UNIVERSIDAD DE PALERMO

Facultad de Ingeniería
Maestría en Tecnología de la Información

Deuda cognitiva y uso de Inteligencia Artificial en la educación superior de grado en Argentina

Tesis de Maestría

Autor: Nahuel Luis Gentile

Tutor: Diego Branca

Director: Claudio Zamoszczyk

Julio, 2026

Índice

Resumen....4
Abstract....5
1. Introducción....6
2. Antecedentes y Planteamiento del Problema....9
2.1 Historia de la adopción tecnológica en educación superior....11
2.2 La IA generativa como ruptura cualitativa....13
2.3 Contexto argentino y planteamiento del problema central....15
3. Objetivos....17
3.1 Objetivo General....17
3.2 Objetivos Específicos....17
4. Justificación....18
5. Marco Teórico....19
5.1 La expansión de la IA y la dependencia emergente....20
5.2 IA generativa y delegación cognitiva....24
5.3 Deuda cognitiva: definición y fundamentos....30
5.4 Riesgos cognitivos en la educación superior....34
5.5 Competencias cognitivas en la era digital....39
6. Metodología....43
6.1 Diseño del cuestionario....44
6.2 Dimensiones propuestas....45
6.3 Construcción del índice de deuda cognitiva....50
6.4 Validación preliminar....51
6.5 Modelado exploratorio y automatización....52
7. Resultados Esperados....54
8. Análisis de Resultados....57
8.1 Estadísticas descriptivas del índice de deuda cognitiva....59
8.2 Frecuencia de uso y su relación con la deuda cognitiva....61
8.3 Análisis por carrera....63
8.4 Análisis por universidad....65
8.5 Perfil de respuestas por ítem y por dimensión....66
8.6 Validación del instrumento....67
8.7 Verificación de hipótesis....73
8.8 Modelado exploratorio: regresión....76
9. Conclusiones....80
9.1 Síntesis de hallazgos....80
9.2 Alcances y limitaciones....82
9.3 Propuesta de mitigación....83
9.4 Contribución del trabajo....86

10. Próximos pasos para una investigación futura..... 87
10.1 Ampliación y mejora del instrumento..... 87
10.2 Diseño longitudinal..... 88
10.3 Extensión a nuevas poblaciones y contextos..... 88
10.4 Modelado avanzado..... 89
10.5 Diferenciación entre modelos y proveedores de IA..... 90
11. Percepción personal..... 91
Bibliografía..... 92
Anexo A: Cuestionario aplicado..... 96

Resumen

La rápida expansión de la Inteligencia Artificial generativa transformó a fondo la forma en que los estudiantes aprenden, estudian y producen conocimiento en la educación superior. Ese mismo progreso, sin embargo, trajo un fenómeno emergente: la deuda cognitiva, entendida como la pérdida progresiva de esfuerzo mental, autonomía cognitiva y pensamiento crítico derivada de la delegación excesiva del razonamiento humano hacia sistemas de IA.

Una revisión de la literatura reciente demuestra que el uso intensivo de IA puede generar dependencia, disminución de la activación cognitiva y dificultades para retomar tareas sin asistencia tecnológica (Kosmyna et al., 2025; Fan, Y., et al., 2025). La evidencia indica que los propios modelos de lenguaje presentan limitaciones significativas en las condiciones reales de uso estudiantil: Laban et al. (2025) documentan una caída del 39% en el rendimiento de los principales LLMs en conversaciones de múltiples turnos, mientras que Leviathan et al. (2025) demuestran que variables arquitectónicas ajenas al usuario afectan sustancialmente la calidad de las respuestas. En otras palabras, hay factores de los LLMs que determinan la calidad de la respuesta que están completamente fuera del control y del conocimiento del usuario promedio. Estas limitaciones amplían la brecha entre la eficiencia percibida y el aprendizaje real. Si bien estas tecnologías ofrecen beneficios evidentes en términos de productividad y acceso al conocimiento, sus efectos sobre los procesos cognitivos profundos aún no se encuentran plenamente comprendidos ni sistemáticamente medidos en contextos educativos reales, particularmente en el ámbito de la educación superior latinoamericana.

El presente trabajo analiza los riesgos e impactos de la deuda cognitiva en la educación superior de grado en la Argentina y propone un aporte original: el diseño de un cuestionario estructurado de catorce preguntas (trece de ellas, ítems de medición) para evaluar niveles de deuda cognitiva en estudiantes universitarios de universidades públicas y privadas de Rosario y Buenos Aires, considerando seis dimensiones: frecuencia de uso, esfuerzo mental, dependencia tecnológica, autonomía cognitiva, conducta de revisión y fatiga cognitiva. El instrumento

permite clasificar el nivel de deuda cognitiva en cinco categorías (muy baja a muy alta) mediante un índice estandarizado de 0 a 100, y explorar de manera complementaria su estimación automatizada mediante un modelo analítico simple, como la regresión lineal múltiple. De este modo, el trabajo busca aportar una herramienta práctica para el análisis del impacto cognitivo del uso de IA en el ámbito académico y para el diseño de estrategias educativas orientadas al desarrollo del pensamiento crítico y la autonomía intelectual.

Abstract

The rapid expansion of generative Artificial Intelligence has profoundly transformed the way students learn, study, and produce knowledge in higher education. This progress, however, has introduced an emerging phenomenon: cognitive debt, understood as the progressive loss of mental effort, cognitive autonomy, and critical thinking resulting from the excessive delegation of human reasoning to AI systems.

A review of recent literature shows that intensive use of AI can generate dependency, reduced cognitive activation, and difficulty resuming tasks without technological assistance (Kosmyna et al., 2025; Fan, Y., et al., 2025). The evidence further indicates that language models themselves present significant limitations under real conditions of student use: Laban et al. (2025) document a 39% drop in the performance of leading LLMs in multi-turn conversations, while Leviathan et al. (2025) demonstrate that architectural variables beyond the user's control substantially affect response quality. These limitations widen the gap between perceived efficiency and actual learning. While these technologies offer clear benefits in terms of productivity and access to knowledge, their effects on deep cognitive processes remain neither fully understood nor systematically measured in real educational settings, particularly within Latin American higher education.

This work analyzes the risks and impacts of cognitive debt in undergraduate higher education in Argentina and proposes an original contribution: the design of a

structured fourteen-question questionnaire (thirteen of them scored items) to assess levels of cognitive debt in university students from public and private universities in Rosario and Buenos Aires, considering six dimensions: frequency of use, mental effort, technological dependence, cognitive autonomy, reviewing behavior, and cognitive fatigue. The instrument classifies the level of cognitive debt into five categories (very low to very high) through a standardized index from 0 to 100, and additionally explores its automated estimation through a simple analytical model such as multiple linear regression. In this way, the work aims to provide a practical tool for analyzing the cognitive impact of AI use in academic contexts and for designing educational strategies oriented toward the development of critical thinking and intellectual autonomy.

Keywords: cognitive debt; generative artificial intelligence; higher education; critical thinking; cognitive autonomy; educational assessment.

1. Introducción

La Inteligencia Artificial (IA) se ha consolidado como una de las tecnologías más disruptivas del siglo XXI, modificando de manera profunda la forma en que las personas acceden a la información, producen conocimiento y toman decisiones. Dentro de este escenario, la aparición y rápida expansión de los modelos de Inteligencia Artificial generativa, particularmente los *Large Language Models* (LLMs) como ChatGPT, Claude, Deepseek, Gemini o Copilot (entre otras centenas de entidades y modelos) ha generado un impacto significativo en el ámbito educativo, especialmente en la educación superior.

La incorporación de estas herramientas en contextos universitarios ha introducido nuevas dinámicas de aprendizaje y producción académica. Los estudiantes pueden actualmente redactar ensayos, resolver ejercicios, resumir textos, estructurar investigaciones y generar contenido complejo en cuestión de pocos minutos e incluso segundos. Esta capacidad representa una transformación sin precedentes en términos de acceso al conocimiento, productividad y automatización cognitiva.

Diversos autores sostienen que la integración de tecnologías digitales en la educación superior ha sido históricamente presentada como un mecanismo para democratizar el conocimiento, acelerar procesos de aprendizaje y ampliar las capacidades humanas (Casillas Alvarado et al., 2024). Sin embargo, la actual expansión de la Inteligencia Artificial generativa presenta características diferenciales respecto de otras tecnologías educativas previas, dado que ya no se limita a actuar como un repositorio de información o una herramienta de apoyo, sino que participa activamente en tareas tradicionalmente asociadas al razonamiento humano.

Empiezan así a surgir interrogantes sobre los efectos cognitivos del uso intensivo de IA en estudiantes universitarios. Si bien estas tecnologías pueden mejorar la eficiencia y reducir tiempos de resolución de tareas, también podrían estar modificando la manera en que los individuos desarrollan competencias analíticas, pensamiento crítico y autonomía intelectual.

Para nombrar este conjunto de efectos, la literatura reciente ha comenzado a abordar el fenómeno a través del concepto de “deuda cognitiva”, entendido como la acumulación progresiva de efectos derivados de la delegación excesiva de procesos mentales hacia sistemas externos de Inteligencia Artificial. Kosmyna et al. (2025), cuyo trabajo constituye la referencia central de esta investigación, sostienen que el uso prolongado y acrítico de asistentes de IA puede generar una reducción del esfuerzo mental, menor activación de funciones ejecutivas y dificultades para retomar tareas cognitivas sin asistencia tecnológica.

Este fenómeno es particularmente relevante en la educación superior, donde, como se desarrollará a lo largo del trabajo, el aprendizaje no se limita a la obtención de resultados inmediatos, sino que implica el desarrollo de capacidades cognitivas complejas y transferibles. El riesgo no reside únicamente en la utilización de herramientas de IA, sino en la posibilidad de que la automatización cognitiva sustituya progresivamente procesos fundamentales de reflexión, pensamiento crítico, análisis y construcción conceptual. Parafraseando la Taxonomía de Bloom, un estudiante puede producir un trabajo que aparenta alcanzar niveles de análisis, evaluación o síntesis sin haber atravesado los

procesos cognitivos que esos niveles demandan: la IA los ejecutó en su lugar. Este desacople entre la sofisticación aparente del producto y la ausencia real del proceso es, en esencia, lo que la deuda cognitiva describe.

La problemática adquiere especial relevancia en el contexto argentino, foco de este estudio, donde el uso de herramientas de IA se ha expandido rápidamente entre estudiantes universitarios de grado, particularmente en universidades privadas ubicadas en grandes centros urbanos como Rosario y Buenos Aires. Sin embargo, pese a la creciente utilización de estas tecnologías, aún existe escasa evidencia empírica sobre sus efectos cognitivos concretos en contextos educativos reales.

A esta escasez de evidencia se suma una limitación metodológica importante, que constituye el punto de partida del aporte original de este trabajo: la ausencia de instrumentos prácticos y sistemáticos que permitan medir la deuda cognitiva asociada al uso de IA en estudiantes universitarios. Si bien existen estudios relacionados con carga cognitiva, dependencia tecnológica o automatización del aprendizaje, todavía no se dispone de escalas específicas orientadas a evaluar este fenómeno emergente.

En este marco, el presente trabajo propone desarrollar un instrumento preliminar de medición de deuda cognitiva asociado al uso de Inteligencia Artificial en estudiantes universitarios de grado en Argentina. El estudio se centra en universidades públicas y privadas de Rosario y Buenos Aires y busca identificar dimensiones vinculadas al esfuerzo mental, dependencia tecnológica, frecuencia de uso, autonomía cognitiva y pensamiento crítico.

Complementariamente, la investigación incorpora un componente exploratorio basado en un modelo analítico simple, como la regresión lineal, orientado a analizar relaciones entre variables y explorar la viabilidad de una estimación automatizada del nivel de deuda cognitiva como prueba de concepto.

La presente investigación surge también de la experiencia directa del autor como docente universitario en el ámbito de las tecnologías aplicadas a la gestión y como profesional de la transformación digital en organizaciones de escala global. La

observación cotidiana del uso de herramientas de IA generativa por parte de los estudiantes en contextos académicos reales, combinada con años de experiencia implementando tecnologías en entornos organizacionales complejos, generó una pregunta que los instrumentos disponibles en la literatura no permitían responder con precisión: qué ocurre cognitivamente con quienes delegan de manera sostenida su trabajo intelectual a sistemas automatizados. Esta doble perspectiva, la del educador que observa el fenómeno en el aula y la del profesional que lo ve operar en las organizaciones, define el enfoque situado y aplicado con el que se aborda el problema a lo largo de este trabajo.

De esta manera, la investigación busca aportar evidencia conceptual y metodológica sobre un fenómeno aún poco explorado, contribuyendo tanto al debate académico sobre IA y educación como al desarrollo de estrategias orientadas a promover un uso más crítico y equilibrado de estas tecnologías en el ámbito universitario.

2. Antecedentes y Planteamiento del Problema

Antecedentes y planteamiento de la problemática

La Inteligencia Artificial generativa, particularmente los modelos de lenguaje de gran escala más usados de ChatGPT, Claude, Deepseek, Perplexity, Gemini o Copilot, están transformando de manera acelerada la forma en que las personas aprenden, producen conocimiento y resuelven problemas complejos. Tal como señalan Brynjolfsson y McAfee (2024), la incorporación masiva de sistemas de IA en tareas cognitivas ha desplazado procesos que históricamente dependían del razonamiento humano, generando importantes ganancias de eficiencia, pero también modificando en profundidad la dinámica del trabajo intelectual.

En el ámbito de la educación superior, la IA ha introducido una nueva lógica de aprendizaje asistido, en la cual los estudiantes pueden generar textos, resolver ejercicios, sintetizar información o estructurar ideas en cuestión de segundos. Esta automatización cognitiva ofrece ventajas evidentes en términos de productividad, acceso al conocimiento y reducción de tiempos, pero también potencia un

fenómeno que la literatura reciente comienza a describir con mayor precisión: la delegación del pensamiento humano hacia sistemas de IA (Luckin, 2018).

Antes que promover procesos profundos de análisis, reflexión y construcción conceptual, el uso acrítico de estas herramientas puede derivar en una participación cada vez más pasiva del estudiante en los procesos cognitivos. En este contexto, el aprendizaje corre el riesgo de orientarse hacia la obtención de resultados inmediatos, en detrimento del desarrollo de competencias cognitivas complejas y transferibles.

Un aporte central en esta discusión es el concepto de “deuda cognitiva”, desarrollado por Kosmyna et al. (2025). A partir de estudios experimentales y neurocognitivos, los autores evidencian que el uso prolongado y acrítico de asistentes de IA se asocia con una reducción del esfuerzo mental, una disminución de la activación cerebral vinculada a funciones ejecutivas y una menor sensación de agencia intelectual. Cuando los usuarios ya acostumbrados al uso constante de modelos de LLMs deben retomar tareas sin asistencia tecnológica, experimentan frustración, dependencia y una pérdida temporal de autonomía cognitiva. Este fenómeno no se limita a una reducción momentánea de la carga cognitiva, sino que se manifiesta como una acumulación de efectos residuales producto de la delegación sistemática del pensamiento en sistemas externos, y una dependencia creciente.

En la educación superior de grado, estos patrones son especialmente relevantes. Fan, Deng y Liu (2025) demuestran que, aunque los estudiantes perciben beneficios inmediatos en términos de eficiencia y productividad, dichas percepciones no siempre se traducen en mejoras reales del aprendizaje. Se observa una brecha entre la eficiencia percibida y la profundidad cognitiva alcanzada, lo que impacta negativamente en el desarrollo de habilidades analíticas, metacognitivas y de pensamiento crítico.

En el contexto argentinol uso de herramientas de IA se ha incorporado de manera intensiva en las prácticas laborales y académicas cotidianas. Sin embargo, persiste una limitación significativa en términos metodológicos: la ausencia de instrumentos prácticos y sistemáticos que permitan medir la deuda cognitiva de

manera cuantitativa en estudiantes universitarios, sin recurrir a entornos experimentales controlados.

2.1 Historia de la adopción tecnológica en educación superior

La incorporación de tecnologías en la educación superior no es un fenómeno reciente. Desde finales del siglo XX, universidades e instituciones educativas comenzaron a integrar plataformas virtuales, sistemas de gestión del aprendizaje, herramientas multimedia y recursos digitales orientados a optimizar los procesos de enseñanza y aprendizaje. Sin embargo, la expansión de la Inteligencia Artificial generativa representa una nueva etapa dentro de esta evolución tecnológica.

Según Casillas Alvarado, Ramírez Martinell y López-Olivera López (2024), los procesos históricos de adopción tecnológica en educación han seguido patrones progresivos asociados a la incorporación de herramientas que inicialmente complementaban el aprendizaje y posteriormente transformaban las prácticas educativas. En este sentido, la IA generativa introduce una diferencia sustancial respecto de tecnologías anteriores: su capacidad para producir contenido original, interactuar mediante lenguaje natural y resolver tareas complejas tradicionalmente asociadas al razonamiento humano.

En esta misma línea, Luckin (2018) sostiene que la Inteligencia Artificial posee el potencial de redefinir profundamente los modelos educativos contemporáneos. La posibilidad de personalizar procesos de aprendizaje, adaptar contenidos y automatizar tareas rutinarias representa una oportunidad significativa para ampliar el acceso al conocimiento y optimizar la experiencia educativa. No obstante, esta misma autora, junto con otros, advierte que la transformación introduce riesgos asociados a la dependencia tecnológica y a la sustitución progresiva de procesos cognitivos esenciales, un punto que se profundizará a lo largo del presente capítulo.

Esa advertencia se ve reforzada por Selwyn (2019), quien observa que muchas discusiones sobre IA y educación tienden a centrarse exclusivamente en sus beneficios operativos, dejando en segundo plano las implicancias sociales, cognitivas y pedagógicas derivadas de la automatización del aprendizaje. El autor

plantea que la adopción acrítica de sistemas inteligentes puede promover formas de aprendizaje superficial, excesivamente orientadas a la eficiencia y a la obtención inmediata de resultados, un diagnóstico que conecta directamente con el concepto de deuda cognitiva que se desarrolla más adelante en el marco teórico.

Holmes, Bialik y Fadel (2019) matizan este escenario sosteniendo que la IA puede potenciar procesos educativos siempre que exista una integración pedagógica adecuada. Sin embargo, coincidiendo con Selwyn, advierten que la automatización excesiva puede reducir la participación activa del estudiante en procesos fundamentales como el análisis crítico, la resolución autónoma de problemas y la construcción reflexiva del conocimiento. Este punto es central para la presente investigación, dado que la hipótesis que la orienta sostiene precisamente que la delegación intensiva del trabajo cognitivo hacia la IA produce un costo acumulativo que el sistema educativo todavía no está midiendo.

La preocupación se vuelve más concreta a la luz del trabajo de Feng et al. (2025), quienes a partir de una revisión sistemática de una década de investigación sobre adopción tecnológica en educación superior (2015-2024) identifican un patrón preocupante: las instituciones educativas tienden a priorizar criterios de eficiencia, modernización y reducción de costos por sobre evaluaciones profundas acerca de las consecuencias cognitivas y pedagógicas de las tecnologías que adoptan. Esta tendencia genera escenarios donde la innovación tecnológica precede sistemáticamente a la investigación sobre sus efectos reales, lo cual dificulta el diseño de respuestas pedagógicas adecuadas y refuerza el vacío metodológico que el presente trabajo busca contribuir a cerrar.

Esta lectura se ve consolidada por Valverde-Berrocoso, Acevedo-Borrega y Cerezo-Pizarro (2022), quienes en una revisión sistemática sobre tecnología educativa y rendimiento estudiantil concluyen que la mera incorporación de tecnología no garantiza mejoras en el aprendizaje: los efectos positivos dependen críticamente de la integración pedagógica, el diseño instruccional y la capacitación docente. Esta conclusión es especialmente relevante para el presente estudio, dado que, como se desarrolla en la sección 2.3, el uso actual de IA generativa en

las universidades argentinas se produce, en términos generales, sin mediación pedagógica sistemática.

