https://doi.org/10.69639/arandu.v13i4.2691

Cuando la IA responde antes de que aprendamos: Inteligencia artificial generativa, descarga cognitiva y desaparición del esfuerzo intelectual

When AI responds before we learn: Generative artificial intelligence, cognitive offloading, and the disappearance of intellectual effort

Francisco Xavier Rosero Aulestia
panchorosero9612@gmail.com
https://orcid.org/0009-0003-1007-4281
Investigador Independiente
Ibarra – Ecuador

Carmen Genoveva Díaz Paredes
teachercarmita.1@gmail.com
https://orcid.org/0000-0003-4567-6250
Investigador Independiente
Quito – Ecuador

Álvaro Isaías Quinatoa Donoso
alvaroiquinatoa@gmail.com
https://orcid.org/0009-0001-3778-887X
Investigador Independiente
Ibarra, Ecuador

Henry Daniel Muñoz Papa
henda1989@gmail.com
https://orcid.org/0009-0005-4334-2718
Investigador Independiente
Quito, Ecuador

Christian Fernando Poma Sandoya
cf_ps31@hotmail.com
https://orcid.org/0009-0005-8746-8312
Investigador Independiente
Quito, Ecuador

Artículo recibido: 10 septiembre 2026- Aceptado para publicación:16 octubre 2026
Conflictos de intereses: Ninguno que declarar.

RESUMEN

La inteligencia artificial generativa puede ofrecer una respuesta completa sin que el estudiante haya comprendido las operaciones que la sostienen. Se examinan la descarga cognitiva y la posible desaparición del esfuerzo intelectual mediante la relación entre rendimiento asistido y desempeño posterior. Se realizó una revisión crítica con búsqueda focalizada de siete estudios de matemáticas, escritura, programación, indagación científica y tutoría universitaria. Cuatro midieron por separado el rendimiento asistido. Tres de ellos documentaron una evaluación posterior sin asistencia generativa. Dos incluyeron retención diferida. Las ventajas asistidas

coexistieron con desventajas posteriores, diferencias no detectadas o resultados dependientes del diseño. La correspondencia no automática recibe respaldo documental acotado. Los registros de proceso permiten examinar delegación y sustitución de actividad, sin establecer desaparición del esfuerzo intelectual. Los indicios sobre conservación de práctica no aíslan los componentes de la ayuda. El modelo de delegación, esfuerzo residual, rendimiento asistido y aprendizaje independiente (DERA) organiza las exigencias de evidencia y permanece pendiente de validación. Se concluye que disponer de una respuesta y haber aprendido a construirla requieren evidencia distinta. La posible ausencia de una operación central debe comprobarse en la actividad del estudiante, además de evaluar el producto y el desempeño posterior.

Palabras clave: aprendizaje, cognición, pensamiento crítico, proceso de aprendizaje, tecnología educacional

ABSTRACT

Generative artificial intelligence can provide a complete answer without the student having understood the operations that support it. Cognitive offloading and the possible disappearance of intellectual effort are examined through the relationship between assisted performance and subsequent performance. A critical review with a focused search was conducted on seven studies in mathematics, writing, programming, scientific inquiry, and university tutoring. Four separately measured assisted performance. Three of them documented a subsequent evaluation without generative assistance. Two included deferred retention. The assisted advantages coexisted with subsequent disadvantages, undetected differences, or design-dependent outcomes. Non-automatic correspondence receives limited documentary support. Process records allow for the examination of delegation and substitution of activity, without establishing the disappearance of intellectual effort. The indications regarding the preservation of practice do not isolate the components of assistance. The delegation, residual effort, assisted performance, and independent learning (DERA) model organizes the evidence requirements and remains pending validation. It is concluded that having a response and having learned to construct it require different evidence. The possible absence of a central operation must be verified in the student's activity, in addition to evaluating the product and subsequent performance.

Keywords: learning, cognition, critical thinking, learning process, educational technology

Todo el contenido de la Revista Científica Internacional Arandu UTIC publicado en este sitio está disponible bajo licencia Creative Commons Atribution 4.0 International.

INTRODUCCIÓN

Un estudiante puede recibir una respuesta bien construida y tener dificultades para explicar por qué sus argumentos sostienen la conclusión. También puede resolver ejercicios con asistencia y encontrar obstáculos ante un problema semejante cuando esa ayuda termina. La calidad del producto y la competencia de quien lo presenta son dimensiones relacionadas, pero su correspondencia debe comprobarse. Cuando una herramienta interviene en la respuesta, el resultado incorpora operaciones de la persona y operaciones realizadas externamente. El problema aparece cuando disponer de la respuesta se confunde con haber aprendido a construirla.

La inteligencia artificial generativa (IA) amplía esta dificultad de atribución. Puede corregir una expresión, recuperar información, organizar un texto o elaborar un razonamiento completo. Esas intervenciones cumplen funciones educativas distintas. Si una actividad pretende enseñar a argumentar, automatizar la presentación de las referencias puede modificar una tarea auxiliar. Delegar la elección de la evidencia y la construcción de la inferencia puede modificar la práctica mediante la cual se aprende a argumentar. El significado educativo depende de la competencia que se pretende adquirir.

Marcelo (2013) examinó la integración de las tecnologías a partir de las personas, la innovación y su contexto. Su análisis sitúa el cambio educativo en las prácticas de enseñanza, sin atribuirlo a la sola presencia de dispositivos. En el problema aquí planteado, disponer de una herramienta y modificar las condiciones para aprender son aspectos que requieren análisis separado.

Díaz Barriga (2003) sitúa el conocimiento en la actividad, el contexto y las prácticas en que se aprende. Su planteamiento permite examinar la ayuda como parte de una actividad formativa y preguntar qué hace el estudiante con ella. Recibir información y participar en su elaboración requieren observaciones distintas.

Esta separación tiene antecedentes. Salomon et al. (1991) distinguieron los efectos que aparecen al trabajar con tecnología de los efectos que permanecen cuando la persona actúa alejada de ella. Soderstrom y Bjork (2015) examinaron la diferencia entre rendimiento durante la adquisición y aprendizaje manifestado en retención o transferencia. La IA generativa vuelve especialmente relevante esa distinción porque puede intervenir en operaciones que forman parte de la capacidad evaluada.

Las síntesis recientes también muestran la necesidad de precisar las medidas. Deng et al. (2025) señalaron problemas de evaluación del desempeño, uso de autoinformes y seguimiento. Mo et al. (2026) integraron 66 estudios y 129 tamaños de efecto y reportaron una estimación positiva de g de Hedges = 1,14. Ese promedio describe los resultados incluidos en su análisis. No establece que todas las modalidades de asistencia produzcan el mismo efecto ni identifica qué operaciones puede ejecutar después cada estudiante sin la herramienta.

Descarga cognitiva y función del esfuerzo

Risko y Gilbert (2016) examinaron la descarga cognitiva como uso de apoyos externos que modifican las exigencias de procesamiento. Externalizar una operación puede ampliar lo que una persona consigue hacer. En educación, interesa además determinar si esa operación debe practicarse para adquirir la competencia. Corregir ortografía puede ser auxiliar al evaluar razonamiento jurídico y central al enseñar ortografía. La centralidad es una relación entre operación, tarea, conocimiento previo y objetivo formativo.

En este artículo se denomina descarga funcional a la externalización de operaciones cuya ejecución personal no es necesaria para el objetivo definido. La sustitución cognitiva se plantea cuando la herramienta asume una operación necesaria para construir la competencia y el estudiante reduce su práctica o reconstrucción. El acceso a una solución completa ofrece una posibilidad de delegación. La actividad posterior permite examinar cómo se utilizó. La operación debe clasificarse según el objetivo antes de conocer el resultado, para evitar llamar perjudicial a una delegación únicamente porque el postest fue menor.

Fuertes Alpiste (2024) propone entender las aplicaciones generativas como herramientas para la cognición y conserva las funciones ejecutivas en el estudiante. Ese criterio teórico centra el análisis en quién organiza, decide y evalúa durante la tarea. La presencia de una herramienta abre una posibilidad de colaboración, cuya realización requiere examinar la actividad.

