross-sectional study. JMIR Formative Research. 2026;10:e82667. doi: 10.2196/82667

41. Kosmyna N, Hauptmann E, Yuan YT, Situ J, Liao XH, Beresnitzky AV, et al. Your brain on ChatGPT: accumulation of cognitive debt when using an AI assistant for essay writing task [preprint]. arXiv; 2025. doi: 10.48550/arXiv.2506.08872

42. Fang CM, Liu AR, Danry V, Lee E, Chan SWT, Pataranutaporn P, et al. How AI and human behaviors shape psychosocial effects of chatbot use: a longitudinal randomized controlled study [preprint]. arXiv; 2025. doi: 10.48550/arXiv.2503.17473

43. Kestin G, Miller K, Klales A, Milbourne T, Ponti G. AI tutoring outperforms in-class active learning: An RCT introducing a novel research-based design in an authentic educational setting. Scientific Reports. 2025;15:17458. doi: 10.1038/s41598-025-97652-6

44. Jiang H, Wang Z, Shen W, Meng M, Yang D, Li X, et al. AI-powered simulation for nursing education: Mixed methods systematic review. Journal of Medical Internet Research. 2026;28:e95167. doi: 10.2196/95167

45. Henttonen A, Christidis M, Kullenberg H, Sormunen T, Westerbotn M, Hägg-Martinell A. Nurse educators' experiences and perceptions using generative artificial intelligence: A systematic review. BMC Medical Education. 2026;26(1). doi: 10.1186/s12909-026-10113-0

46. Li J, Ai F, Wang J, Cheng B, Li Y, Chen Z. Application of AI-generated content in medical education: Systematic review of the impact on critical thinking abilities of medical students. JMIR Medical Education. 2026;12:e79939. doi: 10.2196/79939

47. Nyakhar S, Wang H. Effectiveness of artificial intelligence chatbots on mental health & well-being in college students: A rapid systematic review. Frontiers in Psychiatry. 2025;16:1621768. doi: 10.3389/fpsyt.2025.1621768

48. Wang W, Wu Y, Fang J, Yang C, Wen L. When cognitive offloading becomes dependence: How AI dependence mediates the pathway from academic stress to burnout and anxiety. BMC Psychology. 2026;14(1). doi: 10.1186/s40359-026-04911-w

49. Gerlich M. AI tools in society: Impacts on cognitive offloading and the future of critical thinking. Societies. 2025;15(1):6. doi: 10.3390/soc15010006

50. Lee HPH, Sarkar A, Tankelevitch L, Drosos I, Rintel S, Banks R, et al. The impact of generative AI on critical thinking: Self-reported reductions in cognitive effort and confidence effects from a survey of knowledge workers. En Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems [Internet]. ACM; 2025. doi: 10.1145/3706598.3713778

51. Zhang S, Zhao X, Zhou T, Kim JH. Do you have AI dependency? The roles of academic self-efficacy, academic stress, and performance expectations on problematic AI usage behavior. International Journal of Educational Technology in Higher Education. 2024;21:34. doi: 10.1186/s41239-024-00467-0

52. Liang W, Yuksekgonul M, Mao Y, Wu E, Zou J. GPT detectors are biased against non-native English writers. Patterns. 2023;4(7):100779. doi: 10.1016/j.patter.2023.100779

53. Bozkurt A, Xiao J, Lambert S, et al. Speculative futures on ChatGPT and generative artificial intelligence (AI): A collective reflection from the educational landscape. Asian Journal of Distance Education. 2023;18(1):53-130. doi: 10.5281/zenodo.7636568

54. Stanković M, Hirche E, Kollatzsch S, Doetsch JN. Commentary on Kosmyna et al. (2025), Your brain on ChatGPT [preprint]. arXiv; 2026. doi: 10.48550/arXiv.2601.00856