2.2 La IA generativa como ruptura cualitativa

Como se anticipó en la sección anterior, la diferencia entre la IA generativa y las tecnologías educativas previas no es de grado: es de naturaleza. La masificación de herramientas como ChatGPT y Claude aceleraron significativamente estos debates. La facilidad con la que los estudiantes puedan producir textos académicos, resolver ejercicios o sintetizar información mediante IA ha generado preocupación en docentes e instituciones educativas respecto de la autenticidad del aprendizaje y el desarrollo real de competencias cognitivas, una preocupación que el autor de este trabajo ha podido constatar de manera directa en su práctica docente universitaria.

En este contexto emerge el concepto de “delegación cognitiva”, entendido como la transferencia progresiva de tareas intelectuales desde el individuo hacia sistemas tecnológicos externos. Si bien la delegación cognitiva ha existido históricamente a través de calculadoras, motores de búsqueda o software especializado, la IA generativa amplifica este fenómeno debido a su capacidad para simular razonamientos complejos y producir respuestas altamente sofisticadas. Esta distinción, apenas esbozada aquí, constituye uno de los ejes que se desarrollan en profundidad en el marco teórico, particularmente en la sección 5.2.

Es precisamente sobre esta base que cobra relevancia el aporte central de esta discusión: el trabajo de Kosmyna et al. (2025), quienes introducen el concepto de “deuda cognitiva” para describir los efectos acumulativos derivados del uso intensivo de asistentes de IA en tareas cognitivas. A partir de estudios experimentales vinculados a escritura asistida mediante IA, los autores identificaron una disminución significativa de la activación cognitiva y del esfuerzo mental en usuarios que delegaban de manera constante procesos intelectuales en sistemas automatizados. Este estudio, que da nombre y sustento empírico al concepto que articula toda la presente investigación, será retomado y analizado en detalle en la sección 5.3.

Los hallazgos de Kosmyna et al. no son un caso aislado. Fan, Deng y Liu (2025), en un estudio realizado con estudiantes de ingeniería en China, concluyen en la misma dirección: la IA generativa produce mejoras percibidas en productividad y rapidez de resolución de tareas, aunque estas mejoras no necesariamente se traducen en una comprensión más profunda de los contenidos. Los autores identifican una brecha entre la eficiencia percibida y la calidad cognitiva real del aprendizaje, brecha que es central para comprender por qué la deuda cognitiva puede pasar desapercibida tanto para el estudiante como para la institución. Esta brecha se profundiza aún más si se considera que, como demostraron Laban et al. (2025) en un estudio premiado en ICLR 2026, los LLMs exhiben un rendimiento un 39% inferior en conversaciones de múltiples turnos respecto de consultas de turno único, y que una vez que toman un camino equivocado en la conversación, no se recuperan. Dado que los estudiantes interactúan con IA predominantemente en modo conversacional e iterativo, la calidad de las respuestas que efectivamente consumen es probablemente inferior a la que asocian con la herramienta, lo que convierte la brecha entre eficiencia percibida y aprendizaje real en un fenómeno de doble capa.

De manera complementaria, y reforzando esta línea de evidencia, Lubbe, Marais y Kruger (2025) analizan el impacto de la IA generativa sobre las habilidades de pensamiento de orden superior, en articulación con la Taxonomía de Bloom. Sus resultados sugieren que el uso frecuente de IA puede afectar procesos vinculados a creatividad, razonamiento analítico y elaboración conceptual cuando no existe mediación pedagógica adecuada, una conclusión que dialoga directamente con lo señalado por Holmes, Bialik y Fadel (2019) en la sección anterior respecto de la necesidad de una integración pedagógica que no quede librada al azar.

Es por esto que incluso estas preocupaciones de la literatura académica son compartidas por los organismos internacionales. Ya en 2021, UNESCO (2021) advirtió sobre la necesidad de desarrollar marcos regulatorios y pedagógicos orientados a garantizar un uso ético y equilibrado de la Inteligencia Artificial en educación, y sostiene que la IA debe complementar el aprendizaje humano y no reemplazar procesos esenciales de razonamiento, creatividad y pensamiento

crítico. Posteriormente, el *Global Education Monitoring Report* 2023 de UNESCO profundizó esta preocupación al señalar que muchas tecnologías educativas son implementadas sin suficiente evidencia sobre sus impactos reales, lo que coincide de manera notable con el patrón que Feng et al. (2025) describieron en la sección anterior: la innovación tecnológica tiende a preceder a la investigación sobre sus efectos.

2.3 Contexto argentino y planteamiento del problema central

Todo lo expuesto hasta aquí adquiere una dimensión particular al trasladarlo al contexto local. En América Latina y particularmente en Argentina, el debate sobre IA y educación superior todavía se encuentra en una etapa temprana y poco madura. La mayoría de las investigaciones locales se concentran en aspectos vinculados a innovación tecnológica, transformación digital o herramientas de apoyo al aprendizaje, mientras que los impactos cognitivos derivados del uso intensivo de IA, que son el foco de este trabajo, continúan escasamente explorados.

Esta brecha resulta especialmente notoria en el ámbito que motiva la presente investigación. En base a las poblaciones encuestadas, se puede confirmar lo siguiente: en universidades públicas y privadas de Rosario y Buenos Aires, el uso cotidiano de herramientas de IA generativa se ha expandido rápidamente entre estudiantes de grado. La utilización de IAs para redactar trabajos prácticos, generar resúmenes, resolver ejercicios o estructurar ideas se ha naturalizado en numerosos contextos académicos, un fenómeno observable como docente. Sin embargo, y tal como se viene argumentando desde el inicio de este capítulo, esta incorporación acelerada no ha sido acompañada por instrumentos capaces de evaluar cómo estas tecnologías afectan dimensiones cognitivas fundamentales.

Este vacío metodológico contrasta con un creciente debate académico sobre la inteligencia artificial en Argentina, y en el mundo, en el que comienzan a configurarse voces locales con peso propio. Una contribución particularmente relevante en este sentido es Ok, Pandora (López et al., 2024), volumen colectivo publicado por El Gato y La Caja que reúne ensayos de seis investigadores

argentinos provenientes de la filosofía, la neurociencia, la computación y las ciencias sociales. Significativamente, el ensayo del neurocientífico Enzo Tagliazucchi, investigador del CONICET en el INFINA y especialista en neurociencia de la cognición, converge de manera directa con la línea empírica desarrollada por Kosmyna et al. (2025): ambos coinciden en que la cuestión central frente a la IA generativa no se juega en las capacidades observables del sistema artificial, sino en los procesos cognitivos subyacentes, tanto los que la máquina simula como, sobre todo, los que el ser humano deja de ejercitar al interactuar con ella. Esta convergencia entre la evidencia neurocognitiva internacional y la producción reflexiva argentina refuerza la consistencia de la hipótesis que orienta el presente trabajo y le otorga anclaje local. De manera complementaria, el análisis de Carolina Aguerre sobre gobernanza tecnológica en América Latina, contenido en el mismo volumen, advierte que las instituciones de la región vienen adoptando estas herramientas sin haber construido aún marcos para evaluar sus impactos cognitivos y sociales, un diagnóstico que coincide con lo señalado por Feng et al. (2025) y por Valverde-Berrocoso et al. (2022) y que justifica la urgencia de estudios situados como el que aquí se propone.

Problema central

De todo lo anterior se desprende el problema que esta investigación busca abordar. A pesar de la creciente relevancia de la deuda cognitiva asociada al uso de Inteligencia Artificial en la educación superior, y pese a la convergencia de evidencia internacional y local reseñada en los párrafos precedentes, las herramientas disponibles para evaluar su impacto siguen siendo limitadas. Existen estudios experimentales, análisis conceptuales y escalas vinculadas a carga cognitiva o dependencia tecnológica, pero aún no se cuenta con un instrumento específico, práctico y contextualizado que permita medir la deuda cognitiva derivada del uso cotidiano de IA en estudiantes universitarios de grado.

La ausencia de indicadores cuantitativos concretos dificulta la identificación de niveles de riesgo, la planificación de intervenciones pedagógicas y el diseño de estrategias educativas orientadas a un uso responsable de la Inteligencia Artificial. Asimismo, la falta de mecanismos de análisis automatizado limita la posibilidad de

identificar patrones cognitivos emergentes a partir de datos empíricos, incluso en estudios de alcance exploratorio. En este sentido, se vuelve necesario desarrollar un instrumento de evaluación que permita no solo medir la deuda cognitiva, sino también explorar su análisis sistemático mediante modelos analíticos simples.

La ausencia de indicadores específicos dificulta identificar niveles de dependencia tecnológica, pérdida de autonomía intelectual o reducción del esfuerzo mental. En consecuencia, surge la necesidad de desarrollar herramientas metodológicas que permitan operacionalizar el concepto de deuda cognitiva y transformarlo en un fenómeno medible dentro de contextos educativos reales.

3. Objetivos

3.1 Objetivo General

Desarrollar indicadores de medición de la deuda cognitiva resultante del uso de IA en estudiantes de educación superior de grado en la Argentina, y explorar la viabilidad de su análisis automatizado mediante modelos analíticos simples.

3.2 Objetivos Específicos

1. Identificar las manifestaciones, causas y dimensiones centrales de la deuda cognitiva asociada al uso de IA en estudiantes universitarios de grado.
2. Analizar el impacto de la delegación cognitiva en los procesos de aprendizaje y pensamiento crítico en la educación superior argentina.
3. Diseñar un cuestionario para medir dimensiones de deuda cognitiva en estudiantes de universidades públicas y privadas.
4. Establecer lineamientos preliminares de validación del instrumento, considerando fiabilidad y consistencia interna.
5. Explorar el uso de un modelo analítico simple, como la regresión lineal, para estimar niveles de deuda cognitiva de manera automatizada en muestras acotadas.

6. Proponer estrategias educativas orientadas a mitigar la deuda cognitiva en contextos universitarios de grado.

4. Justificación

La deuda cognitiva asociada al uso de Inteligencia Artificial constituye un fenómeno emergente que aún no cuenta con suficientes estudios empíricos ni con instrumentos de medición específicos, particularmente en el contexto de la educación superior argentina. A diferencia de la carga cognitiva tradicional, la deuda cognitiva se acumula a lo largo del tiempo como resultado del uso intensivo y prolongado de herramientas de IA, afectando dimensiones clave como el esfuerzo mental, la autonomía cognitiva, la atención sostenida y el pensamiento crítico.

Este estudio aporta conocimiento actualizado sobre un fenómeno aún poco explorado, articulando evidencia neurocognitiva y educativa en el ámbito de la educación superior de grado. Asimismo, introduce un instrumento original para medir la deuda cognitiva en estudiantes universitarios, lo cual representa un aporte innovador para el análisis del impacto cognitivo de la IA en contextos académicos reales.

La incorporación de modelos analíticos simples como complemento del instrumento de medición permite explorar nuevas formas de análisis del fenómeno, sin pretender desarrollar sistemas predictivos complejos. Este enfoque metodológico contribuye a sentar las bases para futuras investigaciones que integren medición cognitiva y análisis de datos en el estudio del uso de la Inteligencia Artificial en el aprendizaje universitario.

La expansión acelerada de la Inteligencia Artificial generativa en contextos universitarios ha generado profundas transformaciones en la forma en que los estudiantes producen conocimiento, resuelven problemas y desarrollan tareas académicas. No obstante, mientras gran parte del debate académico y tecnológico se concentra en las ventajas operativas de estas herramientas, todavía existe

escasa evidencia respecto de sus consecuencias cognitivas a mediano y largo plazo.

Desde el punto de vista científico, la investigación es pertinente porque el trabajo seminal de Kosmyna et al. (2025) fue realizado en condiciones de laboratorio y en un contexto cultural y educativo específico. La generalización de sus hallazgos a otros contextos, como la educación superior latinoamericana, requiere investigaciones situadas que consideren las particularidades institucionales, culturales y tecnológicas de cada entorno. El presente trabajo contribuye a cerrar esta brecha al abordar el fenómeno desde el contexto argentino.

Desde el punto de vista metodológico, la investigación propone un aporte original mediante el diseño preliminar de un cuestionario estructurado orientado a medir dimensiones vinculadas al esfuerzo mental, dependencia tecnológica, frecuencia de uso, revisión crítica y autonomía cognitiva. Si bien existen escalas relacionadas con carga cognitiva, dependencia tecnológica o fatiga mental, todavía no se dispone de herramientas diseñadas específicamente para evaluar este fenómeno emergente en estudiantes universitarios.

Desde el punto de vista pedagógico e institucional, las universidades enfrentan actualmente el desafío de integrar herramientas de IA sin comprometer el desarrollo de capacidades cognitivas fundamentales. Comprender los riesgos asociados a la dependencia tecnológica puede contribuir al diseño de estrategias educativas orientadas a promover un uso más reflexivo, crítico y equilibrado de estas tecnologías. Finalmente, el trabajo busca aportar evidencia que permita enriquecer el debate académico sobre IA y educación superior, contribuyendo a una comprensión más profunda de las transformaciones cognitivas asociadas a la automatización del aprendizaje.

5. Marco Teórico

El presente Marco Teórico se organiza en torno a cinco ejes conceptuales que articulan la base teórica de la investigación. El primero contextualiza la expansión de la IA más allá de la educación y la dependencia que esta ubicuidad genera. El

segundo aborda la IA generativa como fenómeno de delegación cognitiva. El tercero desarrolla el concepto de deuda cognitiva, sus fundamentos neurocognitivos y su relación con las teorías clásicas sobre carga cognitiva, pensamiento dual y metacognición. El cuarto analiza los riesgos cognitivos específicos que el uso intensivo de IA genera en la educación superior. El quinto examina las competencias cognitivas que la era digital demanda preservar y desarrollar.

5.1 La expansión de la IA y la dependencia emergente

El interés por estudiar la deuda cognitiva en la educación superior surge de una observación más amplia: la Inteligencia Artificial generativa dejó de ser una herramienta especializada para convertirse en una capa presente en casi todos los productos y servicios que las personas usan a diario. En pocos años, asistentes como ChatGPT, Claude, Gemini, Microsoft Copilot, Perplexity, Grok, Meta AI, Mistral, DeepSeek, Cursor y otros sistemas basados en inteligencia artificial generativa han transformado la manera en que las personas interactúan con la información y la tecnología, permitiendo automatizar tareas complejas, mejorar la productividad y acceder a conocimientos especializados de forma rápida y accesible. Estos asistentes pasaron de ser aplicaciones independientes a integrarse de manera silenciosa en buscadores, suites de oficina, plataformas de mensajería, herramientas de programación, servicios de atención al cliente y dispositivos móviles. Esta integración acelerada hizo que la consulta a un sistema de IA dejara de ser una decisión deliberada para transformarse en un reflejo cotidiano, muchas veces invisible para quien lo realiza.

Esa ubicuidad genera un fenómeno que excede el ámbito educativo. A medida que la asistencia automatizada se vuelve la opción por defecto, las personas delegan de a poco tareas que antes resolvían solas: redactar un correo, resumir un texto, escribir una fórmula, planificar una actividad, tomar una decisión sencilla. La conveniencia inmediata es real. Pero su acumulación silenciosa puede derivar en una dependencia funcional, es decir, en la dificultad creciente para sostener esas mismas tareas sin asistencia tecnológica. La dinámica recuerda al artículo de

1998, La Mente Extendida, de Clark y Chalmers, según la cual los procesos cognitivos se apoyan en recursos externos hasta volverse inseparables de ellos. La diferencia con las tecnologías anteriores es cualitativa: aquellas almacenaban o recuperaban información; la IA generativa ejecuta el propio acto de pensar que antes correspondía a la persona.

Esta preocupación dejó de ser especulativa. En un estudio en Reino Unido con más de 600 participantes de distintas edades y niveles educativos, Gerlich (2025) halló una correlación negativa significativa entre el uso frecuente de herramientas de IA y la capacidad de pensamiento crítico, mediada por el fenómeno de descarga cognitiva, es decir, la tendencia a delegar en la herramienta el esfuerzo mental que antes asumía la persona. El estudio observó, además, que los participantes más jóvenes presentaban una mayor dependencia de la IA y puntajes más bajos de pensamiento crítico que los de mayor edad. En el ámbito laboral, Lee et al. (2025), en una investigación de Microsoft Research y la Universidad Carnegie Mellon sobre más de 300 trabajadores del conocimiento, encontraron un patrón complementario: cuanta mayor confianza depositaban los profesionales en la IA, menor era el esfuerzo de pensamiento crítico que reportaban, desplazando su actividad mental desde la resolución de problemas hacia la mera supervisión de las respuestas generadas.

El concepto de descarga cognitiva (*cognitive offloading*) que Gerlich emplea no es nuevo: la psicología cognitiva lo estudia desde antes de la IA generativa. Risko y Gilbert (2016) la definieron formalmente como *“the use of physical action to alter the information processing requirements of a task so as to reduce cognitive demand”* (p. 676), es decir, el uso de acciones o recursos externos para reducir las demandas cognitivas de una tarea, desde anotar algo para no memorizarlo hasta delegar un cálculo en una herramienta, y mostraron que, si bien suele ser una estrategia eficiente, tiene un costo cuando reemplaza el procesamiento que sostiene el aprendizaje. Un antecedente influyente es el trabajo de Sparrow, Liu y Wegner (2011) sobre los “efectos Google”: hallaron que, cuando las personas saben que una información quedará disponible en un buscador, la recuerdan peor y, en cambio, recuerdan mejor dónde encontrarla; la memoria se reorganiza para

almacenar el acceso a la información antes que la información misma. En palabras de los autores, *“when people expect to have future access to information, they have lower rates of recall of the information itself and enhanced recall instead for where to access it”* (Sparrow et al., 2011, p. 776). Puesto en términos sencillos, cuando damos por sentado que vamos a poder volver a buscar una información, la recordamos menos y, en cambio, recordamos mejor dónde encontrarla. La deuda cognitiva puede entenderse como la versión amplificada y acumulativa de este fenómeno: si el “efecto Google” afectaba sobre todo la memoria de datos, la IA generativa extiende la descarga a procesos de orden superior, analizar, argumentar, redactar, resolver, que son justamente los que la educación universitaria busca desarrollar.

La evidencia neurocientífica apunta en la misma dirección. El estudio del MIT Media Lab dirigido por Kosmyna et al. (2025), que constituye una de las referencias centrales de esta tesis, registró mediante electroencefalografía la actividad cerebral de 54 participantes mientras escribían ensayos con asistencia de IA, con un buscador o sin ninguna herramienta. Los autores hallaron que quienes utilizaron IA mostraron la menor conectividad neuronal, una memoria más débil de lo que habían escrito y un menor sentido de autoría sobre su propio trabajo, efectos que persistían incluso después de dejar de usar la herramienta. Es justamente esta acumulación de efectos residuales la que dio nombre al concepto de deuda cognitiva que esta investigación retoma y busca medir en el contexto argentino.

A esta dependencia individual se suma un patrón estructural de otra escala: la sobreutilización de la inteligencia artificial como respuesta automática a problemas de toda naturaleza. En el diseño actual de muchos productos y servicios digitales, incorporar un asistente de IA se ha convertido en una norma comercial más que en una necesidad funcional. Herramientas de oficina, plataformas de comunicación, motores de búsqueda, sistemas de gestión empresarial y aplicaciones educativas incorporan capacidades de IA de forma cada vez más visible e insistente. En muchos casos, la pregunta ya no es si una tarea requiere inteligencia artificial, sino cómo integrar inteligencia artificial a una tarea.

Por mencionar un ejemplo, en el ámbito laboral, esta tendencia genera una presión adicional sobre los trabajadores. A medida que las organizaciones adoptan estas tecnologías para mejorar la productividad, reducir costos y acelerar procesos, el uso de asistentes inteligentes deja de ser una opción para convertirse progresivamente en una expectativa implícita. Al igual que ocurrió con otras transformaciones tecnológicas del pasado, como la calculadora, el *boom* de Internet o los *smartphones*, la adaptación se presenta como una condición para mantener la competitividad profesional. Quienes no incorporan estas herramientas corren el riesgo de quedar rezagados frente a colegas capaces de producir más rápido y en mayor volumen.