Tampoco todo esfuerzo tiene valor formativo. Sweller (1988) distinguió exigencias de resolución de problemas y adquisición de esquemas. Andrade-Lotero (2012) examinó el diseño de la información, los ejemplos resueltos y el conocimiento previo como condiciones para interpretar las demandas de aprendizaje. Una explicación externa puede facilitar comprensión si el estudiante estudia sus relaciones. Roediger y Karpicke (2006) encontraron beneficios de la práctica de recuperación para la retención posterior. Estos antecedentes justifican examinar la función de cada actividad, en lugar de valorar toda reducción de dificultad como pérdida de aprendizaje.

El esfuerzo intelectual comprende operaciones de comprensión, verificación, recuperación y aplicación. Tiempo de trabajo, número de mensajes y dificultad percibida son indicadores parciales. Una sesión extensa puede incluir distracción. Una explicación breve puede requerir razonamiento considerable. La revisión distingue esos indicadores de las operaciones cuya realización se pretende inferir.

Como hipótesis de análisis, la desaparición del esfuerzo intelectual se refiere aquí a la ausencia de realización propia de una operación central durante un episodio de trabajo asistido. Exige comprobar que el estudiante tampoco la reconstruyó ni verificó después de recibir la respuesta. Se delimita por la operación y el episodio, porque el esfuerzo puede mantenerse en otras decisiones. Esta hipótesis necesita evidencia de lo que la persona hizo y dejó de hacer. Un registro incompleto no permite distinguir ausencia de actividad de ausencia de observación.

López Aymes (2012) vinculó la autonomía intelectual con las habilidades de pensamiento crítico y la disposición para emplearlas. En una tarea asistida, interesa comprobar si esas habilidades intervienen en las decisiones del estudiante.

Comprensión percibida y competencia demostrada

Comprender y reconocer si se ha comprendido son problemas relacionados. Osses Bustingorry y Jaramillo Mora (2008) distinguieron el conocimiento sobre los propios procesos cognitivos del control que permite planificar, supervisar y evaluar su realización. Esta perspectiva permite examinar si el estudiante identifica sus dificultades y revisa el resultado. Recibir una explicación constituye una condición de la actividad, mientras comprobar qué se puede explicar o aplicar permite valorar su comprensión.

La satisfacción y la seguridad subjetiva informan sobre la experiencia, pero requieren contraste con una actuación propia. Panadero y Alonso-Tapia (2014) analizaron la autorregulación mediante las estrategias, las metas y las condiciones personales y del entorno que intervienen en el aprendizaje. En una interacción con IA, enviar muchos mensajes puede coexistir con adoptar respuestas sin reconstrucción. Revisar pocas sugerencias puede exigir evaluación sustantiva. La frecuencia de uso deja abierta la pregunta sobre lo que el estudiante hizo con la información recibida.

Paciencia cognitiva y condiciones del aprendizaje

Rosero Aulestia et al. (2026) delimitaron la paciencia cognitiva mediante el esfuerzo orientado a comprender durante un intervalo en que el resultado permanece diferido e incierto. La propuesta distingue esa demora de la dificultad y sitúa parte de la explicación en la estructura temporal de la tarea. También separa cumplir una actividad de comprender su contenido. Esta delimitación permite examinar una condición que la respuesta inmediata vuelve relevante: qué procesamiento se conserva mientras el estudiante todavía no comprende, aunque ya disponga de una solución.

Las condiciones del entorno tampoco se reducen a las características del estudiante. Rosero Aulestia et al. (2026) examinaron cómo el tratamiento tarifario de las aplicaciones modifica las condiciones económicas de exploración de espacios digitales. Su categoría de currículo invisible tarifario describe un condicionamiento del acceso, sin demostrar efectos causales sobre el aprendizaje. El antecedente corresponde a otro problema educativo. Se incorpora porque permite distinguir las condiciones externas de la actividad de los resultados cognitivos que todavía deben observarse.

Planteamiento de la investigación

El vacío que aborda esta revisión es analítico y metodológico: precisar qué evidencia permite distinguir reducción de demanda, sustitución de actividad y posible desaparición de una operación central. El esfuerzo declarado, la operación delegada y el desempeño posterior remiten a dimensiones diferentes. Su comparación requiere conservar qué representa cada indicador. La

revisión delimita ese problema en el corpus examinado, sin afirmar una ausencia general de investigaciones sobre el tema.

La pregunta de investigación es: ¿qué evidencia aportan los estudios seleccionados sobre la posible desaparición del esfuerzo intelectual en tareas asistidas por IA generativa y sobre su relación con el rendimiento asistido y el desempeño posterior? Su formulación delimita una investigación documental. No presupone que la revisión pueda estimar un efecto universal o identificar por sí sola un mecanismo causal.

El objetivo general es examinar críticamente la descarga cognitiva y la posible desaparición del esfuerzo intelectual mediante la comparación del rendimiento asistido, el procesamiento propio y el desempeño posterior. El artículo tiene cuatro objetivos específicos. El primero consiste en describir los diseños, comparadores, medidas y condiciones de evaluación. El segundo busca contrastar los resultados durante la asistencia y después de ella, diferenciando actuación inmediata, retención y transferencia. El tercero examina qué evidencia de delegación y actividad propia permite distinguir reducción de demanda, sustitución y ausencia de una operación central. El cuarto pretende formular reglas que integren esas exigencias en el modelo DERA.

El análisis parte de tres hipótesis. La primera sostiene que una mejora del rendimiento asistido no asegura una mejora correspondiente del desempeño independiente. La segunda plantea que la delegación de una operación central, acompañada de ausencia de realización, reconstrucción y verificación propia, puede relacionarse con menor desempeño independiente. La tercera considera que los diseños de asistencia que conservan intentos, explicaciones y revisión del estudiante pueden limitar la sustitución de esa actividad y la divergencia entre rendimiento asistido y desempeño posterior. Las hipótesis se refieren, respectivamente, a la correspondencia de resultados, a la posible desaparición de una operación central y a una condición pedagógica.

Las hipótesis orientan el contraste documental y admiten resultados contrarios o insuficientes. No se someten aquí a una prueba estadística conjunta. La investigación busca precisar qué respalda cada estudio, qué explicaciones alternativas conserva y qué información falta. DERA reúne delegación, esfuerzo residual, rendimiento asistido y aprendizaje independiente. Su contribución es una matriz de interpretación. La distinción entre rendimiento y aprendizaje cuenta con antecedentes y también aparece en investigaciones recientes como Chen (2026).

La revisión se orienta, por tanto, a examinar la descarga cognitiva y la posible desaparición del esfuerzo intelectual mediante la comparación de la actividad propia, el rendimiento asistido y el desempeño posterior.

METODOLOGÍA

Diseño y búsqueda documental

Se realizó una revisión crítica con búsqueda focalizada y síntesis comparativa. Guirao Goris (2015) describe el análisis crítico y la integración conceptual como propósitos de este tipo de revisión. Aquí se aplicó una selección intencional de casos informativos, sin un protocolo de recuperación exhaustiva. El período de búsqueda informado para esta investigación comprende del 10 de julio al 15 de septiembre de 2026. Se consideraron estudios sobre IA generativa publicados desde 2022 hasta el cierre de ese período. Los antecedentes conceptuales se consultaron sin esa restricción temporal. La identidad bibliográfica y las cifras se comprobaron durante la revisión del manuscrito mediante documentos primarios, páginas editoriales, repositorios de texto completo e identificadores DOI.

La localización partió de los cuatro trabajos identificados en el esquema inicial. Su procedencia se distingue en el apéndice y sus diseños y resultados se presentan en la tabla 1. Se amplió con programación, indagación científica y tutoría para incorporar contrastes y explicaciones alternativas. Los términos se relacionaron con descarga cognitiva, rendimiento asistido, aprendizaje independiente, retención y tutoría, junto con nombres de autores y características del diseño. Se utilizaron también sus equivalentes en inglés para localizar publicaciones en ese idioma. El apéndice identifica las versiones y vías de acceso efectivamente verificadas. Las comprobaciones por título o DOI no se contaron como estudios nuevos. La revisión no aplicó un protocolo de recuperación exhaustiva o de selección sistemática.

Selección y unidades de análisis

Se incluyeron siete estudios empíricos con una intervención de IA generativa en una actividad de aprendizaje, comparador identificable y medidas pertinentes para alguna dimensión DERA. Se excluyeron de la comparación principal trabajos centrados únicamente en satisfacción, aceptación o intención de uso. Revisiones y textos teóricos se utilizaron como antecedentes. La selección fue intencional. Privilegió casos informativos y no constituye una muestra representativa de toda la literatura.