Sin embargo, la inteligencia artificial introduce una diferencia relevante respecto de muchas tecnologías anteriores. Mientras que la mayoría de las innovaciones históricas ampliaban la capacidad humana para ejecutar tareas, la IA tiene la capacidad de asumir procesos cognitivos que tradicionalmente requerían razonamiento, análisis, síntesis, creatividad o toma de decisiones. Esto desplaza parte de la actividad mental desde la persona hacia el sistema tecnológico. Es por esto que el riesgo no radica en utilizar la herramienta, sino en la frecuencia y profundidad con la que se delegan funciones intelectuales fundamentales.

De este modo, el usuario no solo elige apoyarse en la IA por conveniencia, sino que además opera dentro de un ecosistema diseñado para incentivar dicho apoyo de manera constante. Cada sugerencia automática, resumen generado, respuesta propuesta o contenido redactado por el sistema reduce la necesidad inmediata de realizar el esfuerzo cognitivo correspondiente. En consecuencia, el desafío contemporáneo no consiste únicamente en determinar qué tareas pueden ser automatizadas, sino también en identificar cuáles deberían seguir requiriendo participación cognitiva activa por parte de las personas. La cuestión central deja de ser tecnológica y pasa a ser humana: cómo aprovechar los beneficios de la inteligencia artificial sin convertir la delegación cognitiva en una sustitución sistemática del pensamiento.

El proceso de enseñanza-aprendizaje en la educación media y superior es un escenario donde estas tendencias se manifiestan con particular nitidez y con

consecuencias específicas. En la vida cotidiana, delegar tareas rutinarias puede ser inocuo o incluso cómodo. En el aula no: ahí el objetivo central es desarrollar las capacidades cognitivas que la IA tiende a sustituir, como la elaboración de argumentos, la síntesis crítica, la resolución autónoma de problemas y el razonamiento sostenido. Por eso la deuda cognitiva preocupa más aquí que en otros ámbitos: incide de lleno sobre el aprendizaje y sobre la formación de las competencias que la universidad tiene la responsabilidad de cultivar. Esa tensión, entre una tecnología que se vuelve ubicua y cómoda en todos los planos de la vida y un ámbito cuya misión es formar la autonomía intelectual, motiva y enmarca esta investigación. Los apartados siguientes profundizan en los mecanismos cognitivos que lo explican.

El marco teórico aporta además un mecanismo que sirve como clave de lectura para toda la investigación. La literatura sobre interacción humano-IA describe la llamada “Paradoja de la confianza en la IA”. Esta explica que a medida que los modelos se vuelven más capaces y sus respuestas más fluidas y convincentes, los usuarios tienden a confiar más y a verificar menos. La mayor calidad del sistema, entonces, no reduce la dependencia; podría incluso intensificarla (Ibrahim et al., 2025). Bajo esta lente, la deuda cognitiva deja de ser un efecto colateral de modelos imperfectos que el avance tecnológico iría corrigiendo. Pasa a ser un fenómeno que se agudiza a medida que las herramientas mejoran, porque un asistente más competente reduce con más eficacia la necesidad percibida de esfuerzo propio. Esta clave interpretativa, que se retoma en las conclusiones, apunta a que el problema es estructural y no pasajero, y a que su relevancia podría crecer con el tiempo.

5.2 IA generativa y delegación cognitiva

La Inteligencia Artificial generativa constituye una de las innovaciones tecnológicas más relevantes de los últimos años debido a su capacidad para producir contenido original mediante procesamiento avanzado del lenguaje natural. A diferencia de sistemas tradicionales orientados únicamente a automatizar tareas mecánicas o repetitivas, los modelos generativos actuales pueden redactar textos complejos,

sintetizar información, responder preguntas y resolver problemas de manera aparentemente similar al razonamiento humano.

El desarrollo de modelos de lenguaje de gran escala (*Large Language Models* o LLMs) representa un cambio significativo en la relación entre seres humanos y tecnología. Los LLMs no solo facilitan el acceso a información, sino que participan activamente en tareas cognitivas vinculadas a escritura, comprensión conceptual y producción intelectual. Cada uno de estos sistemas presenta particularidades en términos de capacidades, sesgos y modos de interacción, pero todos comparten la característica común de operar como agentes cognitivos capaces de generar respuestas elaboradas a partir de instrucciones en lenguaje natural.

Brynjolfsson y McAfee (2024), ya citados previamente, sostienen que la actual revolución asociada a la IA posee características comparables a las grandes transformaciones tecnológicas de la historia moderna. Los autores afirman que la automatización ya no se limita al trabajo físico o administrativo, sino que comienza a impactar directamente sobre actividades cognitivas avanzadas. En educación superior, esta transformación adquiere especial relevancia porque el proceso de aprendizaje depende profundamente del desarrollo de capacidades intelectuales complejas.

No está de más mencionar que esta expansión de la IA generativa excede largamente el ámbito educativo, está reconfigurando también los procesos organizacionales y los mercados laborales a una velocidad sin precedentes, por ejemplo, procesos de reclutamiento de personal. Xu, Li y Jiang (2025) aportan evidencia empírica de un fenómeno emergente con implicancias directas sobre la equidad en las organizaciones: el sesgo de autopreferencia de los LLMs en los procesos de selección de personal. Los autores documentan que, cuando las empresas usan LLMs para evaluar candidatos cuyos currículos fueron a su vez redactados con IA, los modelos tienden a favorecer los documentos generados por el mismo modelo que actúa como evaluador. El experimento fue a gran escala: un experimento de correspondencia de currículos que simuló procesos de selección en veinticuatro ocupaciones. Los candidatos que usaron el mismo LLM que el evaluador tuvieron entre un 23% y un 60% más de probabilidades de ser

preseleccionados que candidatos igualmente calificados cuyos currículos había escrito un humano, con las brechas más pronunciadas en áreas como ventas y contabilidad. El hallazgo ilustra un riesgo que va más allá de los sesgos demográficos tradicionalmente estudiados: la IA puede generar ventajas y desventajas sistémicas según qué herramienta usa cada individuo, lo que abre interrogantes de fondo sobre la equidad y la autonomía intelectual en un mundo cada vez más mediado por sistemas automatizados. La conexión con la deuda cognitiva es directa. Si quienes delegan su producción intelectual a la IA obtienen ventajas institucionales concretas, los incentivos para esa delegación se amplifican, y con ellos los riesgos de erosión de la autonomía intelectual en los entornos formativos.

Ahora bien, para entender por qué la delegación a la IA es cualitativamente distinta de otras formas de externalización tecnológica, ayuda recurrir a la filosofía de la mente. Clark y Chalmers (1998) conceptualizaron el fenómeno con profundidad. Su artículo, de hecho, se abre con una pregunta que condensa todo el problema: *"Where does the mind stop and the rest of the world begin?"* (Clark y Chalmers, 1998, p. 7). Dicho de otro modo, los autores se preguntan dónde termina la mente y dónde empieza el mundo, anticipando su idea de que ese límite no está fijado en el cerebro, sino que se extiende hacia las herramientas que usamos. En su influyente artículo "*The Extended Mind*" sostuvieron que la mente humana no termina en el cerebro: se extiende hacia el entorno a través de las herramientas que usa. Bajo esta Tesis de la Mente Extendida, un recurso externo puede formar parte genuina del sistema cognitivo de un individuo si cumple la función de almacenar y procesar información que de otro modo harían procesos biológicos internos. La IA generativa sería, desde esta perspectiva, la forma más radical de extensión cognitiva jamás disponible: un sistema capaz de realizar casi cualquier proceso mental que un estudiante universitario pueda necesitar. Esa misma extensión se vuelve problemática cuando genera dependencia y atrofia las capacidades internas que el sistema educativo tiene la responsabilidad de cultivar.

Esta problemática se profundiza al notar que la IA generativa, desde la perspectiva del usuario, dista de ser una herramienta pasiva comparable a una calculadora o

un buscador web. Guzman y Lewis (2020), en su propuesta de una agenda de investigación sobre comunicación humano-máquina (*Human-Machine Communication*, HMC), señalan que las interacciones de las personas con sistemas de IA mediante agentes conversacionales y software de generación de lenguaje desbordan los paradigmas clásicos de la teoría de la comunicación, que asumieron históricamente que la comunicación significativa ocurre entre seres humanos. Para los autores, la IA generativa funciona como un agente comunicativo: genera significado, responde de manera contextual y simula la presencia de un interlocutor con criterio propio. La distinción tiene consecuencias directas sobre los procesos cognitivos de quien interactúa con ella. Cuando un estudiante dialoga con un LLM como si fuera un interlocutor inteligente, la barrera que separa la asistencia tecnológica de la delegación del razonamiento se vuelve cada vez más difusa. El carácter conversacional reduce la sensación de que se está externalizando el pensamiento, y así amplifica en silencio el riesgo de acumular deuda cognitiva.

Las dos tesis, la de Clark y Chalmers y la de Guzman y Lewis, a pesar de tener casi 20 años de historia entre medio, se complementan perfectamente. La primera explica que la IA puede integrarse al sistema cognitivo del estudiante, y la segunda explica por qué esa integración ocurre con mayor facilidad y menor resistencia consciente que con cualquier tecnología anterior. Se suma una dimensión técnica que agrava el cuadro: Laban et al. (2025) demostraron que los LLMs rinden bastante peor en conversaciones de múltiples turnos, justo el modo de interacción que Guzman y Lewis describen como el característico de la comunicación humano-máquina con IA generativa. La misma cualidad comunicativa que hace que la delegación pase desapercibida, el diálogo iterativo, contextual y de apariencia inteligente, es la que degrada la calidad del output recibido. El estudiante conversa con un sistema que simula comprensión, pero que pierde precisión a medida que la conversación se extiende. Esa convergencia entre el análisis comunicacional de Guzman y Lewis y la evidencia empírica de Laban y sus colegas apunta a algo incómodo: la delegación cognitiva mediada por IA generativa es más sutil y riesgosa de lo que advertirían, por separado, las miradas puramente técnicas o puramente pedagógicas.

Frente a este panorama, es justo reconocer que esta automatización también puede generar beneficios importantes cuando es utilizada como herramienta de apoyo complementario. Es interesante citar nuevamente a Holmes, Bialik y Fadel (2019) debido a que ellos señalan que la IA puede contribuir a personalizar procesos educativos, optimizar tiempos de aprendizaje y ampliar el acceso al conocimiento. Sin embargo, y este es el punto que articula toda la presente investigación, el problema emerge cuando estas herramientas dejan de complementar el razonamiento humano y comienzan a sustituir procesos fundamentales de reflexión y construcción conceptual.

Llegado este punto importa precisar el concepto de “delegación cognitiva”, el cual se entiende como la transferencia progresiva de funciones intelectuales desde el individuo hacia sistemas tecnológicos externos. No es un fenómeno enteramente nuevo. Distintas tecnologías permitieron históricamente externalizar tareas mentales: las calculadoras redujeron la necesidad de hacer cálculos complejos a mano; los motores de búsqueda cambiaron la forma de acceder a la información y almacenarla; los sistemas de navegación GPS transformaron la orientación espacial, etc. La diferencia, tal como adelantaron Clark y Chalmers y Guzman y Lewis en los párrafos anteriores, es que la IA generativa introduce un salto cualitativo: interviene en tareas tradicionalmente asociadas a la creatividad, el análisis y el razonamiento.

Retomando a Luckin (2018), esta autora plantea que la IA redefine la frontera entre inteligencia humana y automatización tecnológica. Sostiene que el desafío contemporáneo no consiste únicamente en utilizar tecnología, sino en determinar qué procesos deben permanecer bajo control humano y cuáles pueden ser delegados de manera segura. Propone el concepto de “inteligencia aumentada” como horizonte normativo: el objetivo no debería ser la sustitución de la inteligencia humana por la artificial, sino la amplificación de las capacidades humanas mediante la colaboración con sistemas de IA. En la misma línea, Selwyn (2019) advierte que muchas tecnologías educativas promueven una lógica centrada en eficiencia, velocidad y productividad inmediata, desplazando procesos de aprendizaje más profundos y reflexivos. En contextos universitarios, esta

situación puede manifestarse cuando estudiantes recurren sistemáticamente a herramientas de IA para producir contenido académico sin involucrarse activamente en procesos de comprensión conceptual. La automatización cognitiva presenta además una dimensión psicológica relevante: los individuos tienden a confiar crecientemente en respuestas generadas por sistemas automatizados, particularmente cuando dichas respuestas presentan apariencia de autoridad y sofisticación técnica, lo que reduce la tendencia a cuestionar críticamente la información recibida.

En el contexto universitario concreto, la delegación cognitiva mediada por IA generativa se manifiesta cuando estudiantes utilizan LLMs para la resolución de problemas sin haber intentado comprenderlos de manera autónoma, o recurren a sistemas de IA para generar la estructura argumental de un ensayo sin haber realizado el proceso reflexivo que esa construcción debería implicar. En estos casos, la tarea académica se completa de manera eficiente desde el punto de vista operativo, pero el proceso cognitivo que la tarea estaba diseñada para activar no ocurre. Es precisamente en esta brecha donde se acumula la deuda cognitiva, el concepto que se desarrolla en detalle en la sección siguiente. No obstante, este fenómeno no se circunscribe al ámbito académico. Como se señaló en capítulos anteriores, la ejecución de tareas sin análisis profundo en favor de la automatización también caracteriza los entornos laborales. La eficiencia operativa se obtiene, pero el costo cognitivo lo asumen los usuarios, no las máquinas.

Frente a este panorama, los organismos internacionales han comenzado a trazar marcos normativos y pedagógicos para orientar el uso responsable de la IA. La UNESCO (2021) adviertió que la incorporación de IA en educación debe orientarse a fortalecer capacidades humanas y no a reemplazarlas. El organismo enfatiza la necesidad de desarrollar alfabetización digital crítica y promover modelos pedagógicos centrados en pensamiento reflexivo. En consecuencia, la discusión sobre IA generativa en educación superior trasciende cuestiones puramente tecnológicas y se vincula directamente con interrogantes sobre aprendizaje, autonomía intelectual y desarrollo cognitivo.

5.3 Deuda cognitiva: definición y fundamentos

El concepto de deuda cognitiva surge recientemente dentro del debate académico sobre Inteligencia Artificial y automatización del pensamiento humano. La noción se inspira parcialmente en conceptos provenientes de otros campos, como la “deuda técnica” en ingeniería de software, donde decisiones orientadas a obtener beneficios inmediatos generan costos acumulativos futuros. En términos cognitivos, la deuda se refiere a la acumulación progresiva de efectos negativos derivados de la delegación excesiva de procesos intelectuales hacia sistemas tecnológicos externos.

Uno de los trabajos más relevantes sobre esta temática es el estudio desarrollado por Kosmyna et al. (2025), quienes analizan experimentalmente el impacto del uso de ChatGPT sobre procesos de escritura académica. Los autores identifican que la utilización intensiva de asistentes de IA puede reducir significativamente el esfuerzo mental involucrado en tareas complejas. Si bien esto produce mejoras inmediatas en términos de velocidad y productividad, también genera una menor activación de funciones cognitivas vinculadas a memoria de trabajo, razonamiento y planificación. El estudio sostiene que la deuda cognitiva no debe interpretarse simplemente como una reducción momentánea de carga mental, sino como una acumulación de efectos residuales asociados a la externalización sistemática del pensamiento. Cuando los individuos se acostumbran a depender constantemente de sistemas automatizados, experimentan posteriormente dificultades para desempeñarse sin asistencia tecnológica.

Para comprender los fundamentos de este fenómeno es indispensable recurrir a tres marcos teóricos clásicos de la psicología cognitiva que le otorgan sustento científico: la Teoría de la Carga Cognitiva de Sweller, el modelo de pensamiento dual de Kahneman, y la Teoría de la Metacognición de Flavell.

La Teoría de la Carga Cognitiva, desarrollada por Sweller (1994), sostiene que el sistema cognitivo humano opera con recursos de memoria de trabajo limitados. El aprendizaje complejo, aquel que conduce a la formación de esquemas cognitivos duraderos, requiere que estos recursos sean movilizados activamente en el

procesamiento de información nueva. La teoría distingue entre tres tipos de carga cognitiva: la intrínseca, determinada por la complejidad inherente del material; la extrínseca, generada por la forma en que se presenta el material; y la relevante, asociada al esfuerzo cognitivo que contribuye directamente a la formación de nuevos esquemas de conocimiento. Desde esta perspectiva, la deuda cognitiva puede entenderse como el resultado de una reducción sistemática de la carga cognitiva relevante: cuando la IA realiza el trabajo intelectual que debería activar el procesamiento cognitivo del estudiante, ese procesamiento no ocurre y, en consecuencia, los esquemas cognitivos que deberían formarse no se consolidan. Esta distinción es fundamental: mientras la carga cognitiva excesiva es un problema de sobrecarga puntual, la deuda cognitiva es un problema de subejercicio sistemático y prolongado.

A la Teoría de la Carga Cognitiva de Sweller se suma un segundo marco igualmente esclarecedor. Kahneman (2011), en su obra *Thinking, Fast and Slow*, propone una distinción entre dos sistemas de pensamiento que es especialmente relevante para comprender la dinámica de la delegación cognitiva descrita en la sección previa.

El Sistema 1 corresponde al pensamiento rápido, automático e intuitivo. Opera sin esfuerzo consciente y produce respuestas casi inmediatas. Reconocer un rostro familiar, leer un cartel en la calle, responder cuánto es 2+2, detectar el tono de enojo en la voz de alguien o esquivar un objeto que cae son ejemplos típicos de su funcionamiento. Es el sistema que utilizamos por defecto en la mayor parte de nuestras actividades cotidianas. El Sistema 1 es eficiente y rápido, pero también es el responsable de la mayor parte de los sesgos cognitivos, actos fallidos y errores de juicio.

El Sistema 2, en cambio, corresponde al pensamiento lento, deliberado y analítico. Se activa cuando enfrentamos problemas que el Sistema 1 no puede resolver de manera automática. Resolver una multiplicación compleja, comparar dos opciones financieras evaluando tasas e intereses, redactar un párrafo académico bien estructurado, evaluar críticamente un argumento, planificar una estrategia o aprender una habilidad nueva son ejemplos claros de su funcionamiento.

El Sistema 2 demanda concentración, consume energía mental significativa y opera con recursos cognitivos limitados. Es el sistema responsable del autocontrol, la atención sostenida, el razonamiento riguroso y la toma de decisiones deliberadas. Cuando alguien describe una tarea como “mentalmente agotadora” o “que requiere concentración”, se está refiriendo al uso intensivo del Sistema 2.

La diferencia central para comprender la deuda cognitiva radica en que la educación universitaria depende fundamentalmente del Sistema 2. El pensamiento crítico, la argumentación rigurosa, el análisis conceptual, la elaboración original de ideas y la construcción de conocimiento nuevo son operaciones que exigen el funcionamiento sostenido del Sistema 2. Kahneman observa, sin embargo, que los seres humanos tendemos por naturaleza a evitar el esfuerzo del Sistema 2 siempre que sea posible, una característica que el autor denomina ley del mínimo esfuerzo: *“A general ‘law of least effort’ applies to cognitive as well as physical exertion. The law asserts that if there are several ways of achieving the same goal, people will eventually gravitate to the least demanding course of action”* (Kahneman, 2011). Dicho de manera más simple, ante varias formas de alcanzar un mismo objetivo, las personas tienden a elegir la que exige menos esfuerzo mental. Cuando existe una alternativa más fácil que ofrezca un resultado aceptable, el cerebro tiende a optar por ella.