La unidad documental fue el estudio. La unidad de comparación fue el contraste entre condiciones y momentos de evaluación dentro de cada estudio. Se distinguieron participantes, casos analizados, clases y observaciones repetidas. Dos versiones de una publicación se trataron como el mismo trabajo. Las hipótesis no se pre registraron y la ampliación del corpus respondió al propósito crítico del análisis, por lo que se conserva el riesgo de selección ligado al punto de partida. Los trabajos sobre paciencia cognitiva y currículo invisible tarifario se incorporaron como antecedentes para contrastar las condiciones temporales y materiales de la actividad. No se contaron como investigaciones empíricas adicionales sobre rendimiento asistido o desempeño posterior ni como validaciones de DERA.

Extracción y reglas de clasificación

Se extrajeron población, asignación, actividad, modalidad de asistencia, comparador, resultado durante el uso, evaluación posterior, intervalo temporal e indicadores de proceso. Se revisaron pérdidas de seguimiento, dependencia por clases o personas, comparabilidad de instrumentos y recursos permitidos en la prueba. La información de Barcaui (2025) se recuperó parcialmente desde secciones editoriales de método, resultados y limitaciones. No se consultó su archivo completo. El estudio de Fan et al. (2025) se examinó en versión de autor y se contrastó bibliográficamente con la publicación editorial. No se obtuvo el suplemento de Kestin et al. (2025).

La matriz documental aplicó tres reglas. Se registró rendimiento asistido separado cuando existía una puntuación de la tarea realizada durante el acceso a IA, distinta del postest. La retirada se consideró documentada cuando el procedimiento describía expresamente la ausencia del apoyo generativo en la evaluación. Un detalle insuficiente se registró como no establecido. Una prueba supervisada conocida mediante acceso parcial conservó esa condición. Se registró retención diferida cuando existía un intervalo de seguimiento posterior a la actividad de aprendizaje. La semana final de una intervención no se convirtió automáticamente en un intervalo de retención.

Las operaciones delegadas y el procesamiento propio se describieron según la evidencia disponible: interacciones, versiones, conducta observada, autoinforme o inferencia del investigador. En su interpretación se distinguieron reducción de demanda, sustitución de actividad y ausencia documentada de una operación central. La desaparición exigía comprobar que esa operación no se realizó ni se reconstruyó ni se verificó durante el episodio. Una menor carga o duración no satisfizo por sí sola ese criterio. Una medida ausente se registró como no determinada. No se asignó esfuerzo cero por falta de información. Los resultados de aprendizaje reportados con control de asistencia insuficientemente descrito se conservaron como postest, sin clasificarlos automáticamente como aprendizaje independiente demostrado.

Síntesis y alcance del contraste

La síntesis conservó unidades, comparadores e incertidumbre originales. Los recuentos describen únicamente los siete documentos. Las comprobaciones aritméticas se limitaron a diferencias, cambios relativos y pérdidas calculables desde datos publicados. No se combinaron tamaños de efecto ni se realizaron meta regresiones. La heterogeneidad de constructos, tareas y tiempos impidió definir una brecha numérica común entre rendimiento y aprendizaje.

Para cada hipótesis se identificaron observaciones compatibles, resultados que impiden una interpretación uniforme y requisitos todavía ausentes. Una diferencia no detectada no se interpretó como equivalencia. Se separaron efecto de la intervención completa, mecanismo y generalización. La selección y la extracción no fueron duplicadas de forma independiente y no se estimó acuerdo entre evaluadores. Las dimensiones DERA procedían del esquema conceptual

inicial. Sus reglas se precisaron mediante el contraste documental y permanecen pendientes de validación.

Consideraciones éticas

El análisis utilizó documentos publicados y resultados agregados. No se realizaron intervenciones propias con estudiantes ni se trabajó con datos individuales identificables. Se conservaron la atribución de las ideas, la identidad de las fuentes y los límites de acceso. Los cálculos propios se distinguen de las estadísticas reportadas por cada investigación. No se reconstruyeron datos individuales a partir de los resultados publicados.

RESULTADOS

Características del corpus y condiciones de comparación

La investigación documental reunió siete estudios con comparadores en cinco ámbitos de actividad. La tabla 1 presenta sus resultados y límites. Este apartado comunica hallazgos publicados y resultados de la extracción comparativa. No incorpora mediciones nuevas de estudiantes. La descripción distingue diseño, unidad, asistencia y momento de evaluación, en lugar de reunir todas las puntuaciones bajo una sola medida de aprendizaje.

Tabla 1

Estudios seleccionados y resultados documentados

Estudio y muestra	Diseño y asistencia	Resultado durante la asistencia	Resultado posterior y límite principal
Bastani et al. (2025). Cerca de 1.000 estudiantes de secundaria, Turquía	Asignación por aulas. Matemáticas con GPT Base, GPT Tutor o recursos habituales	Mejoras relativas de 48 % y 127 %, respectivamente, frente al control	GPT Base: desventaja relativa de 17 % en examen inmediato sin recursos. GPT Tutor: diferencia no detectada frente al control. Salvaguardas combinadas. Sin retención prolongada
Fan et al. (2025). 117 universitarios	Cuatro condiciones aleatorizadas de revisión escrita. ChatGPT limitado a asesorar	Mayor mejora del ensayo con ChatGPT	Conocimiento y transferencia: diferencias no detectadas. Pruebas dentro de un día. Control de asistencia insuficientemente descrito en el material consultado. Constructos diferentes
Kazemitabaar et al. (2023). 90 participantes iniciales. 69 completaron, de 10 a 17 años	Asignación aleatoria dentro de pares. Generación con Codex y modificación manual	Ventajas en elaboración de código	Exactitud global sin diferencias detectadas al día siguiente y una semana después. En retención, más errores en autoría con Codex: 1,58 frente a 0,99, p = .042. Pérdidas y tareas acotadas

Estudio y muestra	Diseño y asistencia	Resultado durante la asistencia	Resultado posterior y límite principal
Chen (2026). 180 participantes iniciales. 168 analizados en seis clases	Ocho semanas. Apoyo limitado con reflexión, colaboración abierta o ausencia de IA	Mayor media con colaboración abierta. Comparación con apoyo limitado imprecisa por clases	Tarea supervisada sin IA en semana 8: patrón favorable al apoyo limitado. Pruebas exactas por conglomerados: *p* entre .050 y .063. Condiciones combinadas y pocas clases
Barcaui (2025). 120 participantes iniciales. 85 evaluados	Asignación aleatoria. Estudio con ChatGPT o medios tradicionales	Sin puntuación formal separada del producto inicial	Prueba supervisada a los 45 días: 57,5 % frente a 68,5 %. *d* = 0,68 reportado a favor del control. Pérdidas de 29,2 % y consulta documental parcial
Stadler et al. (2024). 92 participantes iniciales. 91 analizados	ChatGPT 3.5 frente a Google. Indagación de 20 minutos	Menor carga auto informada. Sin puntuación separada del producto asistido	Justificaciones sin notas de la búsqueda o conversación: 1,20 frente a 1,87 aspectos pertinentes, *p* = .001. Evaluación inmediata y específica. Mecanismo no aislado
Kestin et al. (2025). Curso de 194 estudiantes. 316 evaluaciones	Diseño cruzado. Asignación por pequeños grupos. Tutor frente a clase activa	Sin resultado asistido comparable separado del postest	Mejores medianas de postest con tutor: 4,5 frente a 3,5. Control de asistencia no plenamente establecido en el material recuperado. Evaluación inmediata y tratamiento combinado

Nota. Las cifras proceden de las fuentes citadas. Los porcentajes de Bastani et al. (2025) son cambios relativos al control. Las 316 evaluaciones de Kestin et al. (2025) no son 316 personas. Las diferencias no detectadas conservan incertidumbre y no demuestran equivalencia. Las versiones y los límites de acceso constan en el apéndice.

La clasificación de la tabla 2 identificó cuatro estudios con una medida separada del resultado durante la asistencia. Tres de esos cuatro describen además retirada explícita de IA en una evaluación posterior. Stadler et al. (2024) también documentan una actuación sin el recurso generativo, pero su estudio carece de una puntuación separada del producto asistido. Barcaui (2025) aporta una prueba supervisada conocida mediante recuperación parcial. En Fan et al. (2025), la independencia no quedó suficientemente establecida en los materiales consultados. Tampoco quedó establecida en el material recuperado de Kestin et al. (2025).