La IA generativa opera, desde el punto de vista del usuario, como una externalización radical del Sistema 2: transfiere la deliberación lenta y esforzada a un sistema automatizado y devuelve un resultado que puede ser incorporado sin activar los procesos analíticos que ese resultado debería haber requerido. A largo plazo, el subuso sistemático del Sistema 2 puede conducir a lo que Kahneman describe como pereza cognitiva: una disposición creciente a evitar el esfuerzo mental deliberado en favor de respuestas rápidas e irreflexivas. Esta pereza, lejos de ser un rasgo de personalidad, es una consecuencia previsible de no ejercitar las funciones del Sistema 2 durante períodos prolongados.

La Teoría de la Metacognición, formulada por Flavell (1979), introduce una dimensión adicional de análisis fundamental. Flavell definió la metacognición como el conocimiento y el control que los individuos ejercen sobre sus propios procesos

cognitivos: la capacidad de monitorear si se está comprendiendo genuinamente lo que se lee, de evaluar si la estrategia de resolución de un problema es la más adecuada, de reconocer cuándo es necesario revisar una respuesta. La metacognición es, en este sentido, el fundamento de la autonomía intelectual. Kosmyna et al. (2025) observaron en sus estudios que los participantes que usaban IA intensivamente mostraban no solo menor rendimiento en tareas cognitivas autónomas posteriores, sino también menor capacidad de monitorear sus propias limitaciones: tendían a sobreestimar la calidad de sus producciones y a subestimar la medida en que dependían de la asistencia tecnológica. Este hallazgo sugiere que la deuda cognitiva afecta no solo las capacidades cognitivas de primer orden, sino también las metacognitivas, comprometiendo la autorregulación del aprendizaje.

Desde una perspectiva neurocognitiva, Kosmyna et al. (2025) identificaron que los participantes que completaban tareas de escritura académica con asistencia de IA mostraban significativamente menor activación en áreas cerebrales vinculadas a funciones ejecutivas (planificación, memoria de trabajo, control inhibitorio) en comparación con quienes realizaban las mismas tareas de manera autónoma. Más relevante aún: cuando los participantes del grupo de IA debían posteriormente realizar tareas similares sin asistencia, no solo mostraban menor rendimiento, sino también mayores niveles de frustración y una menor sensación de agencia intelectual. Lo que importa no son las capacidades observables del producto final, sino los procesos cognitivos subyacentes que el ser humano deja de ejercitar, y son precisamente esos procesos los que el estudio de Kosmyna et al. muestra disminuidos a nivel cerebral. Los propios autores resumen el riesgo con claridad: *“While LLMs offer immediate convenience, our findings highlight potential cognitive costs. Over four months, LLM users consistently underperformed at neural, linguistic, and behavioral levels”* (Kosmyna et al., 2025). Traducido a términos concretos, aunque la IA aporta una comodidad inmediata, tiene un costo cognitivo: a lo largo de cuatro meses, quienes usaron el modelo de lenguaje rindieron peor en lo neuronal, lo lingüístico y lo conductual.

En otro punto, es interesante analizar la perspectiva de Vygotsky (1978) sobre la Zona de Desarrollo Próximo (ZDP) la cual ofrece un marco adicional para comprender el problema educativo que la deuda cognitiva plantea. Vygotsky planteó que el aprendizaje más significativo ocurre en la zona liminar entre lo que el individuo puede hacer de manera autónoma y lo que puede hacer con asistencia. La función del andamiaje sea proporcionado por un docente, un par o una herramienta, es ofrecer el apoyo mínimo necesario para que el estudiante realice la tarea, de modo que el esfuerzo y el aprendizaje siguen siendo del estudiante.

La IA generativa representa un andamiaje radicalmente diferente: en lugar de ofrecer el apoyo mínimo necesario, realiza la tarea completamente. Esto colapsa la ZDP: ya no hay zona de desarrollo, solo zona de consumo de resultados. El estudiante permanece estancado en su nivel actual de desarrollo autónomo, sin el beneficio del esfuerzo que el aprendizaje requiere.

A modo de síntesis de los marcos hasta aquí presentados, entre los principales síntomas asociados a deuda cognitiva identificados por la literatura reciente se encuentran: reducción del esfuerzo mental, dependencia creciente de herramientas de IA, disminución de autonomía intelectual, menor disposición a realizar procesos de análisis profundo, dificultades para resolver tareas sin asistencia tecnológica, reducción de capacidades metacognitivas y tendencia al aprendizaje superficial.

Lubbe et al. (2025) sostienen que el uso intensivo de IA puede afectar particularmente las habilidades de pensamiento de orden superior, como el análisis crítico, la creatividad y la elaboración conceptual. En este sentido, la deuda cognitiva no implica únicamente una pérdida de habilidades operativas, sino también una potencial transformación en la manera en que los individuos interactúan con el conocimiento. Estos síntomas, que aquí se enuncian de manera general, se examinan como riesgos concretos del ámbito universitario en la sección siguiente, y son los que más adelante el instrumento propuesto en la metodología buscará operacionalizar y medir.

5.4 Riesgos cognitivos en la educación superior

Los marcos teóricos desarrollados en la sección anterior, la carga cognitiva de Sweller, los sistemas de pensamiento de Kahneman, la metacognición de Flavell y la Zona de Desarrollo Próximo de Vygotsky, permiten ahora identificar con precisión los riesgos concretos que la incorporación masiva de Inteligencia Artificial generativa introduce en las universidades. Estos desafíos se vinculan al aprendizaje, la evaluación y el desarrollo de competencias cognitivas. Aunque estas tecnologías ofrecen beneficios significativos en términos de acceso a información y eficiencia académica, tal como se reconoció con Holmes, Bialik y Fadel en secciones anteriores, diversos estudios advierten sobre riesgos asociados al uso excesivo y acrítico de herramientas automatizadas.

Uno de los principales riesgos identificados por la literatura es la promoción de aprendizajes superficiales. El aprendizaje profundo implica procesos de comprensión conceptual, análisis crítico y elaboración reflexiva de información. Sin embargo, cuando los estudiantes utilizan IA principalmente para obtener respuestas inmediatas o resolver tareas automáticamente, puede reducirse su involucramiento activo en dichos procesos. Como ya identificaron Fan, Y., et al. (2025), citados previamente, muchos estudiantes perciben mejoras significativas en productividad gracias al uso de IA, aunque dichas mejoras no siempre se reflejan en una mayor comprensión de contenidos: emerge así una diferencia entre eficiencia operativa y aprendizaje real que constituye uno de los rasgos más característicos de la deuda cognitiva. Los autores lo advierten de manera directa: *“AI technologies such as ChatGPT may promote learners' dependence on technology and potentially trigger metacognitive ‘laziness’, which can potentially hinder their ability to self-regulate and engage deeply in learning”* (Fan, Y., et al., 2025). En síntesis, herramientas como ChatGPT pueden fomentar la dependencia tecnológica y una “pereza metacognitiva” que le dificulta al estudiante autorregular su aprendizaje e involucrarse en profundidad.

Otro riesgo relevante se vincula con la disminución del pensamiento crítico. El pensamiento crítico constituye una competencia central en educación superior y se relaciona con la capacidad de evaluar información, identificar inconsistencias,

construir argumentos y cuestionar supuestos. Las respuestas generadas por sistemas de IA suelen presentarse de manera coherente y convincente, lo cual puede disminuir la tendencia de los estudiantes a cuestionar o verificar críticamente la información recibida, un mecanismo que Selwyn (2019) ya anticipó al describir la confianza creciente en respuestas de apariencia autorizada. Fan, Y., et al. (2025) muestran que la dependencia excesiva de herramientas de IA puede afectar habilidades de razonamiento de orden superior, incluyendo la creatividad, el análisis conceptual y la resolución autónoma de problemas.

Esta observación se articula con la conocida Taxonomía de Bloom (Bloom et al., 1956), marco de referencia clásico y muy utilizado para categorizar niveles de aprendizaje.

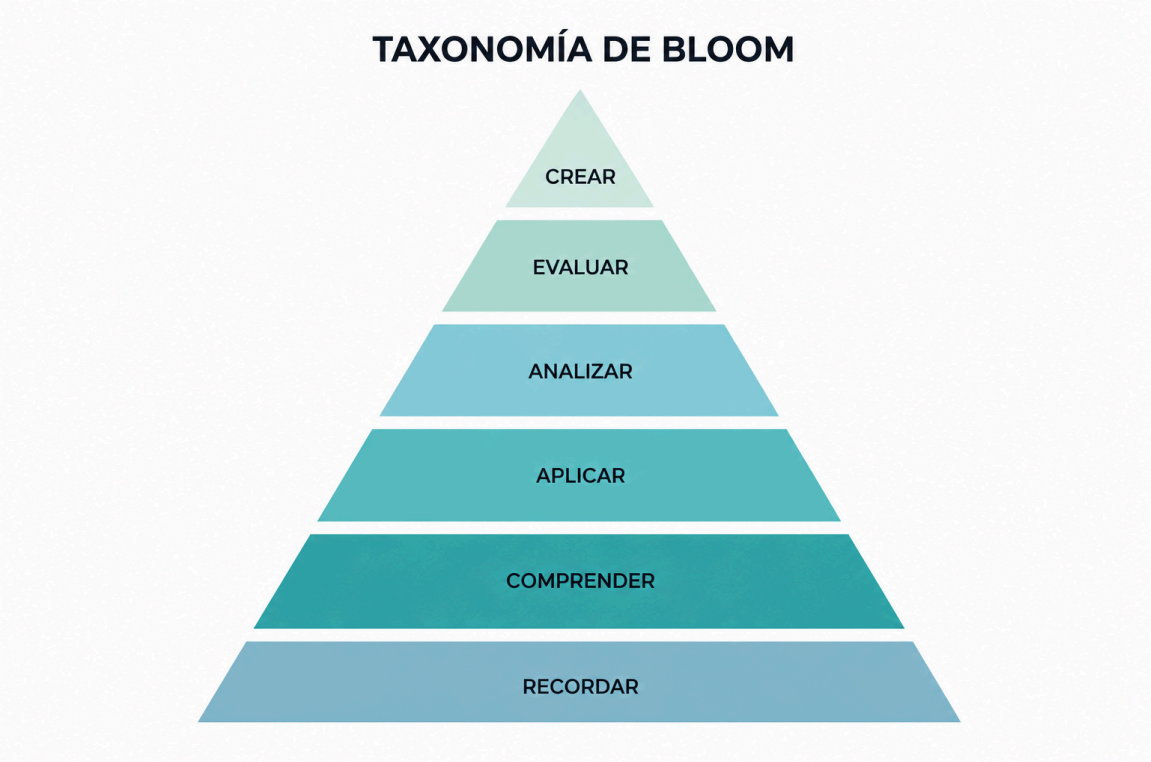


Figura 1. *Los 6 niveles de la Taxonomía de Bloom*

Los 3 niveles superiores de la taxonomía, análisis, evaluación y creación, son precisamente los que la educación universitaria busca desarrollar y los que el uso intensivo de IA amenaza con cortocircuitar. Un estudiante puede producir un trabajo que aparenta alcanzar el nivel de evaluación o síntesis (los niveles más

complejos) sin haber realizado los procesos cognitivos que esos niveles demandan, dado que fue la IA quien realizó ese trabajo.

Estos riesgos de sobredelegación se ven agravados por una paradoja que la literatura técnica reciente comienza a documentar: los propios modelos de lenguaje presentan limitaciones significativas que los usuarios habitualmente no perciben y que ponen en cuestión la confiabilidad del output que consumen. Leviathan, Kalman y Matias (2025) demostraron que algo tan elemental como repetir el *prompt* de entrada mejora significativamente el rendimiento de los principales LLMs en tareas que no requieren razonamiento explícito, lo que revela que la calidad de las respuestas depende de variables arquitectónicas completamente ajenas a la comprensión del usuario promedio. Más significativo aún para el contexto educativo, Laban, Hayashi, Zhou y Neville (2025), en un estudio que recibió el *Outstanding Paper Award* en ICLR 2026, documentaron que todos los principales LLMs exhiben un rendimiento significativamente inferior en conversaciones de múltiples turnos respecto de consultas de turno único, con una caída promedio del 39% en seis tareas generativas. Los autores descubrieron que los modelos tienden a hacer suposiciones prematuras, generar respuestas excesivamente verbosas y, cuando toman un camino equivocado en la conversación, no se recuperan, salvo que el usuario se de cuenta y pida activamente corregirlo.

Dado que el uso estudiantil de IA se produce predominantemente en modo conversacional, iterativo y con instrucciones frecuentemente subespecificadas (tal como se puede constatar en cualquier aula universitaria contemporánea), este hallazgo sugiere que los outputs que los estudiantes efectivamente consumen son de calidad inferior a la que el modelo es capaz de producir en condiciones óptimas. Esta brecha entre la calidad percibida y la calidad real de la interacción, que se suma a la brecha entre eficiencia percibida y aprendizaje real documentada por Fan, Deng y Liu (2025), amplifica el riesgo de deuda cognitiva descrito a lo largo de este trabajo: los estudiantes no solo están delegando su razonamiento, sino que lo están delegando a un sistema que, en el modo de uso más frecuente, rinde por debajo de sus propias capacidades.

La reducción del esfuerzo mental sostenido representa otro aspecto problemático, directamente vinculado a la noción de carga cognitiva relevante de Sweller desarrollada anteriormente. El aprendizaje universitario requiere frecuentemente atravesar procesos complejos de lectura, análisis, escritura y revisión que demandan esfuerzo cognitivo significativo pero son fundamentales para consolidar conocimientos y desarrollar habilidades intelectuales. La automatización excesiva puede reducir la exposición de los usuarios (en este caso estudiantes) a este tipo de procesos. Es que en los estudios ya citados, se han identificado numerosas situaciones donde los usuarios que delegan constantemente tareas complejas en IA experimentan posteriormente mayores dificultades para resolver problemas sin asistencia tecnológica, lo que sugiere que la reducción prolongada del esfuerzo mental podría afectar capacidades asociadas a autonomía intelectual y resolución independiente de tareas.

La utilización intensiva de IA puede también impactar procesos metacognitivos, retomando la dimensión que Flavell (1979) aportó al marco teórico. La metacognición implica que los estudiantes sean capaces de monitorear su propio aprendizaje, identificar errores y evaluar críticamente su desempeño. Cuando gran parte del trabajo intelectual es realizado por sistemas automatizados, puede disminuir la participación activa del estudiante en dichos procesos de autorregulación cognitiva, un efecto que, como se señaló en la sección 5.3, el propio estudio de Kosmyna et al. documentó al observar que los usuarios intensivos de IA tendían a sobreestimar la calidad de sus producciones.

Otro riesgo relevante se relaciona con la homogeneización del pensamiento. Las herramientas de IA generativa producen respuestas basadas en patrones estadísticos derivados de grandes volúmenes de datos, lo que favorece salidas estructuralmente similares y limita los procesos de producción original y pensamiento divergente. En otros términos, estos sistemas no generan conocimiento genuinamente nuevo, sino que recombinan y reorganizan contenido preexistente.

También emergen desafíos vinculados a la evaluación educativa. La utilización masiva de IA dificulta determinar cuánto del trabajo académico corresponde

efectivamente al estudiante y cuánto fue producido mediante automatización. Esto genera tensiones relacionadas con la autoría, integridad académica y legitimidad de procesos evaluativos. La UNESCO (2023), cuya advertencia ya se mencionó, sostiene que muchas instituciones educativas todavía no cuentan con lineamientos claros para integrar IA de manera pedagógicamente responsable, lo que puede generar escenarios donde los beneficios operativos superen temporalmente la capacidad institucional para gestionar sus riesgos cognitivos y educativos.

Finalmente, la automatización cognitiva también puede afectar la motivación y la percepción del aprendizaje. Cuando las tareas académicas son resueltas rápidamente mediante IA, algunos estudiantes pueden percibir menor necesidad de involucrarse activamente en procesos de estudio profundo, generando una visión instrumental del aprendizaje centrada exclusivamente en resultados inmediatos y no en construcción significativa de conocimiento, en otras palabras, desinteres. En conjunto, estos riesgos no son independientes entre sí, sino manifestaciones interrelacionadas del mismo fenómeno de fondo, la deuda cognitiva, y son los que fundamentan las dimensiones del instrumento de medición que se presenta en el capítulo de metodología.

5.5 Competencias cognitivas en la era digital

Si la sección anterior identificó los riesgos cognitivos que la IA introduce en la educación superior, corresponde ahora preguntarse por su contracara: qué competencias deben preservarse y fortalecerse para contrarrestar esos riesgos. La transformación tecnológica contemporánea ha modificado profundamente las competencias requeridas para desenvolverse en entornos educativos, profesionales y sociales. En la era digital, las capacidades cognitivas más valoradas ya no se limitan únicamente a la memorización o reproducción de información, sino que incluyen pensamiento crítico, resolución compleja de problemas, creatividad, adaptabilidad y alfabetización digital.

La expansión de la Inteligencia Artificial generativa intensifica esta transformación al automatizar numerosas tareas tradicionalmente asociadas al trabajo intelectual

humano. En este contexto, diversos autores sostienen que la educación superior debe reconfigurar sus objetivos formativos para priorizar competencias que continúen siendo distintivamente humanas. Retomando una vez más a Luckin (2018), ella plantea que el verdadero desafío educativo de la era de la IA consiste en desarrollar capacidades complementarias a la automatización tecnológica: formar estudiantes capaces de interpretar críticamente información, contextualizar respuestas automatizadas y tomar decisiones reflexivas.

El pensamiento crítico aparece como una de las competencias centrales dentro de este escenario, y su erosión fue señalada como uno de los principales riesgos del uso intensivo de IA. En contextos de abundancia informacional y automatización cognitiva, esta competencia adquiere relevancia aún mayor. Las herramientas de IA pueden generar respuestas técnicamente sofisticadas, pero no necesariamente correctas o contextualizadas, un punto que conecta con el sesgo de autopreferencia documentado por Xu et al. (2025): los outputs de la IA pueden resultar convincentes sin ser por ello confiables. Por este motivo, los estudiantes requieren habilidades para verificar finamente, cuestionar y complementar información producida por sistemas automatizados. Esta capacidad constituye, como señala Luckin (2018), una metacapacidad: la habilidad de comprender cuándo y cómo confiar en los outputs de la IA, y cuándo someterlos a escrutinio propio.

Otra competencia fundamental es la autonomía intelectual. La literatura sobre deuda cognitiva advierte que el uso intensivo de IA podría afectar progresivamente esta capacidad. Cuando los estudiantes delegan sistemáticamente tareas complejas en herramientas automatizadas, existe el riesgo de reducir la práctica activa de procesos intelectuales fundamentales. La autonomía intelectual no es simplemente la capacidad de prescindir de herramientas, ya que nadie razona en el vacío absoluto, sino la confianza y la disposición de involucrarse activamente con los problemas cognitivos en lugar de eludirlos mediante automatización. Esta definición es la que el instrumento de medición del capítulo siguiente operacionaliza como una de sus dimensiones centrales.

La metacognición también adquiere importancia central en la era digital. Los estudiantes necesitan desarrollar capacidad para monitorear críticamente sus propios procesos de aprendizaje y reconocer cuándo el uso de tecnología contribuye efectivamente al desarrollo cognitivo y cuándo puede transformarse en dependencia excesiva. Flavell (1979) anticipó que la metacognición es fundamental para el aprendizaje autónomo a lo largo de la vida; en el contexto actual, esta capacidad incluye necesariamente una metacognición sobre la propia relación con la tecnología.

Por otra parte resulta relevante el concepto de alfabetización digital crítica. La alfabetización digital tradicional se centraba principalmente en habilidades operativas vinculadas al uso de tecnologías. Sin embargo, la expansión de la IA exige enfoques más complejos orientados a comprender cómo funcionan los sistemas automatizados, cuáles son sus limitaciones y qué sesgos pueden presentar. La UNESCO (2021) sostiene que la educación contemporánea debe promover competencias orientadas a la utilización ética, crítica y responsable de la Inteligencia Artificial, formando estudiantes capaces de interactuar con tecnologías avanzadas sin perder capacidad de razonamiento independiente. Esta exigencia se vuelve más urgente a la luz de hallazgos técnicos como los de Leviathan et al. (2025), que demuestran que algo tan simple como la estructura del *prompt* afecta significativamente la calidad de las respuestas, y los de Laban et al. (2025), que documentan cómo el modo conversacional, lejos de ser un entorno de refinamiento progresivo, degrada la precisión del modelo. Una alfabetización digital que no incorpore este tipo de comprensión deja al estudiante sin herramientas para evaluar críticamente lo que la IA produce, lo que refuerza el ciclo descrito por Bloom, siendo *outputs* que aparentan sofisticación sin que el proceso cognitivo los respalde.

Estas consideraciones conducen a un principio que la literatura sobre diseño de sistemas de inteligencia artificial ha formalizado bajo el nombre de “*human in the loop”*, es decir, la idea de que el ser humano debe permanecer como participante activo dentro del proceso de decisión y no como mero receptor de sus resultados. El concepto, sistematizado por Monarch (2021) en el campo del aprendizaje

automático pero con raíces previas en la teoría de control y la ingeniería de factores humanos, sostiene que los sistemas automatizados deben incorporar puntos de intervención humana, especialmente en decisiones de alto impacto o difícil reversión. Es significativo que, según documenta la propia literatura del campo, la noción de *human in the loop* en el aprendizaje automático se inspiró en técnicas pedagógicas cuyo propósito era convertir al estudiante en un socio activo del proceso de aprendizaje para evitar que se volviera excesivamente dependiente. Trasladado al problema que ocupa a esta tesis, el principio de *human in the loop* ofrece un horizonte normativo complementario al de la inteligencia aumentada de Luckin (2018): el estudiante no debería situarse al final de la cadena, recibiendo pasivamente el producto de la IA, sino permanecer dentro del circuito de producción del conocimiento, interrogando, corrigiendo y re-elaborando aquello que la herramienta genera. Como se demostrará empíricamente en el capítulo de resultados, esta permanencia activa del estudiante en el proceso, operacionalizada en la dimensión de conducta de revisión, constituye el principal factor protector frente a la acumulación de deuda cognitiva.

Otra dimensión relevante se relaciona con creatividad y pensamiento divergente. Aunque los sistemas de IA pueden generar contenido sofisticado, su funcionamiento se basa en patrones derivados de datos previamente existentes, lo que conecta con el riesgo de homogeneización del pensamiento mencionado anteriormente. Por este motivo, la creatividad humana continúa siendo una capacidad distintiva fundamental. Lubbe et al. (2025) advierten que el uso excesivo de IA podría afectar los procesos de elaboración original cuando los estudiantes se acostumbran a depender constantemente de respuestas automatizadas.

En consecuencia, la educación superior enfrenta el desafío de integrar IA sin debilitar procesos creativos y reflexivos. Las universidades deben asumir un rol activo en la construcción de modelos pedagógicos orientados a equilibrar innovación tecnológica y preservación de capacidades cognitivas fundamentales. La incorporación de IA no debería orientarse solo a la eficiencia operativa; también

tiene que fortalecer competencias humanas difíciles de automatizar. Valverde-Berrocoso et al. (2022) subrayan que los efectos de las tecnologías educativas sobre el aprendizaje dependen críticamente del diseño instruccional que las acompaña: una misma herramienta puede tener efectos positivos o negativos sobre las competencias cognitivas dependiendo de cómo se integra en los procesos de enseñanza y evaluación.

Finalmente, la discusión sobre competencias cognitivas en la era digital remite a un interrogante central que recorre todo este marco teórico: cómo utilizar tecnologías avanzadas sin comprometer los procesos intelectuales que constituyen la base del aprendizaje universitario. Responder esta pregunta requiere desarrollar estrategias pedagógicas, instrumentos de evaluación y marcos conceptuales capaces de comprender los efectos cognitivos de la automatización tecnológica. En este sentido, el concepto de deuda cognitiva representa una herramienta teórica relevante para analizar críticamente la relación entre Inteligencia Artificial, aprendizaje y autonomía intelectual en la educación superior contemporánea. Es precisamente este concepto el que el capítulo de metodología se propone transformar en un fenómeno medible, traduciendo los marcos teóricos aquí expuestos en dimensiones operativas concretas.

6. Metodología

La presente investigación se estructura como un estudio descriptivo-exploratorio con un diseño transversal y de carácter no experimental. El enfoque adoptado es predominantemente cuantitativo, dado que el objetivo central consiste en operacionalizar el concepto de deuda cognitiva mediante un instrumento estructurado de medición que pueda ser aplicado a estudiantes universitarios de grado en un contexto educativo real. La elección de un diseño transversal responde a la naturaleza exploratoria del estudio: se busca obtener una primera caracterización del fenómeno en la población de interés, sin pretender en esta etapa establecer relaciones causales ni realizar seguimientos longitudinales.

El estudio reconoce desde su planteamiento metodológico que el tamaño muestral previsto es acotado. Se estima alcanzar al menos cien respuestas válidas, obtenidas mediante muestreo no probabilístico por conveniencia a partir de estudiantes de universidades públicas y privadas de Rosario y Buenos Aires. Esta limitación muestral, lejos de invalidar el estudio, define con claridad su alcance: el trabajo se propone como una prueba de concepto y una primera aproximación al diseño y validación preliminar del instrumento, dejando sentadas las bases para investigaciones posteriores que puedan ampliar la muestra y profundizar los análisis estadísticos.

6.1 Diseño del cuestionario

El instrumento propuesto consiste en un cuestionario estructurado autoadministrado, diseñado para ser distribuido a través de medios digitales (Google Forms). La modalidad online responde a varias consideraciones prácticas: facilita la distribución entre estudiantes universitarios, permite la recolección automatizada de respuestas en formato tabular apto para análisis posterior, reduce costos operativos y elimina la necesidad de transcripción manual de datos.

El cuestionario se presenta como un instrumento unificado sin división en secciones diferenciadas. El instrumento comienza con una pregunta inicial consultando por universidad a la cual el alumno pertenece y continúa de manera directa con las preguntas de medición de la deuda cognitiva, agrupadas en dos subconjuntos diferenciados por su formato de respuesta. El primero comprende dos preguntas de frecuencia con escala ordinal de cinco categorías: la primera sobre el uso académico de herramientas de IA y la segunda sobre el uso fuera del ámbito universitario (trabajo, vida personal, entretenimiento), en un rango que va desde “Nunca” hasta “Varias veces al día”. El segundo subconjunto comprende once preguntas con escala tipo Likert de cinco puntos, donde el participante manifiesta su nivel de acuerdo con afirmaciones específicas (1 = Totalmente en desacuerdo, 2 = En desacuerdo, 3 = Ni de acuerdo ni en desacuerdo, 4 = De acuerdo, 5 = Totalmente de acuerdo).

La elección de la escala Likert de cinco puntos se fundamenta en su amplia validación en investigación social y psicométrica, su facilidad de comprensión para los participantes y su compatibilidad con análisis cuantitativos posteriores. Las preguntas fueron formuladas en lenguaje claro y directo, evitando tecnicismos y construcciones ambiguas, con el objetivo de minimizar la cantidad de errores y facilitar la respuesta de estudiantes que pueden no estar familiarizados con conceptos especializados de psicología cognitiva o tecnología.

Algunas de las preguntas están redactadas en sentido inverso al constructo medido. Es decir, en estas preguntas un puntaje alto indica baja deuda cognitiva (por ejemplo, “Releo críticamente lo que la IA me genera buscando errores antes de usarlo”).

Esta estrategia, denominada de ítems invertidos, busca reducir el sesgo de aquiescencia, un sesgo común en cuestionarios donde los participantes tienden a responder afirmativamente sin atender al contenido de la afirmación. Para el cálculo del índice final, las respuestas a estos ítems se invierten matemáticamente, de modo que en la base de datos final todas las variables apuntan en la misma dirección.

Se prevé que el cuestionario pueda ser completado en un tiempo aproximado de cinco a siete minutos. Este parámetro es importante porque la extensión de un cuestionario afecta directamente la tasa de respuesta y la calidad de las respuestas: cuestionarios excesivamente largos generan abandono o respuestas apresuradas.

Antes de la distribución definitiva, se contempla la realización de una prueba piloto con un grupo reducido de estudiantes (entre cinco y diez participantes), con el objetivo de detectar ambigüedades en la redacción, ítems confusos, problemas de comprensión o errores técnicos en el formulario.

En términos éticos, la participación será voluntaria y anónima. El cuestionario incluirá un consentimiento informado al inicio, en el cual se explicará el propósito del estudio, el uso previsto de los datos, el carácter anónimo de las respuestas y el derecho del participante a interrumpir su participación en cualquier momento.

6.2 Dimensiones propuestas

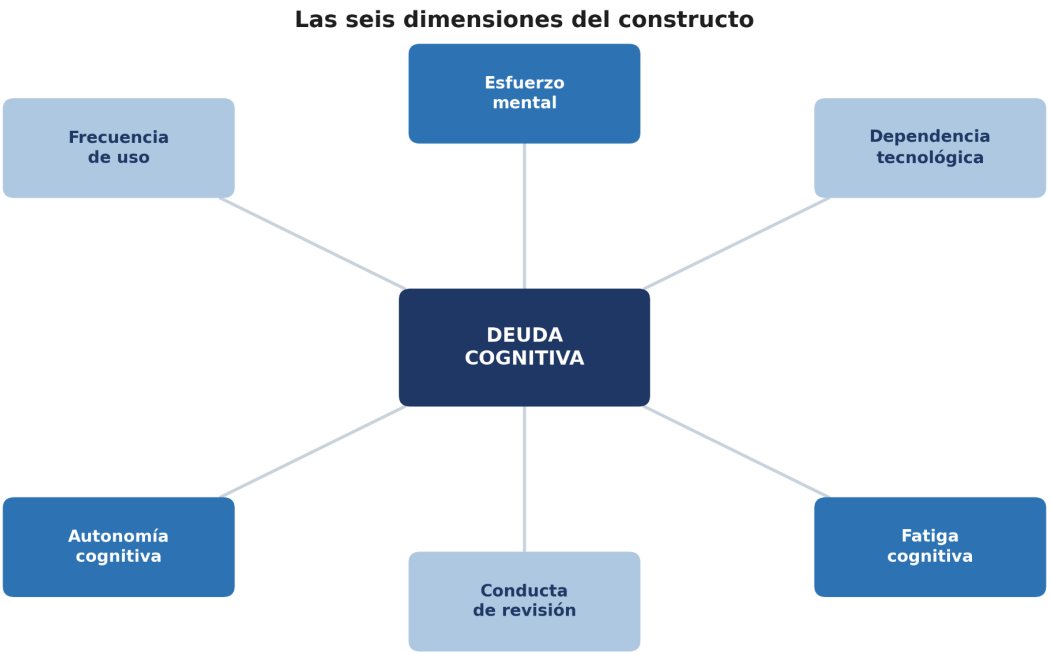


Figura 2. Las seis dimensiones que componen el constructo de deuda cognitiva.

La operacionalización de la deuda cognitiva se realiza mediante seis dimensiones derivadas directamente de la revisión teórica desarrollada en el capítulo anterior. No se trata de categorías elegidas de manera arbitraria: cada una traduce a un plano medible alguno de los marcos conceptuales o riesgos discutidos en las secciones del marco teórico. Cada dimensión captura un aspecto específico del fenómeno y se mide a través de uno o dos ítems del cuestionario. La combinación de las seis dimensiones permite construir una visión multidimensional del constructo, en lugar de reducirlo a un único indicador.

Dimensión 1: Frecuencia de uso

Esta dimensión mide la intensidad con la que el estudiante recurre a herramientas de IA generativa tanto en el ámbito académico como fuera de él. Se operacionaliza mediante dos preguntas de frecuencia con escala ordinal: la primera sobre uso académico y la segunda sobre uso cotidiano general. La distinción entre ambos contextos permite identificar si el patrón de dependencia trasciende el entorno

universitario. La frecuencia de uso no es por sí sola un indicador de deuda cognitiva (está claro que es posible usar IA frecuentemente de manera reflexiva y crítica), pero opera como factor de riesgo: a mayor frecuencia de uso, mayor exposición a los efectos cognitivos acumulativos.

Dimensión 2: Esfuerzo mental

Esta dimensión captura la disposición del estudiante a involucrarse en tareas cognitivamente exigentes y su tolerancia al esfuerzo intelectual sostenido. Se relaciona directamente con la Teoría de la Carga Cognitiva de Sweller y con el concepto de Kahneman de Sistema 2: medir el esfuerzo mental es, en términos operativos, medir la disposición del estudiante a activar su Sistema 2 en lugar de delegarlo. Se mide a través de dos ítems: el primero evalúa la tendencia a consultar IA antes de intentar resolver autónomamente; el segundo, incorporado en la versión revisada del instrumento, evalúa si el estudiante utiliza IA como primera opción para resolver actividades académicas en lugar de buscadores tradicionales como Google. Esta sustitución refleja una forma de delegación del proceso de búsqueda e investigación activa, que históricamente requería del estudiante mayor criterio selectivo y esfuerzo cognitivo para evaluar fuentes.

Dimensión 3: Dependencia tecnológica

Esta dimensión evalúa el grado en que el estudiante percibe que su rendimiento académico depende de la disponibilidad de herramientas de IA. Refleja la pérdida de autonomía funcional y la sensación de bloqueo ante la ausencia de la herramienta, fenómenos que Kosmyna et al. (2025) documentaron experimentalmente y que se reseñaron en la sección 5.3. Se mide mediante tres ítems: el primero indaga sobre la percepción de caída en el rendimiento sin IA; el segundo, sobre la sensación de bloqueo cuando la herramienta no está disponible o no entrega respuestas útiles; el tercero, incorporado en la versión revisada del instrumento, evalúa si la ausencia de IA hace que las tareas se perciban como significativamente más demandantes en tiempo o dificultad. Este tercer ítem reemplaza una formulación anterior que contenía un anclaje temporal (“me canso más rápido que antes”) que podía sesgar la respuesta al presuponer un cambio previo, y lo sustituye por una medición de esfuerzo percibido más neutra y directa.

Dimensión 4: Autonomía cognitiva

Esta dimensión evalúa la confianza y disposición del estudiante para resolver tareas académicas complejas sin asistencia tecnológica. Es una dimensión clave porque la autonomía intelectual constituye uno de los objetivos formativos centrales de la educación superior y una de las competencias más amenazadas por la delegación cognitiva. Se mide mediante dos ítems, uno formulado en sentido directo (preferencia por pedir resúmenes en lugar de leer) y otro en sentido inverso (capacidad percibida de redactar trabajos extensos sin IA).

Dimensión 5: Conducta de revisión

Esta dimensión mide el grado en que el estudiante somete a escrutinio crítico la información generada por la IA antes de utilizarla. Captura la calidad de la interacción con la herramienta y opera como factor protector: estudiantes que revisan, modifican y complementan las respuestas de la IA tienen menor probabilidad de acumular deuda cognitiva. Esta dimensión es especialmente relevante a la luz de dos hallazgos discutidos en el marco teórico: por un lado, la Taxonomía de Bloom que permite identificar que sin revisión crítica el estudiante puede producir trabajos que aparentan niveles altos de pensamiento sin haberlos alcanzado; por otro, los hallazgos de Laban et al. (2025) y Leviathan et al. (2025) demuestran que la calidad de los outputs de IA es más variable e impredecible de lo que el usuario promedio percibe, lo que vuelve la revisión una necesidad cognitiva, más que una simple buena práctica.

Esta dimensión se vincula estrechamente con la conducta metacognitiva que Flavell (1979) describió y con la metacapacidad de discernir cuándo confiar en los outputs de la IA que Luckin (2018) postuló, ambas presentadas en el marco teórico. Se mide mediante tres ítems, todos formulados en sentido inverso: el primero evalúa si el estudiante relee críticamente el output de IA buscando errores conceptuales o inconsistencias; el segundo, si lo modifica, reescribe o complementa con ideas propias antes de entregarlo; el tercero, incorporado en la versión revisada del instrumento, indaga si el estudiante utiliza IA para generar un primer borrador sobre el cual trabaja posteriormente. Este último ítem captura una forma de uso que combina delegación inicial con involucramiento cognitivo

posterior: el componente activo “luego trabajo sobre él” lo distingue del uso pasivo de outputs de IA y lo caracteriza como factor protector parcial frente a la deuda cognitiva.

No obstante, cabe señalar que esta dimensión presenta una limitación relevante vinculada a la subjetividad en la interpretación del propio comportamiento de revisión. El umbral que cada estudiante utiliza para considerar que "revisó" o "modificó" un *output* de IA puede variar sustancialmente: para algunos, agregar una línea de texto, cambiar el formato o introducir un ajuste superficial puede percibirse como una revisión suficiente, mientras que para otros la revisión implica un proceso de análisis crítico, contraste con fuentes propias y reformulación sustancial del contenido.

Dado que los ítems de esta dimensión se basan en autorreporte, no es posible garantizar que todos los estudiantes estén respondiendo bajo un mismo criterio de lo que constituye "revisar" o "complementar con ideas propias". Esta heterogeneidad en la interpretación introduce un componente de subjetividad que limita la comparabilidad directa de las respuestas entre sujetos.

En consecuencia, se considera que esta dimensión constituye un punto de partida valioso pero perfectible, y que futuras investigaciones podrían profundizar en su operacionalización, por ejemplo, mediante observación directa de las interacciones con la IA, análisis de las versiones sucesivas de un trabajo, o escalas que distingan niveles de profundidad en la revisión, con el fin de captar de manera más precisa la naturaleza y el grado del involucramiento cognitivo posterior al uso de la herramienta.

Dimensión 6: Fatiga cognitiva

Esta dimensión evalúa el grado de agotamiento mental percibido tras el uso intensivo de IA y la dificultad para retomar tareas que exigen razonamiento propio. Se relaciona con la noción de pereza cognitiva propuesta por Kahneman y con los hallazgos neurocognitivos de Kosmyna et al. (2025) sobre la dificultad para retomar tareas autónomas después del uso de IA.

Tabla 1. Correspondencia entre las seis dimensiones del constructo, los ítems del cuestionario, su fundamento teórico y su codificación.

Dimensión	Ítem(s) del cuestionario	Fundamento teórico	Codificación
1.Frecuencia de uso	P2 (uso académico) y P3 (uso general)	Factor de riesgo: exposición acumulativa	Directa
2.Esfuerzo mental	P4 (primera reacción: consultar IA) y P5 (IA en vez de Google)	Carga cognitiva (Sweller); Sistema 2 (Kahneman)	Directa
3.Dependencia tecnológica	P6 (rendimiento bajaría), P7 (bloqueo sin IA) y P8 (más tiempo sin IA)	Dependencia funcional (Kosmyna et al., 2025)	Directa
4.Autonomía cognitiva	P9 (capaz de redactar sin IA) y P10 (prefiere resumen)	Autonomía intelectual; ZDP (Vygotsky)	P9 inversa; P10 directa
5.Conducta de revisión	P11 (relee crítico), P12 (modifica) y P13 (IA como borrador)	Metacognición (Flavell); human in the loop	Inversa
6.Fatiga cognitiva	P14 (cuesta reconcentrarse tras usar IA)	Pereza cognitiva (Kahneman); evidencia neuro (Kosmyna)	Directa

6.3 Construcción del índice de deuda cognitiva

La construcción del índice de deuda cognitiva sigue un procedimiento estandarizado que permite transformar las trece respuestas individuales del cuestionario en un puntaje único interpretable. El procedimiento se desarrolla en cuatro pasos.

En primer lugar, se realiza la codificación inversa de los ítems formulados en sentido contrario al constructo. Para cada respuesta a un ítem invertido, se aplica la transformación: valor nuevo = 6 menos valor original. Esto significa que una respuesta de 1 (totalmente en desacuerdo) se convierte en 5, una respuesta de 5 se convierte en 1, y los valores intermedios se invierten de manera

correspondiente. Tras esta transformación, todos los ítems apuntan en la misma dirección: a mayor puntaje, mayor deuda cognitiva.