Tabla 2

Disponibilidad de medidas para contrastar rendimiento y desempeño posterior

Estudio	Producto asistido separado	Condición posterior documentada	Momento y retención diferida
Bastani et al. (2025)	Sí	Sin IA ni recursos durante el examen	Misma sesión. Sin seguimiento diferido
Fan et al. (2025)	Sí	Retirada de IA no suficientemente establecida	Dentro de un día. Sin seguimiento diferido definido
Kazemitabaar et al. (2023)	Sí	Sin generador ni documentación	Día siguiente y una semana. Retención diferida reportada
Chen (2026)	Sí	Tarea supervisada sin IA	Semana 8 al terminar la intervención. No ocho semanas de retención
Barcaui (2025)	No	Prueba supervisada. Acceso documental parcial	45 días. Retención diferida reportada
Stadler et al. (2024)	No	Justificación sin ChatGPT, notas ni páginas de búsqueda	Inmediata. Sin seguimiento diferido
Kestin et al. (2025)	No	Retirada de IA no plenamente establecida	Postest inmediato. Sin seguimiento diferido

Nota. Elaboración a partir de los métodos publicados. Sí requiere una puntuación durante el acceso, distinta del postest. La clasificación refleja lo documentado y no imputa condiciones ausentes. Un estudio puede aportar evidencia pertinente sin permitir un contraste completo entre rendimiento asistido y actuación independiente.

Dos estudios del corpus incluyeron una prueba de retención diferida, como registra la tabla 2. El registro de la semana 8 en Chen (2026) corresponde al final de ocho semanas de intervención, no a un seguimiento de ocho semanas después de terminarla. Estos recuentos son resultados descriptivos del corpus seleccionado. No estiman la frecuencia de esos diseños en toda la literatura sobre IA.

Correspondencia entre rendimiento asistido y desempeño posterior

El contraste de resultados identificó patrones que requieren interpretaciones distintas. Bastani et al. (2025) documentaron una ventaja durante la práctica y una desventaja posterior bajo GPT Base. El cambio de signo dentro del mismo diseño permite distinguir ejecución asistida y actuación individual. Con GPT Tutor, el contraste posterior frente al control no detectó una diferencia. Esa ausencia de detección no establece equivalencia ni una ganancia independiente.

En programación, Kazemitabaar et al. (2023) encontraron ventajas durante la elaboración de código y diferencias globales no detectadas en exactitud posterior. El indicador de errores en

autoría durante retención siguió otro patrón, como muestra la tabla 1, y coexistió con más tareas intentadas por el grupo Codex. Por ello, los errores no se interpretaron como prueba aislada de menor conocimiento. Exactitud, errores y tareas completadas conservaron sus diferencias. La generación coexistió con modificación manual, lo que mantiene una explicación alternativa a la sustitución completa de la práctica.

Fan et al. (2025) aportaron una mejora del ensayo sin diferencias detectadas en conocimiento y transferencia. El resultado compara constructos distintos y conserva incertidumbre sobre los recursos de la prueba. Chen (2026) mostró una ordenación de medias diferente durante la asistencia y en la tarea final sin IA. La incertidumbre por seis clases impide convertir esa ordenación en una estimación causal precisa de cada componente pedagógico.

Los tres estudios sin puntuación separada del producto asistido aportaron información complementaria. Barcaui (2025) reportó una desventaja en la prueba diferida. Stadler et al. (2024) reportaron menos aspectos pertinentes en la justificación inmediata. Kestin et al. (2025) documentaron un postest favorable después de tutoría. Ninguno de esos tres permite calcular una brecha R–A a partir de dos resultados separados. Su inclusión evita reducir el análisis a una sola dirección de efectos y exige mantener el alcance de cada medida.

La diferencia de Barcaui (2025) equivale a 68,5 − 57,5 = 11 puntos porcentuales a favor del comparador. La pérdida de seguimiento fue (120 − 85) / 120 × 100 = 29,2 %. En el estudio de Kazemitabaar et al. (2023) fue (90 − 69) / 90 × 100 = 23,3 %. Estas comprobaciones describen cifras publicadas, no análisis nuevos de datos individuales. Los porcentajes relativos de Bastani et al. (2025) conservan un denominador distinto: por ejemplo, 0,137 / 0,284 × 100 ≈ 48 % para la mejora asistida de GPT Base. No se convierten en puntos porcentuales ni se suman entre estudios.

Evidencia sobre delegación y esfuerzo intelectual

El examen de los procesos permitió distinguir la información observada de las conclusiones sobre mecanismos. Bastani et al. (2025) examinaron solicitudes y respuestas. Kazemitabaar et al. (2023) registraron código generado e incorporación. Fan et al. (2025) analizaron actividad y secuencias de autorregulación. Chen (2026) integró interacciones, versiones y reflexión. Tales registros aportan conductas observables con niveles distintos de aproximación al procesamiento. No equivalen a una medición directa común de esfuerzo intelectual.

Stadler et al. (2024) midieron carga cognitiva auto informada. Barcaui (2025) recuperó tiempo de estudio declarado, y Kestin et al. (2025) describieron condiciones de tutoría y tiempo de actividad. La revisión identificó así una diferencia documental: disponer de indicadores de demanda o duración no implica conocer qué razonamiento se delegó. Tampoco la incorporación de texto demuestra ausencia de comprensión. Las fuentes permitieron describir procesos y asociaciones, pero no establecer un mecanismo causal compartido entre los siete estudios.

El contraste obtuvo un resultado de síntesis documental: los indicadores revisados describen cambios de demanda y formas de delegación, pero no establecen desaparición del esfuerzo intelectual. La ausencia de una operación central requiere distinguir lo realizado por la herramienta de su ejecución, reconstrucción o verificación por el estudiante. Una desventaja posterior no acredita por sí sola esa ausencia. Con las reglas propuestas, esta condición permanece sin establecerse en la síntesis.

La integración de los hallazgos permitió establecer cuatro exigencias de análisis. Se necesita identificar la operación efectivamente delegada, documentar la actividad que conserva el estudiante, medir por separado la ejecución conjunta y describir la evaluación posterior, su independencia y su intervalo. Estas exigencias fundamentan la formulación operativa de DERA en la discusión. Constituyen una síntesis conceptual del contraste, no una validación de una escala ni un hallazgo obtenido mediante observación nueva de participantes.

Reglas de análisis derivadas del contraste

La comparación documental precisó cuatro condiciones de evidencia para las dimensiones DERA del esquema inicial. Se identificó la necesidad de registrar la operación delegada, la actividad propia, el resultado durante la asistencia y la actuación posterior con sus recursos e intervalo. La tabla 3 presenta esa matriz de preguntas y límites. Constituye el resultado analítico del cuarto objetivo, sin asignar puntuaciones a participantes ni validar un instrumento.

Tabla 3

Dimensiones DERA y evidencia necesaria para su aplicación

Dimensión	Pregunta y evidencia	Límite de interpretación
D Delegación	Qué operación realizó la IA. Solicitudes, respuestas y cambios en el producto	Disponibilidad de ayuda y operación efectivamente delegada son datos diferentes
E Esfuerzo residual	Qué actividad conservó el estudiante. Intentos, versiones, verificaciones y explicaciones	Duración, mensajes y dificultad son indicadores parciales. Una medida ausente no significa esfuerzo nulo
R Rendimiento asistido	Qué produjo durante el acceso. Calidad, exactitud o ejecución con comparador definido	El resultado del sistema estudiante–IA no se atribuye íntegramente a la persona
A Aprendizaje independiente	Qué pudo hacer sin la ayuda examinada. Actuación inmediata, retención o transferencia delimitadas	La independencia y el intervalo deben documentarse. Una puntuación es evidencia acotada del constructo aprendizaje

Nota. Las dimensiones organizan preguntas y no se suman para producir una puntuación global. Una ausencia de registro no se clasifica como desaparición del esfuerzo. Las reglas propuestas requieren evaluación de validez y acuerdo entre evaluadores antes de utilizarlas para decisiones individuales.

Balance documental de las hipótesis

La tabla 4 responde expresamente a las hipótesis. Su balance conserva la diferencia entre respaldo a una distinción de resultados y demostración de un mecanismo. No utiliza el número de resultados significativos como prueba conjunta ni interpreta una diferencia no detectada como aceptación de una hipótesis nula.

Tabla 4

Respuesta a las hipótesis y alcance de la evidencia

Hipótesis	Resultado del contraste documental	Evidencia no establecida
Correspondencia entre rendimiento asistido y aprendizaje	Bastani et al. (2025) registran ventaja asistida y desventaja posterior bajo GPT Base. Otros resultados muestran correspondencias diferentes o imprecisas	No se obtuvo una estimación conjunta ni un efecto general para todos los tipos de asistencia
Posible desaparición de una operación central	Se recuperaron registros de delegación y práctica propia. La ausencia de realización, reconstrucción y verificación no quedó establecida con las medidas disponibles	No se establecieron la condición de ausencia de la operación ni su relación causal con el desempeño posterior
Conservación de práctica en el diseño de la ayuda	Indicios compatibles en intervenciones con salvaguardas y reflexión. Componentes no aislados	No prueba el efecto específico del intento previo, las preguntas o la reflexión. La incertidumbre exige estudiar cada condición por separado

Nota. El balance es analítico y corresponde a una selección intencional. Respaldo documental no equivale a confirmación estadística de una relación general. Las limitaciones son parte de la respuesta, no datos imputados para completar el contraste.

DISCUSIÓN

Respuesta a la pregunta de investigación

La pregunta planteada recibe una respuesta delimitada: los estudios aportan evidencia sobre delegación y cambios de actividad, pero no establecen desaparición del esfuerzo intelectual. El rendimiento asistido y el desempeño posterior tampoco presentan una correspondencia automática. El resultado conjunto puede mejorar mientras una actuación individual posterior sigue otro patrón. Esa divergencia justifica separar las medidas. Para explicar sus causas se necesita documentar la operación asumida por la herramienta y la práctica que dejó de realizar o conservó el estudiante.

La primera hipótesis encuentra su respaldo más claro en un contraste que observa resultados durante el uso y después de retirar la asistencia. Ese respaldo tiene una consecuencia

concreta para la evaluación: una calificación alta del producto asistido no debe convertirse por sí sola en una afirmación de competencia personal. La conclusión no exige que todas las intervenciones produzcan una desventaja. Basta con reconocer que calidad del producto y capacidad individual requieren evidencia diferente. Salomon et al. (1991) permiten distinguir los efectos de trabajar con tecnología de los que permanecen en la persona. Soderstrom y Bjork (2015) precisan la diferencia entre rendimiento durante la adquisición y aprendizaje observado en retención o transferencia. Ambas distinciones conservan utilidad al examinar la IA generativa.

Las medidas también delimitan lo que se puede atribuir a aprendizaje. Una prueba inmediata permite observar una actuación en condiciones específicas. Una evaluación diferida añade información sobre conservación. Una tarea diferente añade información sobre transferencia. Ninguna de esas medidas agota por sí sola una competencia. La duración de una intervención describe cuánto tiempo se ofreció asistencia, mientras el intervalo de seguimiento describe cuánto tiempo transcurrió antes de comprobar lo aprendido. Confundir ambos períodos puede ampliar indebidamente la conclusión.

La falta de correspondencia puede depender asimismo de lo evaluado. Revisar un ensayo y contestar preguntas de conocimiento requieren operaciones relacionadas, pero diferentes. Una mejora en la primera medida no asegura una mejora de la segunda. Tampoco una diferencia no detectada en la segunda elimina lo aprendido durante la revisión. La interpretación debe conservar la finalidad y sensibilidad de los instrumentos. El problema de atribución exige comprobar qué capacidad se observa, antes de relacionar dos puntuaciones por compartir un estudiante o una intervención.

Cuando la respuesta precede a la comprensión

Disponer de una respuesta antes de poder explicarla plantea una dificultad de atribución. La herramienta puede entregar el resultado mientras el estudiante todavía trabaja sobre las relaciones que lo sostienen. En ese episodio, respuesta disponible y comprensión alcanzada son condiciones distintas. El contraste entre producto asistido y actuación posterior permite examinar sus consecuencias, aunque no registra por sí solo el momento preciso en que se produjo la comprensión.

La desaparición de una operación central constituye una hipótesis más exigente que la reducción de dificultad. Una persona puede emplear menos tiempo porque elimina una búsqueda auxiliar y concentrarse en evaluar la evidencia. También puede adoptar una inferencia sin reconstruirla. Los dos episodios pueden terminar con un producto de calidad semejante. Para distinguirlos se requiere conocer qué operación estaba prevista y qué actividad personal se mantuvo. Estos ejemplos delimitan la interpretación y no agregan observaciones al corpus.

La actividad puede desplazarse después de recibir la respuesta. Detectar un error, comparar una alternativa o justificar una decisión conserva procesamiento propio, aunque la primera solución haya sido generada externamente. Ese desplazamiento impide equiparar

respuesta inmediata y ausencia de esfuerzo. En cambio, si una operación necesaria para aprender deja de realizarse y no se reconstruye ni verifica, corresponde examinar qué oportunidad de práctica se retiró y qué puede hacer después el estudiante.

La evidencia del corpus sostiene una precaución concreta: la respuesta disponible no acredita por sí solo aprendizaje. Los indicadores de duración, carga o incorporación tampoco permiten afirmar desaparición de actividad intelectual. La explicación necesita vincular el episodio asistido con una actuación propia y reconocer el procesamiento que el registro no logra observar. Esa exigencia conserva el problema planteado y delimita lo que la revisión puede responder.

La explicación por sustitución cognitiva

La segunda hipótesis plantea una condición que la revisión no establece de manera general. La sustitución puede ser parcial, mientras la desaparición de una operación central exige comprobar que no hubo realización, reconstrucción ni verificación propia de esa operación. Un historial puede mostrar que se solicitó una solución completa. Para examinar lo ocurrido después debe comprobarse si la persona comparó alternativas, revisó premisas, rehizo el procedimiento o incorporó el resultado sin esa práctica. Solicitud y uso efectivo son momentos distintos.

La centralidad evita una clasificación circular. Si se pretende aprender a construir inferencias, esa operación debe reconocerse como parte del objetivo antes de observar el postest. Un menor resultado posterior no basta para convertir cualquier operación externalizada en central. Del mismo modo, una buena puntuación no transforma automáticamente la delegación en auxiliar. Definir la competencia con anticipación permite que el modelo admita resultados que contradigan su explicación inicial.

La divergencia R–A conserva explicaciones alternativas. El asistente puede introducir errores. La actividad puede ofrecer menos oportunidades de práctica. La prueba puede evaluar habilidades distintas. Y el tiempo disponible puede cambiar la estrategia. Una condición experimental combina además contenido, interfaz e instrucciones. La asignación aleatoria permite examinar el efecto de esa condición completa dentro del estudio. No separa automáticamente cuánto corresponde a delegación, exactitud de respuestas, retroalimentación o distribución del tiempo.

Esta distinción también afecta a los análisis de mediación. Que una medida de incorporación se relacione con una puntuación posterior no demuestra que modificarla produzca el mismo cambio en aprendizaje. Conocimiento previo, dificultades de la tarea o decisiones del estudiante pueden influir en ambos. En Chen (2026), la comparación de procesos resulta informativa, pero conserva la reflexión y la práctica propia como explicaciones relacionadas. Su aportación consiste en hacer observables partes del proceso y delimitar rivales, sin resolver por sí sola toda la atribución causal.

Esta hipótesis perdería respaldo si una medición adecuada mostrara que la ausencia documentada de una operación central, sin reconstrucción ni verificación propia, se acompaña de resultados independientes iguales o superiores con precisión suficiente. Un caso así exigiría examinar vías de aprendizaje omitidas, como comprender ejemplos o evaluar respuestas. También podría indicar que la clasificación de centralidad fue inadecuada. La propuesta debe admitir esa posibilidad. De otro modo, cualquier resultado terminaría convertido en confirmación del mismo mecanismo.

Las condiciones de la asistencia

La tercera hipótesis recibe indicios compatibles, aunque no una prueba aislada de sus componentes. El contraste entre acceso abierto y asistencia con salvaguardas plantea que importa la forma en que la herramienta interviene. Sin embargo, las salvaguardas pueden cambiar simultáneamente exactitud, contenido y tipo de respuesta. La reflexión obligatoria puede modificar tanto el procesamiento como el tiempo de práctica. Por eso, el patrón favorable de una condición completa no identifica cuál de sus elementos produjo la diferencia.