En segundo lugar, se calcula el puntaje crudo del índice como la suma simple de los trece ítems. Dado que cada ítem puede tomar valores entre 1 y 5, el puntaje crudo se encuentra entre 13 (mínima deuda cognitiva posible) y 65 (máxima deuda cognitiva posible).

En tercer lugar, se realiza una estandarización del puntaje a una escala de 0 a 100, mediante la fórmula: índice estandarizado = (puntaje crudo menos 13) dividido 52, multiplicado por 100. Esta estandarización facilita la interpretación y la comparación con eventuales estudios posteriores que utilicen instrumentos similares.

En cuarto lugar, se realiza la clasificación del estudiante en una de cinco categorías de deuda cognitiva, según los puntos de corte preliminares siguientes: de 0 a 20 puntos, deuda cognitiva muy baja; de 21 a 40, baja; de 41 a 60, media; de 61 a 80, alta; de 81 a 100, muy alta. Es importante señalar que estos umbrales tienen carácter preliminar y orientativo.

6.4 Validación preliminar

En el marco de la presente investigación, dada la limitación muestral y el carácter exploratorio del trabajo, se contempla una validación preliminar centrada en la consistencia interna.

Esta se evaluará mediante el cálculo del coeficiente alfa de Cronbach, un coeficiente estadístico que estima el grado en que los ítems de una escala miden el mismo constructo subyacente. Se considera aceptable un valor de alfa igual o superior a 0,70, el valor convencional en investigación psicométrica. Valores inferiores podrían indicar problemas en la formulación de algún ítem o en la coherencia interna de las dimensiones. El cálculo del alfa de Cronbach se realizará tanto para el conjunto completo del instrumento como para cada dimensión por separado, lo que permitirá identificar dimensiones que requieran ajuste en futuras versiones.

Adicionalmente, se realizará un análisis descriptivo de las respuestas: distribución de frecuencias por ítem, media, desvío estándar, asimetría y curtosis. Estos análisis permiten detectar ítems problemáticos (por ejemplo, ítems con varianza extremadamente baja que no discriminan entre participantes, o ítems con distribuciones marcadamente sesgadas).

Es importante explicitar las limitaciones del proceso de validación previsto. Una validación psicométrica completa requeriría análisis factorial confirmatorio, validación de constructo mediante correlación con instrumentos relacionados, evaluación de fiabilidad test-retest y muestras significativamente más amplias. El alcance del presente trabajo, en cambio, se limita a establecer una primera base evidencial sobre la utilidad del instrumento, dejando explícitamente abierto el camino para investigaciones posteriores que profundicen este proceso.

6.5 Modelado exploratorio y automatización

Una vez recolectados los datos y construido el índice de deuda cognitiva, el trabajo contempla una etapa exploratoria de análisis cuantitativo. El propósito de esta etapa es evaluar la viabilidad de estimar el nivel de deuda cognitiva mediante modelos analíticos simples, utilizando como variables predictoras un conjunto de factores de contexto independientes del propio cuestionario: la frecuencia de uso de IA, el perfil de la carrera y el tipo de universidad. Esta etapa tiene carácter de prueba de concepto y no busca generar capacidades predictivas robustas en sentido estricto, sino explorar relaciones entre variables y sentar bases metodológicas para investigaciones futuras con muestras más amplias.

El modelo: regresión lineal múltiple

Antes de describir el modelo en detalle vale precisar qué se entiende por modelo en este contexto. Un modelo es una representación simplificada de la forma en que se relacionan distintos factores. En este caso, el objetivo es describir cómo el nivel de deuda cognitiva de un estudiante puede asociarse con un conjunto de características observables de ese estudiante. La idea de fondo es que el puntaje de cada persona no es completamente azaroso, sino que responde en parte a factores identificables: por ejemplo, cuánto utiliza la IA, qué tipo de carrera cursa y

en qué clase de institución estudia. Un modelo estadístico expresa esa intuición de manera ordenada y permite estimar cuánto aporta cada factor.

La regresión lineal es la forma más elemental de construir esa representación. El término lineal indica que el modelo asume una relación proporcional y constante entre cada predictor y el índice: a mayor frecuencia de uso, mayor deuda cognitiva esperada, con un incremento parejo a lo largo de toda la escala. El modelo calcula de manera automática cuánto peso asignar a cada factor para aproximarse lo mejor posible a los puntajes observados, y separa la parte del fenómeno que logra explicar de la parte que queda sin explicar. Se eligió este modelo, por encima de alternativas más complejas, precisamente por su transparencia: en una investigación exploratoria como esta, donde el objetivo es comprender qué factores se relacionan con el fenómeno y no construir un predictor de alta precisión, la simplicidad y la interpretabilidad directa de la regresión lineal resultan más valiosas que la sofisticación de métodos cuyos resultados serían más difíciles de leer y de justificar ante un lector no especializado.

El primer modelo propuesto es una regresión lineal múltiple, en la cual la variable dependiente es el índice de deuda cognitiva (puntaje estandarizado de 0 a 100) y las variables independientes son un conjunto de factores externos al cuestionario: la frecuencia de uso académico de IA, la frecuencia de uso general, el perfil de la carrera y el tipo de universidad.

Es importante aclarar por qué no se utilizan como predictoras las respuestas individuales a las preguntas que componen el índice. Dado que el índice se construye sumando esas mismas respuestas, usarlas para predecirlo equivaldría a explicar un total a partir de sus propias partes, lo que arrojaría un ajuste artificialmente perfecto y carente de valor interpretativo. Por esa razón el modelo se construye exclusivamente con variables de contexto independientes del índice, lo que permite evaluar en qué medida factores ajenos al instrumento ayudan a anticipar el nivel de deuda cognitiva. La regresión lineal es el modelo más adecuado para una primera aproximación exploratoria al fenómeno: produce coeficientes directamente interpretables que indican en qué medida y en qué sentido cada variable se asocia con el índice de deuda cognitiva, y permite

cuantificar mediante el coeficiente de determinación R cuadrado qué proporción de la variabilidad del índice es explicada por el conjunto de predictores.

En este caso este modelo aporta dos productos analíticos centrales. El primero es el coeficiente asociado a cada variable independiente, que indica en qué medida y en qué sentido (positivo o negativo) cada variable se asocia con el índice de deuda cognitiva, manteniendo constantes las demás variables.

Este coeficiente permite identificar qué dimensiones o preguntas tienen mayor poder explicativo. El segundo producto es el coeficiente de determinación, denominado R cuadrado, que indica qué proporción de la variabilidad del índice es explicada por el conjunto de las variables predictoras. Dado el tamaño muestral acotado, no se espera obtener valores de R cuadrado elevados, y los resultados deben interpretarse con prudencia.

Limitaciones del enfoque y carácter exploratorio

Es indispensable explicitar las limitaciones del enfoque de modelado propuesto. Con un tamaño muestral de 114 observaciones es dificil producir resultados predictivos generalizables. Los modelos sirven en este contexto como herramienta exploratoria para describir asociaciones presentes en la muestra particular estudiada, identificar patrones preliminares e ilustrar la viabilidad metodológica de una aproximación cuantitativa al fenómeno.

Cabe destacar que, cualquier conclusión derivada de los modelos debe interpretarse con cautela y formular hipótesis sujetas a verificación en estudios posteriores con muestras significativamente mayores.

Asimismo, el muestreo no probabilístico por conveniencia introduce limitaciones de validez externa: los resultados no son extrapolables al conjunto de estudiantes universitarios argentinos; valen exclusivamente para la muestra específica obtenida. Estas limitaciones definen con claridad el alcance y la contribución del estudio: el trabajo aporta un instrumento de medición preliminar y una metodología de análisis que puede ser replicada y ampliada en investigaciones de mayor escala, donde modelos de mayor complejidad y muestras más amplias permitirán profundizar el análisis.

7. Resultados Esperados

La presente investigación se propone obtener un conjunto de resultados articulados en torno a cuatro dimensiones: el diseño y validación preliminar del instrumento de medición, la caracterización del nivel de deuda cognitiva en la muestra estudiada, la identificación exploratoria de variables asociadas al fenómeno y la formulación de orientaciones pedagógicas e institucionales basadas en los hallazgos.

En primer lugar, se espera obtener un instrumento de medición de la deuda cognitiva con propiedades psicométricas preliminares aceptables. Concretamente, se prevé alcanzar un coeficiente alfa de Cronbach igual o superior a 0,70 sobre el conjunto del cuestionario, valor que indicaría que los ítems miden de manera consistente un constructo subyacente común.

En segundo lugar, se espera obtener una caracterización descriptiva del nivel de deuda cognitiva en la muestra de estudiantes universitarios de grado de universidades públicas y privadas de Rosario y Buenos Aires. Sobre la base de la literatura disponible y de la observación del uso intensivo de IA en este sector, se anticipa que una proporción significativa de los estudiantes presentará niveles de deuda cognitiva entre medios y altos. Específicamente, se hipotetiza que entre el cincuenta por ciento de la muestra se ubicará en las categorías media o alta del índice. Esta proyección se sustenta en la convergencia de múltiples líneas de evidencia desarrolladas en el marco teórico: la brecha entre eficiencia percibida y aprendizaje real documentada por Fan, Deng y Liu (2025), la degradación del rendimiento de los LLMs en modo conversacional que Laban et al. (2025) demostraron, y el mecanismo identificado a partir de la Taxonomía de Bloom por el cual los estudiantes pueden producir resultados que aparentan niveles altos de pensamiento sin haber activado los procesos cognitivos correspondientes. La combinación de estos factores configura un escenario en el que la deuda cognitiva puede acumularse de manera silenciosa, sin que el estudiante ni la institución dispongan de señales de alerta tempranas.

En tercer lugar, se espera identificar asociaciones exploratorias entre el nivel de deuda cognitiva y diferentes variables. A partir de la literatura revisada, se anticipan tres hipótesis principales. La primera es que la frecuencia de uso académico de herramientas de IA mostrará una asociación positiva con el índice de deuda cognitiva: a mayor frecuencia de uso, mayor nivel de deuda. La segunda es que la conducta de revisión mostrará una asociación negativa con el índice: estudiantes que revisan, cuestionan y modifican las respuestas generadas por IA presentarán menores niveles de deuda cognitiva, incluso si su frecuencia de uso es alta. La tercera, de carácter más exploratorio, es que la comparación entre frecuencia de uso académico y frecuencia de uso general (fuera de la academia) permitirá identificar si los estudiantes cuyo uso de IA trasciende el contexto universitario presentan niveles de deuda cognitiva más pronunciados, lo cual indicaría que la dependencia tecnológica se extiende más allá del entorno de estudio.

En cuarto lugar, respecto del componente de modelado exploratorio, se espera obtener modelos cuya capacidad predictiva sea limitada en términos absolutos (con valores de R cuadrado moderados en el caso de la regresión lineal), pero útiles para identificar las variables con mayor peso explicativo. No se espera ni se busca un modelo de alta precisión predictiva, dado el tamaño muestral. El valor del componente de modelado reside, fundamentalmente, en la prueba de concepto que aporta sobre la viabilidad metodológica del enfoque, no en la capacidad predictiva alcanzada.

En quinto lugar, se espera que los hallazgos del trabajo permitan formular un conjunto de orientaciones pedagógicas e institucionales para universidades que enfrenten el desafío de integrar herramientas de IA en sus prácticas académicas. Estas orientaciones se desarrollarán en mayor profundidad en la sección posterior dedicada a estrategias de mitigación, pero anticipan recomendaciones en cuatro ejes: diseño de actividades académicas que requieran explícitamente involucramiento cognitivo del estudiante incluso cuando se utiliza IA; desarrollo de competencias de alfabetización digital crítica que incluyan comprensión de los sesgos y limitaciones de los LLMs; reformulación de criterios y prácticas de

evaluación que minimicen la posibilidad de delegación total del trabajo cognitivo; y promoción de estrategias metacognitivas que permitan a los estudiantes regular reflexivamente su propio uso de la tecnología.

Finalmente, se espera que el conjunto del trabajo aporte evidencia conceptual y metodológica que enriquezca el debate académico sobre IA y educación superior, contribuyendo a una comprensión más rigurosa de los efectos cognitivos derivados del uso intensivo de estas tecnologías y, en particular, sentando las bases para una línea de investigación local sobre deuda cognitiva en estudiantes universitarios argentinos. El trabajo asume explícitamente su carácter de etapa inicial dentro de una agenda de investigación más amplia: lo que aquí se propone como prueba de concepto puede, en investigaciones posteriores, escalar hacia estudios longitudinales, muestras representativas y validación psicométrica completa.

El presente trabajo adopta una estrategia gradual y exploratoria, optando por un modelo simple y adecuado al alcance del estudio: una regresión lineal múltiple. Por lo tanto, el diseño adoptado en este trabajo deja explícitamente sentadas las bases metodológicas y los criterios de codificación necesarios para que futuras investigaciones con muestras de mayor escala puedan implementar otros modelos de ensamble más avanzados dando continuidad al análisis iniciado aquí sin necesidad de reformular el constructo ni el cuestionario.

8. Análisis de Resultados

El presente capítulo expone los resultados obtenidos a partir de la aplicación del cuestionario de deuda cognitiva diseñado en el capítulo de metodología. La recolección de datos se llevó a cabo durante el mes de mayo de 2026 mediante un formulario digital distribuido a estudiantes universitarios y de tecnicaturas en las ciudades de Rosario y Buenos Aires. El instrumento alcanzó estudiantes de universidades públicas y privadas.

Del total de respuestas recibidas a través del formulario digital, se excluyeron tres por estar incompletas o no corresponder a la población de estudio, de modo que el

análisis se realizó sobre 114 respuestas válidas, provenientes de 18 instituciones educativas y 35 carreras o programas de formación distintos. La totalidad de estas respuestas resultó completa, sin valores faltantes en ninguno de los trece ítems que componen el índice de deuda cognitiva, lo que permitió incluir la muestra íntegra en el análisis sin descartar observaciones adicionales.

La composición de la muestra refleja una diversidad institucional significativa. Del total de respuestas, 89 (78,1%) corresponden a estudiantes de universidades privadas y 25 (21,9%) a estudiantes de universidades públicas. Las instituciones con mayor representación son la Universidad Católica Argentina con 43 respuestas (37,7%), la Universidad Nacional de Rosario con 22 respuestas (19,3%) y la Universidad del Centro Educativo Latinoamericano con 11 respuestas (9,6%). El resto se distribuye entre quince instituciones adicionales, entre ellas el Instituto Universitario Italiano de Rosario, la Universidad Abierta Interamericana, la Universidad de Palermo, la Universidad Tecnológica Nacional y la Universidad FASTA.

En cuanto a la distribución por carrera, la mayor concentración se observa en Licenciatura en Recursos Humanos con 30 respuestas (26,3%), seguida por Ingeniería en Sistemas con 14 respuestas (12,3%). Otras carreras con representación relevante incluyen Comunicación Periodística, Licenciatura en Administración de Empresas y Medicina. Esta heterogeneidad resultó decisiva para el análisis, ya que permitió contrastar patrones de deuda cognitiva entre disciplinas con perfiles formativos distintos, una comparación que, como se verá, arrojó uno de los hallazgos más significativos del estudio.

La totalidad del análisis estadístico presentado en este capítulo se realizó con JASP, software libre y de código abierto desarrollado en la Universidad de Ámsterdam y de uso creciente en investigación académica.

El análisis se organiza en siete subsecciones. La primera presenta las estadísticas descriptivas del índice y su distribución por niveles. La segunda examina la relación entre frecuencia de uso de IA y deuda cognitiva. La tercera analiza las diferencias por carrera, la cuarta por universidad. La quinta presenta el perfil de respuestas por ítem y por dimensión. La sexta evalúa la validez del instrumento

mediante el coeficiente alfa de Cronbach y el análisis de discriminación de ítems. La séptima contrasta las hipótesis planteadas en el capítulo de resultados esperados con la evidencia empírica obtenida.

8.1 Estadísticas descriptivas del índice de deuda cognitiva

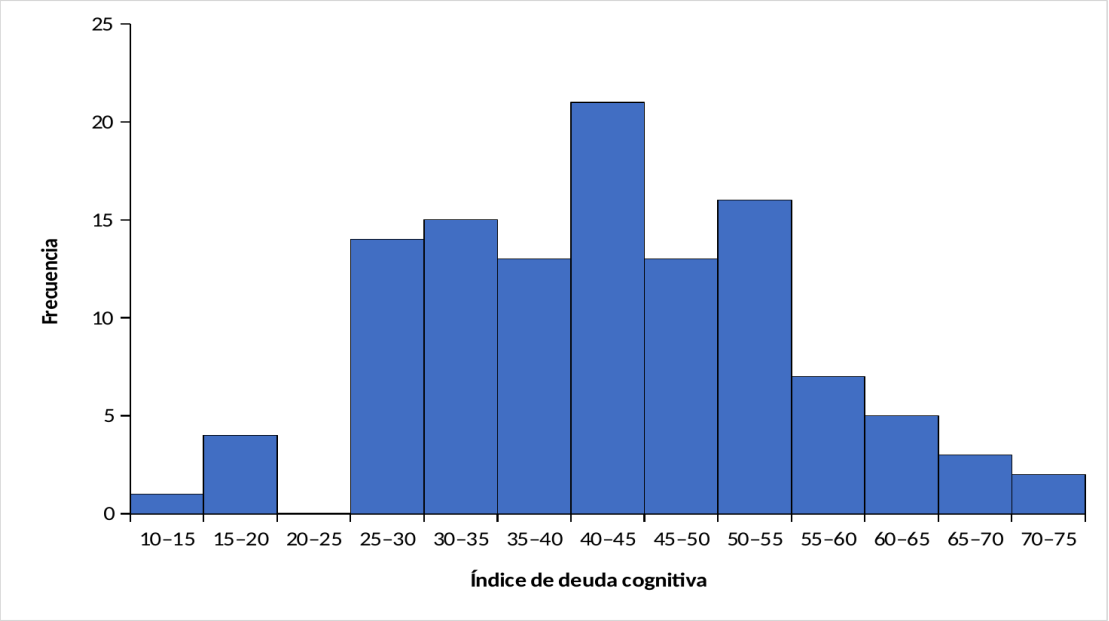


Figura 3. *Distribución del índice de deuda cognitiva en la muestra (n = 114). La media (42,6) y la mediana (43,3) prácticamente coinciden, indicando una distribución aproximadamente simétrica.*

Antes de presentar los valores vale recordar cómo se construye el índice, cuyo procedimiento detallado se desarrolló en la sección 6.3. Cada estudiante responde trece ítems puntuados de 1 a 5; los cuatro ítems formulados en sentido inverso se transforman aplicando (6 menos el valor original) para que todos apunten en la misma dirección; los trece valores se suman para obtener un puntaje crudo entre 13 y 65; y ese puntaje se estandariza a una escala de 0 a 100 mediante la fórmula índice = ((crudo − 13) / 52) × 100. Un valor de 0 corresponde a un estudiante que respondió con el mínimo de deuda en todos los ítems, y un valor de 100 a uno que respondió con el máximo.

La Figura 3 resume la distribución del índice de deuda cognitiva en la muestra. Calculado según el procedimiento de la sección 6.3, el índice arroja una media de 42,6 puntos sobre 100, una mediana de 43,3 y un desvío estándar de 12,5. El

valor más bajo observado es 11,5 y el más alto, 75,0, con un rango intercuartílico que va de 34,6 a 51,9. Que la media y la mediana casi coincidan indica una distribución bastante simétrica, sin sesgos marcados: el fenómeno se reparte de forma relativamente pareja entre los estudiantes y no depende de un puñado de casos extremos que arrastren el promedio.

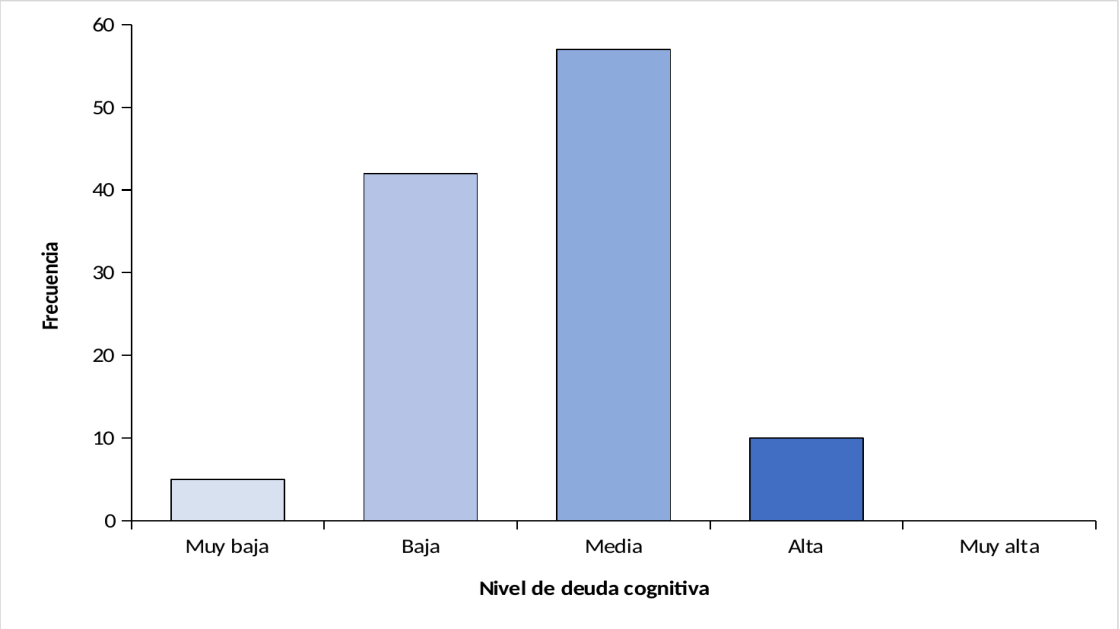


Figura 4. Distribución de los estudiantes según las cinco categorías del nivel de deuda cognitiva. El 50 % se concentra en el nivel medio y no se registran casos en el nivel muy alto.

La clasificación según los puntos de corte definidos en la metodología arroja la siguiente distribución: 5 estudiantes (4,4%) se ubican en deuda cognitiva muy baja, 42 (36,8%) en baja, 57 (50,0%) en media, 10 (8,8%) en alta y ninguno alcanza la categoría muy alta. La concentración del 50% de la muestra en el nivel medio configura una distribución con forma de campana centrada ligeramente por encima del punto medio teórico de la escala. La Figura 4 representa gráficamente esta distribución por categorías.

El dato más relevante de esta distribución es que el 58,8% de los estudiantes presenta niveles de deuda cognitiva media o alta. Esta cifra se ubica dentro del rango del 40% al 60% anticipado en los resultados esperados y constituye la primera evidencia cuantitativa sobre la magnitud del fenómeno en el contexto

universitario argentino. La ausencia de casos en la categoría muy alta admite una interpretación matizada: la deuda cognitiva es un fenómeno extendido pero aún no extremo en la población estudiada. La mayoría de los estudiantes se encuentra en una zona intermedia donde la delegación cognitiva es significativa pero no total, y donde, como se demostrará en la subsección 8.5, las conductas de revisión crítica todavía operan como mecanismo de contención. Este matiz es importante: el estudio no documenta un escenario de dependencia consumada, sino un proceso en curso, lo que ubica a las estrategias de mitigación en un momento oportuno para intervenir.

8.2 Frecuencia de uso y su relación con la deuda cognitiva

El análisis de la frecuencia de uso de IA en el ámbito académico revela un patrón de adopción masiva. El 86,8% de los estudiantes utiliza herramientas de IA generativa al menos algunas veces por semana para tareas académicas: el 42,1% algunas veces por semana, el 34,2% todos los días y el 10,5% varias veces al día. Solo el 1,8% declara no usar IA nunca. Estos datos confirman empíricamente lo que la sección 2.3 planteaba a nivel descriptivo: la IA generativa se ha integrado de manera profunda y cotidiana en las prácticas académicas de los estudiantes argentinos. El uso fuera del ámbito académico es también elevado, alcanzando al 78,1% de la muestra al menos semanalmente.

Para examinar la relación entre la frecuencia de uso y el índice se utiliza el coeficiente de correlación de Pearson (r), formalizado originalmente por Pearson (1895) y constituido desde entonces en el estadístico convencional para medir la fuerza y dirección de la asociación lineal entre dos variables cuantitativas. Este coeficiente toma valores entre −1 (asociación perfectamente negativa) y +1 (asociación perfectamente positiva); valores cercanos a cero indican ausencia de asociación. Siguiendo las convenciones expuestas por Cohen, Cohen, West y Aiken (2003) para las ciencias sociales, valores en torno a 0,3 se consideran débiles, en torno a 0,5 moderados y por encima de 0,6 fuertes.

La relación entre frecuencia de uso académico e índice de deuda cognitiva es marcadamente positiva. La correlación de Pearson entre ambas variables es r =

0,638, un valor que en ciencias sociales se considera fuerte. El gradiente es claro y monotónico, sin excepciones: los estudiantes que nunca usan IA presentan un índice medio de 26,9; quienes la usan algunas veces por mes, 28,4; algunas veces por semana, 38,5; todos los días, 49,5; y quienes la usan varias veces al día alcanzan 54,5. La diferencia entre los extremos es de 27,6 puntos, más de la mitad del rango teórico de la escala. El hecho de que cada incremento en la frecuencia de uso se corresponda con un incremento en la deuda cognitiva, sin que aparezca ningún punto de inflexión o meseta, es coherente con la noción de acumulación progresiva que define al constructo: la deuda no es un umbral que se cruza, sino una magnitud que crece de manera continua con la exposición.

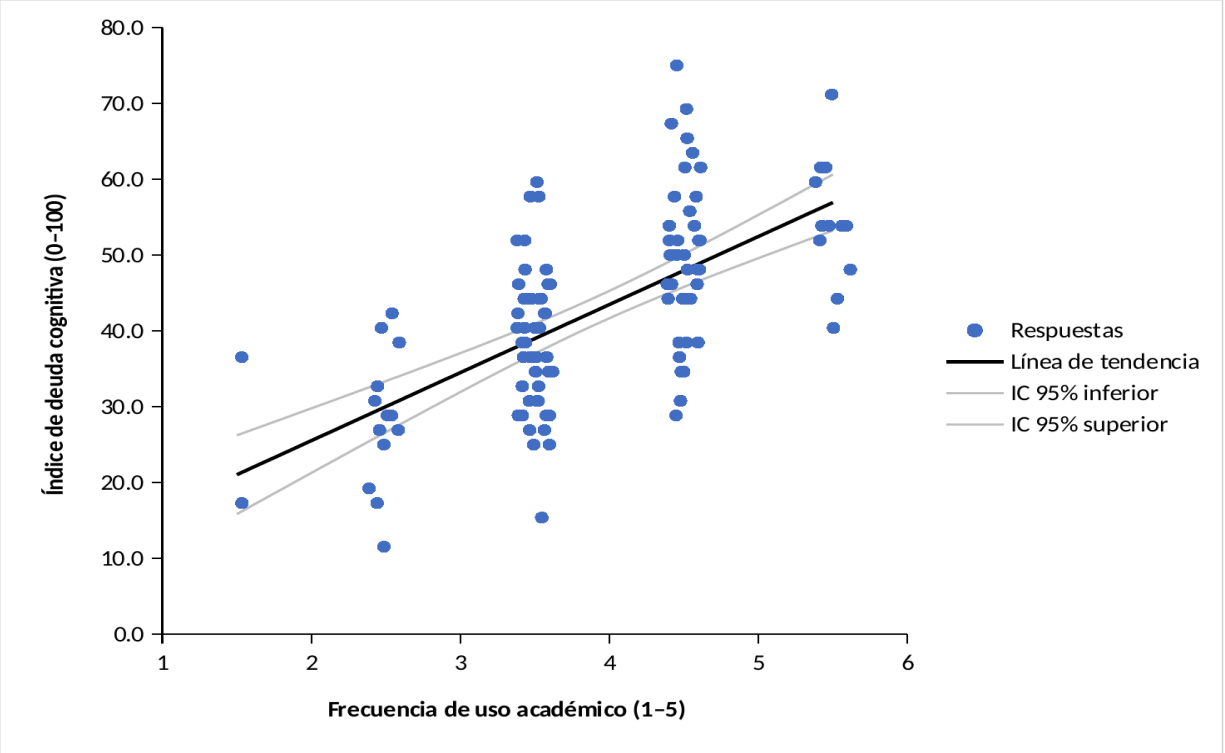


Figura 5. Relación entre la frecuencia de uso académico de IA y el índice de deuda cognitiva (n = 114). *Cada punto es un estudiante (con una leve dispersión horizontal para evitar superposición) y la recta roja muestra la tendencia general. A mayor frecuencia de uso, mayor deuda cognitiva, en línea con la correlación positiva (r = 0,638) reportada en esta sección.*

La frecuencia de uso general, fuera del ámbito académico, también muestra correlación positiva con el índice (r = 0,528). Este resultado, algo menor que el de la frecuencia académica pero igualmente sustantivo, indica que la deuda cognitiva

no es un fenómeno circunscripto al estudio, sino que se vincula con una relación más amplia y generalizada del estudiante con la tecnología. Quien delega intensivamente en la IA en su vida cotidiana tiende también a hacerlo en el plano académico, lo que sugiere que el patrón de delegación es disposicional antes que situacional.

8.3 Análisis por carrera

El análisis por carrera revela el hallazgo más robusto del estudio desde el punto de vista estadístico. Agrupando las carreras según su perfil formativo en dos grandes categorías, disciplinas discursivas o humanísticas (orientadas a la producción escrita, la argumentación y la gestión) y disciplinas técnicas o cuantitativas (ingenierías, contabilidad, arquitectura, diseño), emerge una diferencia clara: el grupo de carreras discursivas presenta un índice medio de 45,1, mientras que el grupo de carreras técnicas presenta 37,1. La diferencia de 8,1 puntos es estadísticamente significativa (prueba t de Student, t = 3,34, p = 0,001), lo que descarta que se trate de una variación producto del azar muestral. La Figura 6 representa este contraste mediante un gráfico de barras con el índice promedio de cada grupo.

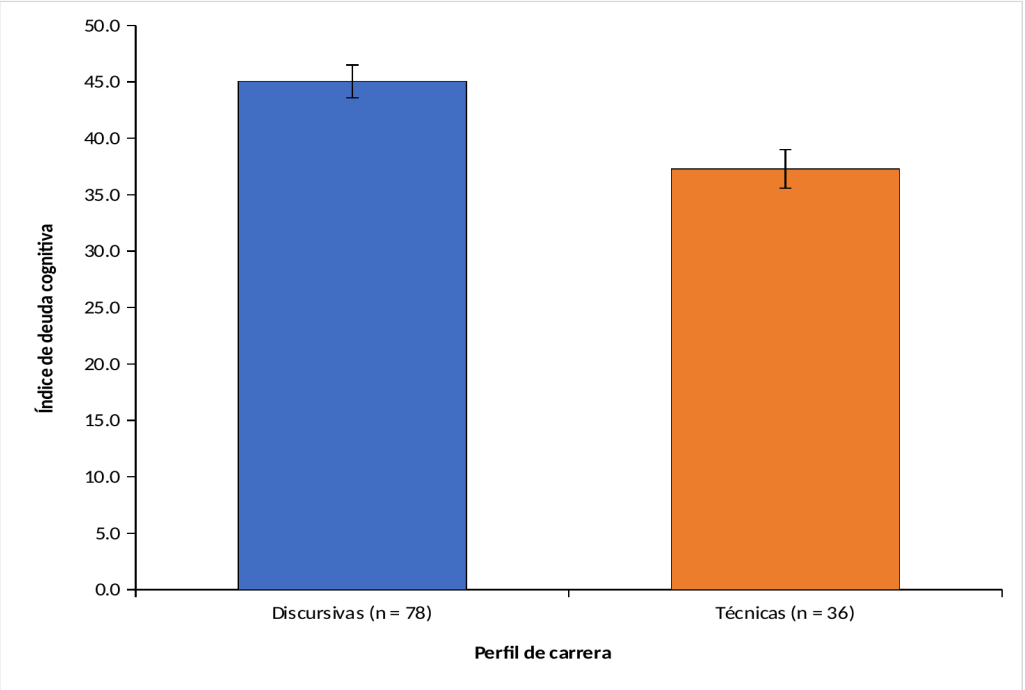


Figura 6. Comparación del índice de deuda cognitiva entre carreras discursivas y técnicas (prueba t: t = 3,34; p = 0,001).

La prueba t de Student es el contraste habitual para decidir si la diferencia entre las medias de dos grupos es real o puede deberse al azar: su valor p expresa esa probabilidad y, por convención, la diferencia se considera significativa cuando es menor a 0,05. Este criterio se aplica de manera consistente en todo el análisis.

Examinando las carreras individuales con al menos tres respuestas, las de mayor índice son Diseño y Comunicación Transmedia (49,4), Administración de Empresas (47,3), Comunicación Periodística (46,4) y Recursos Humanos (45,0). En el extremo opuesto se ubican Contador Público (21,8), Diseño de Interiores (33,3), Ingeniería en Sistemas (35,6) y Arquitectura (38,1). El patrón es consistente: cuanto más depende la carrera de la producción escrita y la elaboración conceptual, mayor es la deuda cognitiva promedio de sus estudiantes.

Este hallazgo se articula de manera directa con la Taxonomía de Bloom discutida previamente. En las carreras discursivas, las tareas académicas se concentran en los niveles superiores de la taxonomía, análisis, evaluación y síntesis, que se materializan precisamente en la escritura de ensayos, informes y argumentaciones. Son justamente esas tareas las que la IA generativa puede producir con mayor verosimilitud, y por lo tanto las que ofrecen mayor tentación de delegación. En cambio, en las carreras técnicas, una proporción mayor de las tareas involucra resolución de problemas con respuesta verificable, cálculo o diseño proyectual, donde el output de la IA es más fácilmente contrastable contra un criterio objetivo y, por ello, más difícil de delegar sin que el error se vuelva evidente.

El caso de Ingeniería en Sistemas merece atención particular y admite una segunda lectura. A pesar de ser la carrera donde cabría esperar el mayor uso de IA por afinidad técnica, su índice de deuda cognitiva es uno de los más bajos de la muestra. Una interpretación plausible es que el conocimiento sobre cómo funcionan internamente los modelos de lenguaje opera como factor protector. Quienes comprenden que un modelo puede alucinar, que su rendimiento se degrada en conversaciones extensas (Laban et al., 2025) y que la calidad de su

salida depende de variables ajenas al usuario (Leviathan et al., 2025) están mejor equipados para mantener una distancia crítica frente a sus respuestas. Este resultado refuerza la noción de alfabetización digital crítica desarrollada en la sección 5.5: el problema no se resuelve usando menos IA, sino comprendiendo mejor qué se está usando.

8.4 Análisis por universidad

El análisis por universidad, restringido a instituciones con al menos tres respuestas, muestra variaciones que reflejan tanto diferencias institucionales como la composición de carreras de cada submuestra. Entre las instituciones con muestra suficiente, la Universidad de Palermo presenta el índice más alto (59,6), seguida por el Instituto Universitario Italiano de Rosario (50,2) y la Universidad FASTA (49,4). En el rango medio se ubican la Universidad Católica Argentina (44,3) y la Universidad Nacional de Rosario (40,3). Con menor índice aparecen la Universidad Tecnológica Nacional (39,7) y la Universidad del Centro Educativo Latinoamericano (37,8).

Es tentador interpretar estas diferencias como un contraste entre gestión pública y privada, pero el análisis estadístico no respalda esa lectura. Comparando el conjunto de universidades privadas (índice medio 43,3) contra el de públicas (40,2), la diferencia de 3,1 puntos no alcanza significación estadística (prueba t de Student, $t = 1{,}07$, $p = 0{,}287$). En otras palabras, la diferencia observada es compatible con la variación esperable por azar muestral y no permite concluir que el tipo de gestión institucional, por sí mismo, determine el nivel de deuda cognitiva. Esta es una distinción metodológicamente importante: el factor que sí muestra una asociación robusta no es la titularidad de la institución, sino el perfil disciplinar de la carrera, como se demostró en la subsección anterior. Las diferencias entre universidades deben interpretarse, por lo tanto, con cautela y atendiendo a la composición de carreras de cada submuestra, dado el carácter no probabilístico del muestreo.

8.5 Perfil de respuestas por ítem y por dimensión

Mirar los datos dimensión por dimensión, más que quedarse con el índice global, permite entender dónde se concentra realmente la deuda cognitiva en la muestra. Para hacerlo, se normalizó el puntaje de cada dimensión a una escala de 0 a 100, de modo que un valor de 50 corresponde al punto neutro de la escala Likert y cualquier valor superior indica presencia del constructo.

El resultado es claro y algo paradójico. Las dos dimensiones con mayor presencia de deuda son precisamente las más silenciosas: la frecuencia de uso (57,5) y el esfuerzo mental (56,9). No son las que el estudiante percibe como problemáticas, sino las que definen su relación cotidiana con la herramienta. Le siguen la dependencia tecnológica (46,1), la fatiga cognitiva (39,9) y la autonomía cognitiva comprometida (38,6). La dimensión con menor presencia de deuda es la conducta de revisión (23,2), y aquí la lectura debe invertirse: un puntaje bajo en esta dimensión no indica poca revisión, sino mucha, dado que sus ítems fueron formulados en sentido inverso. Dicho de otro modo, la mayoría de los estudiantes todavía revisa lo que la IA produce antes de usarlo. Ese es, posiblemente, el dato más alentador de todo el estudio.

A nivel de ítems individuales, las dos conductas protectoras más extendidas son modificar o reescribir el output antes de entregarlo (media 4,36 sobre 5) y releer críticamente lo generado buscando errores (media 4,02). Que ambos ítems provengan de la dimensión de revisión y que sus valores sean los más altos del cuestionario no es un detalle menor: indica que una mayoría de estudiantes no incorpora pasivamente lo que la IA genera, sino que interviene sobre ello. La delegación existe, pero no es total.

En el otro extremo aparecen los resultados que más preocupan. La percepción de que sin IA las tareas llevan más tiempo o resultan más complicadas alcanza una media de 3,54, y el uso de IA en reemplazo de buscadores clásicos como Google llega a 3,44. Ambos superan el punto neutro de la escala. Este segundo dato merece atención especial: la sustitución del proceso de búsqueda activa de información, que históricamente exigía criterio selectivo y evaluación de fuentes, por la consulta directa a un sistema que entrega una respuesta ya elaborada, es

una de las manifestaciones más concretas de la delegación cognitiva descripta en la sección 5.2. No es un riesgo futuro: ya está ocurriendo.

Hay además tres conductas específicas que, aunque no son mayoritarias, delinean un subgrupo que merece atención. El 23,7% de los estudiantes prefiere pedirle a la IA un resumen antes que leer un texto académico completo. El 22,8% no se siente capaz de redactar un trabajo extenso sin asistencia de IA. Y el 10,5% declara sentirse bloqueado para avanzar cuando la herramienta no está disponible. Uno de cada cuatro estudiantes, aproximadamente, ya muestra rasgos de dependencia funcional. La deuda cognitiva, en esos casos, dejó de ser una tendencia para convertirse en una condición de trabajo.

8.6 Validación del instrumento

Para que un cuestionario sea algo más que una colección de preguntas, sus ítems deben medir juntos el mismo fenómeno subyacente. Esa coherencia interna se evalúa mediante el coeficiente alfa de Cronbach, propuesto por Cronbach (1951), que oscila entre 0 y 1 y cuyo umbral convencional de aceptabilidad en investigación psicométrica es 0,70. El instrumento diseñado para esta tesis alcanza un alfa de 0,719, superando ese umbral. No es un valor excepcional, pero para un instrumento exploratorio y nuevo, desarrollado sin antecedentes directos en el campo, es un resultado que respalda su uso como herramienta preliminar de medición.