La programación ofrece una razón adicional para evitar una lectura uniforme. Un estudiante puede recibir código y después probarlo, modificarlo y justificar su funcionamiento. La generación externa y la práctica propia pueden coexistir en una secuencia. El resultado global sin diferencias detectadas no demuestra que todos los procesos sean iguales, pero impide deducir un perjuicio uniforme únicamente porque existió generación de respuestas. Debe examinarse qué operaciones permanecieron y qué indicador posterior las representa.

El postest favorable de Kestin et al. (2025) cumple una función crítica en la comparación: muestra un resultado educativo positivo bajo un diseño tutorial específico. Esa intervención reúne ritmo personal y apoyos preparados. El material recuperado conserva límites sobre la independencia de la prueba. Su resultado cuestiona una conclusión general de perjuicio, sin resolver permanencia o transferencia amplia. Incorporar este contraste exige el mismo rigor que se aplica a los hallazgos desfavorables.

La consecuencia de este contraste es investigar condiciones de ayuda y conservar evidencia de su funcionamiento. Panadero y Alonso-Tapia (2014) permiten situar esas ayudas en relación con las metas y el control que ejerce el estudiante. Puede compararse respuesta completa, pista gradual, retroalimentación a un intento y explicación de ejemplos, manteniendo constantes otras características. Si se modifica todo a la vez, la conclusión debe referirse al paquete pedagógico. Presentar cada componente como una intervención comprobada excedería el alcance de la revisión.

El planteamiento de Fuertes Alpiste (2024) ofrece un criterio para interpretar este contraste: las herramientas pueden apoyar la cognición si el estudiante conserva las funciones ejecutivas. DERA exige precisar cómo se observa esa conservación. Esta relación es conceptual.

No convierte el marco teórico en evidencia experimental del efecto de una modalidad particular de asistencia.

Función formativa del esfuerzo y conocimiento previo

La cantidad de esfuerzo no ofrece una medida suficiente de aprendizaje. Una demanda alta puede reflejar dificultad innecesaria y una demanda menor puede permitir atender a relaciones relevantes. Sweller (1988) permite distinguir la demanda de resolver un problema de la adquisición de esquemas. Andrade-Lotero (2012) relaciona el diseño de la ayuda con el conocimiento previo y el procesamiento que exige la tarea. La pregunta útil es qué operación se facilitó y qué procesamiento quedó disponible. Reducir la búsqueda auxiliar puede liberar atención. Roediger y Karpicke (2006) muestran por qué la recuperación merece conservarse cuando la retención pertenece al objetivo de aprendizaje.

El conocimiento previo modifica esta relación. Andrade-Lotero (2012) analiza por qué una ayuda puede tener un valor diferente según la experiencia del aprendiz. Quien todavía no dispone de esquemas puede necesitar ejemplos y explicaciones que para una persona experta resultarían redundantes. Pedir un intento antes de recibir ayuda tiene sentido cuando ese intento permite movilizar recursos suficientes. Exigir una exploración prolongada sin apoyo puede elevar dificultad sin construir comprensión. Conservar práctica formativa significa ajustar su función y sus condiciones, no prolongar indiscriminadamente el trabajo.

La facilidad percibida requiere interpretación. La satisfacción puede acompañar comprensión o dependencia de una respuesta externa. Osses Bustingorry y Jaramillo Mora (2008) sitúan la planificación, la supervisión y la evaluación de la actividad entre las funciones metacognitivas. Esa delimitación orienta el examen del proceso, pero no demuestra un efecto específico de la IA. Cuando interés y seguridad aumentan, resulta pertinente comprobar qué puede explicar o aplicar el estudiante. Una experiencia agradable y una actuación autónoma constituyen resultados distintos que pueden valorarse de forma complementaria.

La paciencia cognitiva ante una respuesta inmediata

La delimitación de Rosero Aulestia et al. (2026) permite distinguir cuánto trabajo exige una actividad del tiempo durante el cual debe mantenerse antes de comprender. La tarea puede ser exigente y ofrecer señales frecuentes de avance. También puede tener dificultad moderada y mantener incierto su resultado. La rapidez de la respuesta de IA modifica una condición de la actividad, pero no informa por sí sola sobre la comprensión que alcanzó el estudiante. Disponer de una explicación y comprenderla siguen siendo actuaciones distintas.

La relación con DERA debe mantenerse en ese nivel. El esfuerzo residual describe operaciones que el estudiante conserva. La paciencia cognitiva se refiere a un episodio en que ese esfuerzo se orienta a comprender mientras el resultado permanece diferido e incierto. No todo esfuerzo residual reúne esas condiciones. Verificar una solución inmediatamente disponible

puede exigir procesamiento sin constituir el mismo episodio de espera. La distinción evita usar ambos conceptos como nombres diferentes para una sola conducta.

Una pista puede orientar el avance sin entregar el razonamiento completo. Una explicación resuelta puede permitir estudiar relaciones que el estudiante todavía no reconoce. La decisión educativa depende de la operación y del conocimiento previo. La propuesta sobre paciencia cognitiva añade una pregunta sobre el momento y la información que acompaña la ayuda. No convierte la demora en un valor por sí mismo. Mantener al estudiante sin señales de progreso puede hacer incierto el proceso sin facilitar comprensión.

Los siete estudios empíricos examinados no permiten establecer un efecto de la IA sobre este constructo. Su incorporación es conceptual. Para comprobar esa relación sería necesario distinguir disponibilidad de respuesta, comprensión alcanzada y actividad durante el intervalo, manteniendo controladas dificultad y competencia previa. Esa exigencia amplía las preguntas de una investigación posterior, sin atribuir a la presente revisión una medición que no realizó.

DERA como resultado analítico y propuesta operativa

La matriz presentada en la tabla 3 integra las condiciones de evidencia mediante DERA. Su utilidad consiste en reconocer qué parte de una explicación está observada y cuál sigue inferida. Las cuatro dimensiones permiten organizar el registro y la lectura de un episodio asistido. Su aplicación exige comprobar la competencia objetivo y la procedencia de cada indicador. El resultado analítico no constituye una escala validada, un diagnóstico de dependencia ni una estimación del daño de usar IA.

Para D se propone registrar cada episodio de uso asociado a una decisión de la tarea. El registro identifica competencia objetivo, operación solicitada, respuesta y acción posterior. Puede clasificar expresión superficial, recuperación, organización, generación, razonamiento o evaluación. Una respuesta admite varias categorías. Un resumen puede seleccionar contenido y reorganizar relaciones a la vez. La centralidad se justifica como central, auxiliar o no determinada según el objetivo formativo, sin convertir las categorías en grados universales de daño.

E describe adopción, selección, transformación, verificación o reconstrucción. Para examinar desaparición se necesita identificar la operación central y comprobar su ausencia, en lugar de inferirla desde una baja frecuencia de uso o de una puntuación. Esas conductas no forman necesariamente una escala ordinal. Elegir entre argumentos puede requerir más evaluación que modificar extensamente el estilo. Una explicación larga puede ser incorrecta. Para cada observación interesa conservar tarea, exactitud, criterio y procedencia: registro de versiones, conducta observada, autoinforme o interpretación. La coincidencia textual puede señalar incorporación, mientras la comprensión necesita una evidencia distinta. La procedencia permite valorar la fuerza de cada inferencia.

R debe corresponder a la actividad y a un comparador definido. A requiere describir qué ayuda se retiró y qué recursos permanecieron. Un desempeño profesional puede incluir consulta

de normas o calculadora. Retirar esos recursos quizá cambie la competencia objetivo. La evaluación independiente se refiere a la ausencia del apoyo generativo cuya contribución se intenta distinguir. Si esa condición no consta, DERA conserva el postest reportado y registra el vacío. No completa la información mediante una suposición.

La relación entre R y A se expresa como perfil, conforme a la tabla 5. No se define una brecha numérica universal. Estandarizar puntuaciones no vuelve equivalentes tareas, constructos o comparadores. Una diferencia cuantitativa entre efectos requeriría calibración para una competencia comparable, mismo contraste de condiciones y tratamiento de la dependencia. Su incertidumbre exige la covarianza cuando las medidas proceden de las mismas personas o grupos. El corpus no reúne una base común para ese cálculo.

Tabla 5

Perfiles de rendimiento asistido y desempeño posterior

Resultado durante la asistencia	Resultado posterior	Interpretación admisible
Ventaja	Ventaja	Compatible con apoyo productivo. Precisar tarea, intervalo e incertidumbre
Ventaja	Diferencia no detectada	Ventaja asistida sin evidencia suficiente de ganancia posterior adicional. No demuestra equivalencia
Ventaja	Desventaja	Divergencia que requiere examinar proceso y explicaciones alternativas. No identifica una causa por sí sola
Desventaja o menor fluidez	Ventaja	Compatible con práctica formativa cuando medidas y diseño sostienen esa lectura
Sin medida separada	Cualquier resultado	Se informa el resultado posterior. El estudio no establece una brecha R–A

Nota. Son posibilidades analíticas, no un recuento de estudios ni categorías diagnósticas. La retirada del apoyo debe documentarse antes de interpretar el resultado posterior como actuación independiente.

Una aplicación conceptual en argumentación jurídica permite apreciar la diferencia. Automatizar bibliografía puede aliviar una operación auxiliar. Pedir a la herramienta que seleccione el hecho relevante y justifique la inferencia puede delegar una parte central del aprendizaje. Si el estudiante contrasta la propuesta, detecta un error y reconstruye la conclusión, conserva actividades que el producto final quizá no revela. Este ejemplo explica el registro. No constituye evidencia empírica sobre estudiantes de Derecho ni generaliza resultados educativos a esa población.

Consecuencias para la enseñanza y la evaluación

La consecuencia principal es distinguir la finalidad de cada evaluación. Marcelo (2013) sitúa la integración tecnológica en el cambio de las prácticas y en su contexto. Esa perspectiva permite examinar las decisiones docentes que acompañan el acceso a una herramienta. Una tarea con IA puede medir calidad de colaboración, verificación y uso competente de herramientas. Una

actuación posterior puede examinar la operación que la persona conserva. Ambas capacidades tienen valor cuando pertenecen al objetivo del curso. Su identificación en la calificación permite saber qué se observó y evita atribuir una misma puntuación a capacidades distintas sin comprobación.

Un procedimiento aplicable comienza definiendo la competencia y una tarea breve de base. Durante el trabajo asistido se conserva un intento o explicación inicial, las ayudas y las decisiones de aceptar, modificar o rechazar. Después se propone una tarea comparable sin el apoyo generativo examinado. Si interesa permanencia, se añade seguimiento. Si interesa transferencia, se modifica una característica relevante y se declara cuál. Este procedimiento es una propuesta derivada del análisis y necesita evaluación propia antes de presentarse como intervención eficaz.

García Peñalvo et al. (2024) proponen ampliar la evaluación hacia el proceso seguido y complementar los productos escritos con otras actuaciones. Esa orientación coincide con la necesidad de distinguir ejecución conjunta y competencia personal. Su propuesta no sustituye la comprobación de independencia, permanencia y transferencia que exige el presente contraste.

Las decisiones deben atender a los resultados y sus instrumentos. Si mejora el producto asistido y aparece dificultad en la actuación personal, corresponde revisar qué operación se automatizó y qué práctica se conservó. Si ambas actuaciones mejoran, el apoyo merece examinarse para conocer sus condiciones. Si la diferencia posterior es imprecisa, no procede concluir igualdad ni asignar una causa. La respuesta educativa depende de la competencia evaluada y de la calidad de la evidencia, no de un solo signo estadístico.

La retirada de ayuda tampoco debería convertirse en un criterio uniforme ajeno al objetivo. Si la competencia incluye trabajar con herramientas, interesa comprobar su uso y la detección de errores. Si incluye argumentar por cuenta propia, debe existir una actuación que lo permita observar. En ambos casos, las condiciones deben ser transparentes y comparables. La evaluación debe distinguir apoyos generativos de recursos que forman parte habitual de la tarea, considerando las necesidades de acceso del alumnado sin confundir adaptación y sustitución de la operación evaluada.

La perspectiva de Díaz Barriga (2003) añade que la actividad y su contexto forman parte del aprendizaje que se examina. Retirar indiscriminadamente todos los recursos puede transformar la tarea. Por ello, la condición independiente debe definirse según la operación formativa y conservar los apoyos compatibles con ella, como se propone en DERA.

La aplicación de estas evaluaciones requiere describir también las condiciones de acceso. El análisis de Rosero Aulestia et al. (2026) muestra por qué estar conectado no equivale a disponer de las mismas posibilidades de exploración digital. En el presente problema, corresponde registrar qué herramienta estuvo disponible, durante cuánto tiempo y con qué restricciones. La consecuencia es metodológica: si las condiciones materiales difieren, deben conservarse en la

interpretación. El antecedente tarifario no demuestra un efecto de esas diferencias sobre el aprendizaje con IA.

Contrastabilidad y límites de la investigación

Validar DERA exige indicadores aplicables con criterios compartidos y evidencia de que D y E representan dimensiones distintas. Un estudio puede registrar competencia previa, asignar modalidades de asistencia, conservar interacciones y versiones y evaluar actuaciones inmediatas y diferidas. Para aislar componentes pedagógicos debe mantener constantes otros elementos relevantes. La asignación por aulas y las evaluaciones repetidas requieren conservar su dependencia en el análisis. No corresponde compensar pocas clases tratando cada puntuación como una asignación independiente.

La capacidad explicativa del modelo debe compararse con alternativas que incluyan conocimiento previo, errores, tiempo y práctica. Si D y E no añaden información con medición adecuada y precisión suficiente, habrá que revisar la propuesta. Un resultado impreciso no resolvería esa prueba. Si se observa una ganancia independiente bajo delegación central, interesa identificar qué actividad permitió aprender y qué supuesto necesita modificación. Esos resultados posibles hacen que DERA sea contrastable, en lugar de convertirlo en una explicación compatible con cualquier desenlace.

Moya Cano y Moya Espinosa (2026) presentan AIthos EDU como un modelo proyectivo para alfabetización en IA, competencia digital crítica, pensamiento crítico y reflexión ética, cuya eficacia no evalúan empíricamente. DERA comparte la atención a la actividad personal, pero delimita condiciones para atribuir el producto asistido y el desempeño posterior. La diferencia entre proponer un modelo y validar su funcionamiento se mantiene en ambos casos.

El alcance de esta investigación permanece acotado. La selección de siete estudios partió de un esquema y se amplió de manera intencional. Puede omitir trabajos relevantes y no permite estimar frecuencias de la literatura. La extracción no fue duplicada de forma independiente. El trabajo de Barcaui (2025) se consultó de forma parcial y ciertos controles de evaluación no quedaron suficientemente documentados. Esas condiciones limitan la síntesis y permanecen visibles en los resultados.

Las poblaciones, disciplinas y herramientas tampoco ofrecen una base común para una conclusión causal general. Retención prolongada y transferencia distante están cubiertas de forma desigual. La revisión aporta una comparación documental, respuestas diferenciadas a las hipótesis y reglas de interpretación. No valida DERA ni demuestra desaparición del esfuerzo intelectual. Las limitaciones precisan qué nueva evidencia hace falta para avanzar desde una explicación plausible hacia una relación comprobada.

CONCLUSIONES

La pregunta de investigación se responde mediante una distinción: los estudios seleccionados permiten examinar descarga cognitiva, delegación y cambios en la actividad, pero no establecer desaparición del esfuerzo intelectual. La mejora del producto conjunto tampoco constituye por sí sola evidencia de una competencia personal equivalente. Disponer de una respuesta y haber aprendido a construirla requieren comprobaciones diferentes. Para atribuir aprendizaje se necesita identificar la actuación posterior, los recursos permitidos y el alcance de la tarea y del seguimiento.

La descripción del corpus permitió precisar los diseños y las condiciones de medida. Cuatro estudios evaluaron por separado un producto durante la asistencia y tres de ellos documentaron una prueba posterior sin IA. El contraste mostró que esas medidas pueden seguir patrones diferentes y que solo dos estudios del corpus incluyeron retención diferida. Los resultados de una intervención de ocho semanas, una prueba inmediata y una evaluación a los 45 días responden a preguntas temporales distintas. Sus puntuaciones no permiten construir un efecto agregado común.

La hipótesis de correspondencia no automática encuentra respaldo documental acotado: existe evidencia de ventaja asistida acompañada de desventaja posterior. La segunda hipótesis permanece sin comprobar. Los registros examinados permiten estudiar sustitución de actividad, pero no establecen ausencia de realización, reconstrucción y verificación de una operación central ni su relación causal con el desempeño. Se conservan errores, conocimiento previo y condiciones pedagógicas como explicaciones alternativas. La hipótesis sobre conservación de práctica recibe indicios compatibles en intervenciones con salvaguardas o reflexión, sin prueba separada de sus componentes. Ninguna de estas respuestas sostiene un perjuicio o beneficio universal de la IA.

La formulación de reglas se concreta en DERA como propuesta para registrar operación delegada, actividad conservada, ejecución asistida y actuación independiente. Su contribución es delimitar evidencia e inferencias, no asignar una puntuación diagnóstica. Requiere validación, instrumentos pertinentes y evaluación de acuerdo entre observadores antes de utilizarse para decisiones individuales.

La consecuencia educativa es evaluar conjuntamente producto y proceso cuando el objetivo incluye capacidad personal. Conservar evidencia de las ayudas y comprobar una actuación posterior permite distinguir qué logró el sistema estudiante–IA y qué puede realizar el estudiante. La decisión sobre un apoyo debe atender a la operación que facilita, a la práctica que permanece y a los resultados que efectivamente pueden observarse. La paciencia cognitiva añade una pregunta sobre el intervalo previo a la comprensión, mientras las condiciones de acceso delimitan las oportunidades de realizar la actividad. Ambos antecedentes orientan el análisis y requieren mediciones específicas para relacionarlos con el desempeño.

Disponibilidad de datos y materiales

El trabajo analiza información publicada y no incorpora mediciones originales de estudiantes. Los estudios se identifican en la tabla 1 y en las referencias. Las reglas de extracción se describen en el método. Su aplicación se presenta en las tablas 2, 3 y 4. Las condiciones DERA se exponen en los resultados y su interpretación se desarrolla en la discusión. El apéndice identifica la procedencia y las versiones de las fuentes. Los recuentos y las comprobaciones aritméticas son reproducibles a partir de esa información.

REFERENCIAS

Andrade-Lotero, L. A. (2012). Teoría de la carga cognitiva, diseño multimedia y aprendizaje: Un estado del arte. magis, Revista Internacional de Investigación en Educación, 5(10), 75–92. https://doi.org/10.11144/Javeriana.m5-10.tccd

Barcaui, A. (2025). ChatGPT as a cognitive crutch: Evidence from a randomized controlled trial on knowledge retention. Social Sciences & Humanities Open, 12, Artículo 102287. https://doi.org/10.1016/j.ssaho.2025.102287

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

Chen, X. (2026). Layer-sensitive cognitive offloading in generative AI-assisted writing: Supported performance and independent no-AI outcomes. Frontiers in Psychology, 17, Artículo 1906199. https://doi.org/10.3389/fpsyg.2026.1906199

Deng, R., Jiang, M., Yu, X., Lu, Y., & Liu, S. (2025). Does ChatGPT enhance student learning? A systematic review and meta-analysis of experimental studies. Computers & Education, 227, Artículo 105224. https://doi.org/10.1016/j.compedu.2024.105224

Díaz Barriga, F. (2003). Cognición situada y estrategias para el aprendizaje significativo. Revista Electrónica de Investigación Educativa, 5(2), 1–13. https://redie.uabc.mx/redie/article/view/85

Fan, Y., Tang, L., Le, H., Shen, K., Tan, S., Zhao, Y., Shen, Y., Li, X., & Gašević, D. (2025). Beware of metacognitive laziness: Effects of generative artificial intelligence on learning motivation, processes, and performance. British Journal of Educational Technology, 56(2), 489–530. https://doi.org/10.1111/bjet.13544

Fuertes Alpiste, M. (2024). Enmarcando las aplicaciones de IA generativa como herramientas para la cognición en educación. Pixel-Bit. Revista de Medios y Educación, (71), 42–57. https://doi.org/10.12795/pixelbit.107697

García Peñalvo, F. J., Llorens-Largo, F., & Vidal, J. (2024). La nueva realidad de la educación ante los avances de la inteligencia artificial generativa. RIED-Revista Iberoamericana de Educación a Distancia, 27(1), 9–39. https://doi.org/10.5944/ried.27.1.37716

Guirao Goris, S. J. A. (2015). Utilidad y tipos de revisión de literatura. Ene, 9(2). https://doi.org/10.4321/S1988-348X2015000200002

Kazemitabaar, M., Chow, J., Ma, C. K. T., Ericson, B. J., Weintrop, D., & Grossman, T. (2023). Studying the effect of AI code generators on supporting novice learners in introductory programming. En A. Schmidt, K. Väänänen, T. Goyal, P. O. Kristensson, A. Peters, S. Mueller, J. R. Williamson, & M. L. Wilson (Eds.), Proceedings of the 2023 CHI

Conference on Human Factors in Computing Systems (pp. 1–23). Association for Computing Machinery. https://doi.org/10.1145/3544548.3580919

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

López Aymes, G. (2012). Pensamiento crítico en el aula. Docencia e Investigación, (22), 41–60. https://ruidera.uclm.es/server/api/core/bitstreams/da191441-9a5a-4c98-a2d0-1491b06cda81/content

Marcelo, C. (2013). Las tecnologías para la innovación y la práctica docente. Revista Brasileira de Educação, 18(52), 25–47. https://doi.org/10.1590/S1413-24782013000100003

Mo, F., Huang, J., Yang, Y., Özen, Z., Maeda, Y., & Olenchak, F. R. (2026). Undergraduate students' learning outcomes with ChatGPT: A meta-analytic study. Computers and Education: Artificial Intelligence, 10, Artículo 100536. https://doi.org/10.1016/j.caeai.2025.100536

Moya Cano, A., & Moya Espinosa, M. (2026). Uso crítico de la inteligencia artificial generativa en educación superior: Diseño de un modelo pedagógico para el desarrollo de la competencia digital y el pensamiento crítico. Interdisciplinary Journal of Didactics, (5), 111–134. https://doi.org/10.14198/ijd.33415

Osses Bustingorry, S., & Jaramillo Mora, S. (2008). Metacognición: Un camino para aprender a aprender. Estudios Pedagógicos, 34(1), 187–197. https://doi.org/10.4067/S0718-07052008000100011

Panadero, E., & Alonso-Tapia, J. (2014). Teorías de autorregulación educativa: Una comparación y reflexión teórica. Psicología Educativa, 20(1), 11–22. https://doi.org/10.1016/j.pse.2014.05.002

Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. Trends in Cognitive Sciences, 20(9), 676–688. https://doi.org/10.1016/j.tics.2016.07.002

Roediger, H. L., III, & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. Psychological Science, 17(3), 249–255. https://doi.org/10.1111/j.1467-9280.2006.01693.x

Rosero Aulestia, F. X., Oviedo Fraga, D., & López Jara, J. (2026). Tarifas de datos y aprendizaje informal: El zero-rating como currículo invisible en el Ecuador digital. Arandu UTIC, 13(3), 1372–1393. https://doi.org/10.69639/arandu.v13i3.2478

Rosero Aulestia, F. X., Zambrano Trujillo, S. L., Díaz Paredes, C. G., Oviedo Fraga, D., & Meza Romero, A. E. (2026). La paciencia cognitiva: Definición y criterios para la docencia universitaria. LATAM Revista Latinoamericana de Ciencias Sociales y Humanidades, 7(4), 3663–3684. https://doi.org/10.56712/latam.v7i4.6488

Salomon, G., Perkins, D. N., & Globerson, T. (1991). Partners in cognition: Extending human intelligence with intelligent technologies. Educational Researcher, 20(3), 2–9. https://doi.org/10.3102/0013189X020003002

Soderstrom, N. C., & Bjork, R. A. (2015). Learning versus performance: An integrative review. Perspectives on Psychological Science, 10(2), 176–199. https://doi.org/10.1177/1745691615569000

Stadler, M., Bannert, M., & Sailer, M. (2024). Cognitive ease at a cost: LLMs reduce mental effort but compromise depth in student scientific inquiry. Computers in Human Behavior, 160, Artículo 108386. https://doi.org/10.1016/j.chb.2024.108386

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. Cognitive Science, 12(2), 257–285. https://doi.org/10.1207/s15516709cog1202_4