El análisis de correlación ítem-total, que mide cuánto contribuye cada pregunta a discriminar entre distintos niveles de deuda cognitiva, confirma que la mayoría de los ítems funciona bien. Los tres más potentes son también los más intuitivos: la afirmación de que ante una tarea compleja la primera reacción es consultar una IA ($r = 0,62$), la frecuencia de uso académico ($r = 0,55$) y la percepción de que el rendimiento bajaría sin la herramienta ($r = 0,54$). Son, en cierto modo, los que capturan el núcleo del fenómeno con mayor precisión.

Donde el análisis arroja resultados más incómodos es en la dimensión de conducta de revisión. El ítem sobre relectura crítica del output presenta una

correlación baja (r = 0,13), y el que indaga si el estudiante usa la IA para generar un primer borrador sobre el cual trabaja luego presenta correlación negativa (r = −0,31). Si este último ítem se eliminara, el alfa subiría de 0,719 a 0,777.

La decisión de mantenerlo de todos modos no es arbitraria. La conducta que describe, usar la IA como punto de partida y luego reelaborar, es una de las más frecuentes entre los estudiantes y, como se argumentó en la sección 6.2, es también una de las más ambivalentes: combina delegación inicial con involucramiento posterior, y esa ambivalencia es precisamente lo que el instrumento no logra capturar en un solo ítem. El alfa global supera el umbral aceptable con el ítem incluido, de modo que la decisión de conservarlo responde también a no perder información sobre una conducta relevante. La recomendación para versiones futuras del instrumento es desdoblar ese ítem en dos preguntas separadas: una que mida la delegación inicial y otra que mida la profundidad de la reelaboración posterior. Esa distinción, que los datos sugieren necesaria, permitiría capturar con más precisión si el borrador de IA es realmente un punto de partida o un destino.

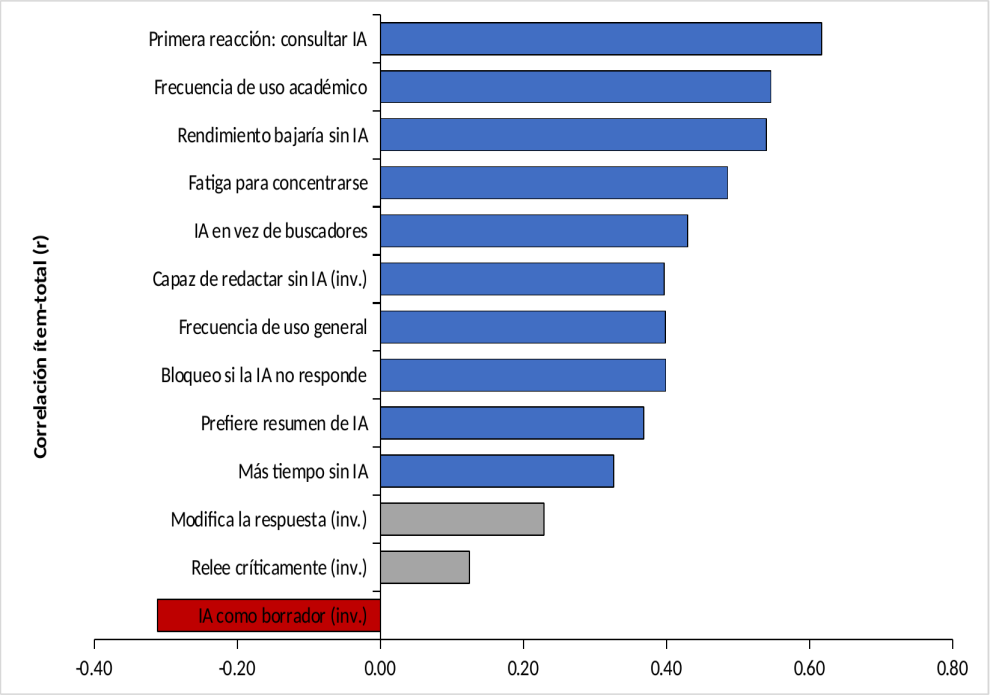


Figura 7. Correlación ítem-total corregida de cada reactivo. En rojo, el ítem con correlación negativa; en gris, ítems por debajo del umbral de 0,30. Los reactivos señalados como "ítem

inverso" fueron recodificados (6 − valor) al construir el índice, por lo que un acuerdo alto indica menor deuda cognitiva.

La Figura 7 permite ver de un vistazo lo que los números describen. Las barras azules, que concentran la mayoría de los ítems, representan las preguntas que discriminan bien entre niveles de deuda cognitiva. La barra gris corresponde al ítem de relectura crítica, que queda por debajo del umbral aceptable de 0,30. Y la barra roja, que se extiende hacia la izquierda, marca al ítem del borrador como el único con correlación negativa en todo el instrumento: el único que se comporta en sentido contrario al constructo. Esa imagen resume con precisión dónde el instrumento es sólido y dónde todavía necesita trabajo.

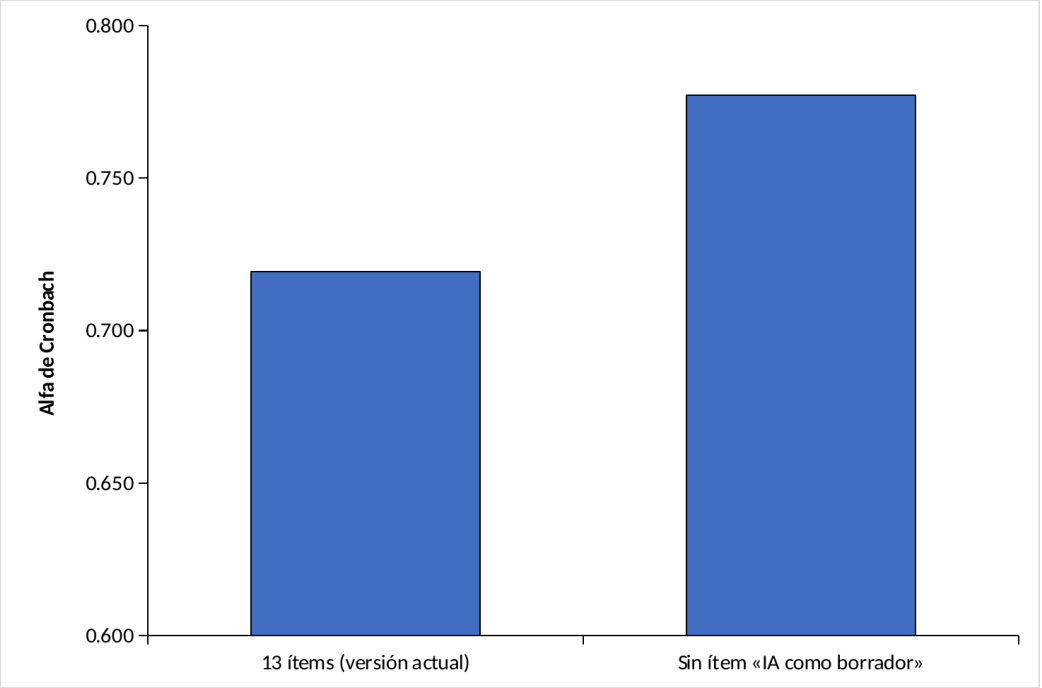


Figura 8. Coeficiente alfa de Cronbach del instrumento completo y tras la eliminación del ítem problemático. Ambos valores superan el umbral de 0,70.

Por otra parte, la Figura 8 ilustra exactamente esa tensión. Las dos barras muestran el alfa actual del instrumento completo (0,719) y el alfa que se obtendría eliminando el ítem problemático (0,777). Ambos valores superan el umbral mínimo de 0,70, lo que significa que la decisión de conservar el ítem no compromete la validez del cuestionario. La diferencia entre las dos barras no es una razón para

corregir ahora, sino una indicación de hacia dónde debería evolucionar el instrumento en versiones futuras.

Estructura interna del instrumento: análisis factorial exploratorio

Para comprobar si las seis dimensiones propuestas se reflejan efectivamente en los datos, se aplicó un análisis factorial exploratorio. En términos sencillos, esta técnica observa qué preguntas tienden a responderse de manera parecida y las agrupa: si varias preguntas “se mueven juntas”, es probable que estén captando un mismo aspecto de fondo. Dos pruebas previas confirmaron que los datos eran adecuados para este análisis: la medida de adecuación muestral KMO fue de 0,744 (por encima del umbral aceptable de 0,70) y el test de esfericidad de Bartlett resultó significativo ($p < 0,001$).

El análisis mostró que, más que seis dimensiones separadas, las respuestas se organizan en torno a tres grandes factores, que en conjunto explican el 37,2 % de la variabilidad. La Tabla 2 resume cómo se agrupan las preguntas. El primer factor reúne las conductas de uso intensivo y delegación (frecuencia de uso, recurrir a la IA como primera reacción, usarla en lugar de un buscador). El segundo agrupa las señales de dependencia y fatiga (sentirse bloqueado o incapaz de trabajar sin IA, preferir un resumen antes que leer, perder la concentración). El tercero corresponde con mucha nitidez a la conducta de revisión (releer, modificar y reelaborar lo que la IA produce).

Este resultado no contradice el marco conceptual, sino que lo refina: las seis dimensiones teóricas siguen siendo útiles para pensar el fenómeno, pero en los datos se consolidan en tres facetas más amplias, algo habitual en instrumentos en etapa exploratoria y con muestras acotadas. Dos puntos merecen destacarse. Que la conducta de revisión aparezca como un factor propio y bien definido refuerza su papel central como comportamiento protector. Y que el uso intensivo y la dependencia surjan como factores distintos sugiere que usar mucho la IA y sentirse dependiente de ella no son exactamente lo mismo. La recomendación para versiones futuras es sumar más preguntas por dimensión, de modo que cada faceta quede mejor representada y la estructura pueda confirmarse, en muestras mayores, mediante un análisis factorial confirmatorio.

Tabla 2. Agrupación de los ítems en los tres factores identificados por el análisis factorial exploratorio (rotación oblicua). Se indica la carga de cada ítem en su factor principal.

Ítem	F1: Uso intensivo y delegación (16,9 %)	F2: Dependencia y fatiga (11,0 %)	F3: Revisión crítica (9,3 %)
Frecuencia de uso académico	0,71		
Primera reacción: consultar una IA	0,66		
Usa la IA en vez de un buscador	0,59		
Frecuencia de uso general	0,56		
Estudiar sin IA le lleva más tiempo	0,51		
Sin IA su rendimiento bajaría	0,42		
Se siente capaz de trabajar sin IA (inv.)		0,63	
Pierde concentración tras usar IA		0,63	
Se bloquea cuando la IA no responde		0,59	
Prefiere un resumen de la IA antes que leer		0,41	
Modifica o reescribe lo que genera la IA (inv.)			0,74
Usa la IA como primer borrador (inv.)			0,60
Relee críticamente lo que genera la IA (inv.)			0,50

8.7 Verificación de hipótesis

Los resultados permiten contrastar las tres hipótesis planteadas en el capítulo de resultados esperados, las tres con respaldo empírico.

La primera hipótesis anticipaba una asociación positiva entre frecuencia de uso académico de IA e índice de deuda cognitiva. Se confirma con una correlación de r = 0,638 y un gradiente monotónico de 26,9 a 54,5 puntos a lo largo de los cinco niveles de frecuencia. La magnitud y la consistencia de esta asociación son coherentes con los hallazgos de Kosmyna et al. (2025) sobre la acumulación de efectos residuales derivados del uso prolongado de IA, y con la advertencia de Laban et al. (2025) sobre la degradación del rendimiento en el uso conversacional sostenido.

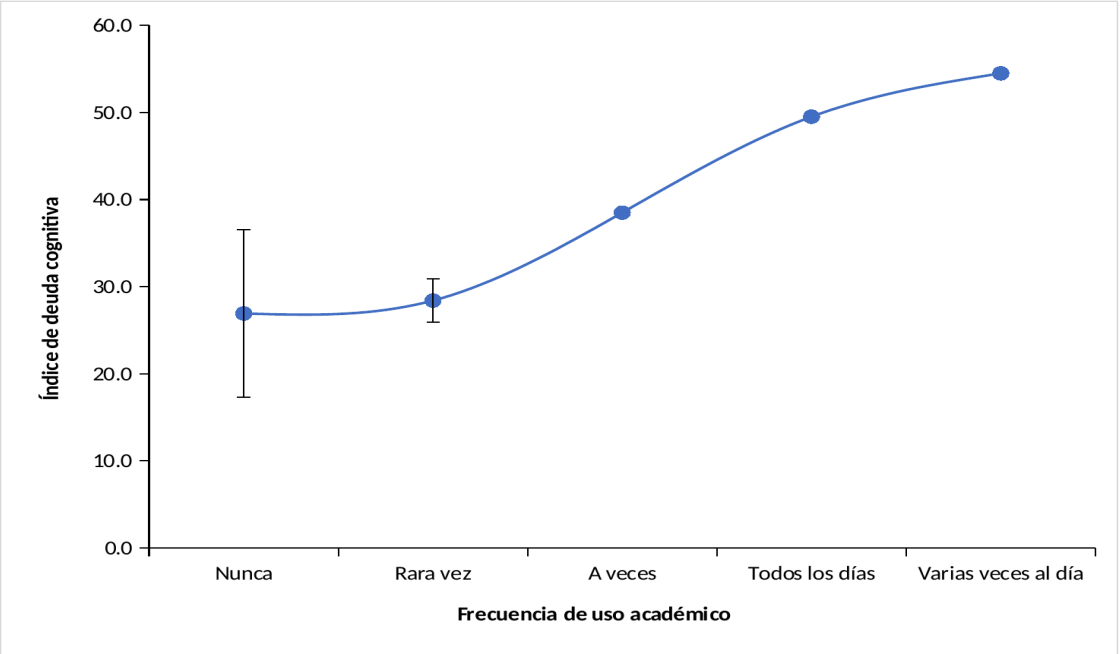


Figura 9. Índice medio de deuda cognitiva según el nivel de frecuencia de uso académico de IA. El gradiente es estrictamente creciente (r = 0,638).

Para evaluar la segunda hipótesis hubo que separar el efecto de la revisión del de la frecuencia de uso, ya que ambos actúan al mismo tiempo sobre el mismo estudiante. La estrategia fue acotar el análisis al subgrupo de mayor riesgo: los 51 estudiantes que usan IA todos los días o varias veces al día. Dentro de ese grupo, parejo en nivel de exposición, se comparó el índice de quienes revisan mucho con

el de quienes revisan poco. El resultado es claro: los primeros promedian 48,5 puntos y los segundos, 54,6, una diferencia de 6,1 puntos estadísticamente significativa (prueba t, p = 0,043). Lo relevante es que el factor protector de la revisión opera justamente en el grupo más expuesto: no anula el efecto de la alta frecuencia, pero lo atenúa de manera medible, lo que confirma el papel de la metacognición que Flavell (1979) describe en la sección 5.3 (Figura 10).

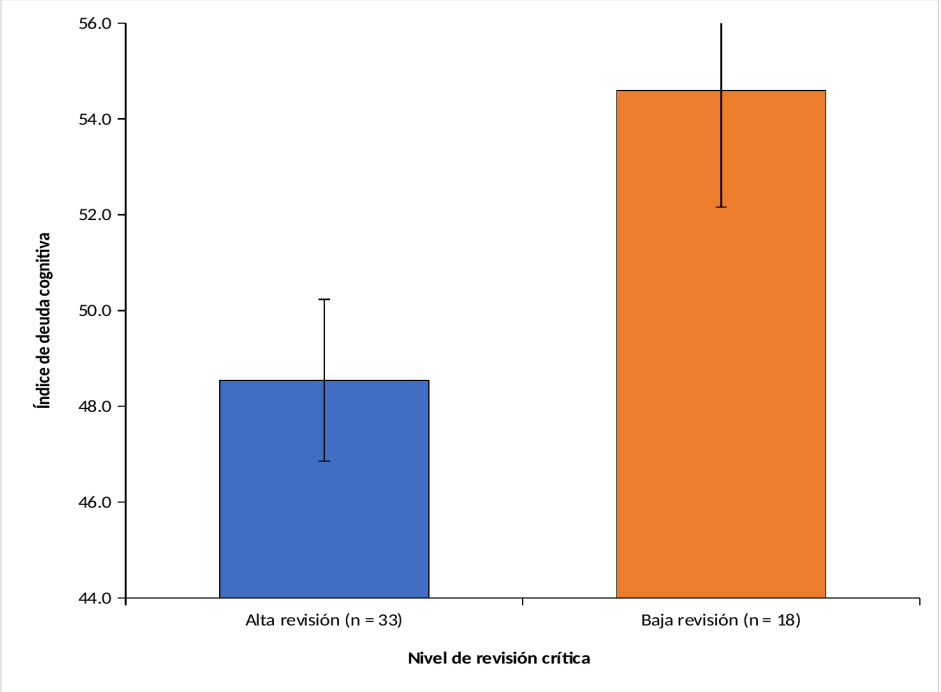


Figura 10. Índice de deuda cognitiva entre usuarios de alta frecuencia según su nivel de revisión crítica. La revisión reduce la deuda 6,1 puntos (p = 0,043) incluso en el grupo de mayor exposición.

La tercera hipótesis proponía que la frecuencia de uso general, fuera de la academia, se asociaría con mayor deuda cognitiva en el contexto de estudio. Se confirma con una correlación de r = 0,528, lo que indica que los patrones de dependencia tecnológica trascienden el entorno académico y reflejan una relación más amplia del estudiante con la tecnología.

Por último, el estudio documentó un hallazgo no anticipado explícitamente en las hipótesis pero de gran relevancia teórica: la diferencia estadísticamente significativa entre carreras discursivas y técnicas. Este resultado, articulado con la Taxonomía de Bloom, sugiere que la vulnerabilidad a la deuda cognitiva no es

uniforme sino que depende del tipo de operación intelectual que cada disciplina demanda, y abre una línea de indagación que las hipótesis iniciales no habían contemplado. En conjunto, la convergencia entre las hipótesis confirmadas y este hallazgo emergente otorga al instrumento una validez de criterio preliminar alentadora: el índice no solo es internamente consistente, sino que se comporta en relación con otras variables de la manera en que la teoría predice que debería hacerlo.

Tabla 3. Síntesis de las hipótesis planteadas y su contraste con la evidencia empírica.

Hipótesis	Predicción	Qué muestran los datos	Estado
H1	Mayor frecuencia de uso académico = mayor deuda	Cuanto más seguido usan IA para estudiar, más deuda: quienes la usan varias veces al día casi duplican a quienes casi no la usan.	Confirmada
H2	La revisión crítica actúa como factor protector	Entre los que más usan IA, los que revisan lo que la IA genera tienen menos deuda que los que no revisan.	Confirmada
H3	Mayor uso general (fuera de la academia) = mayor deuda	Lo mismo ocurre con el uso fuera de la facultad: a más uso en la vida diaria, más deuda.	Confirmada
Hallazgo emergente	Las carreras discursivas presentan más deuda que las técnicas	Las carreras de escritura y argumentación muestran bastante más deuda que las técnicas o de números.	No previsto, confirmado

8.8 Modelado exploratorio: regresión

Tal como se anticipó en la sección 6.5, esta investigación incluye una etapa exploratoria de modelado. Su objetivo no es construir una herramienta de predicción definitiva, sino comprobar si el nivel de deuda cognitiva puede estimarse con un método sencillo y dejar un punto de partida metodológico para

estudios futuros con muestras más grandes. Para ello se emplea una regresión lineal múltiple, que busca entender qué factores del contexto y del comportamiento del estudiante, con qué frecuencia usa IA, qué tipo de carrera cursa, en qué universidad estudia, ayudan a anticipar su nivel de deuda cognitiva.

Conviene aclarar una decisión metodológica. El diseño original (sección 6.5) preveía explicar el índice de deuda cognitiva a partir de las trece respuestas que lo componen. Pero como el índice se calcula directamente a partir de esas trece respuestas, predecirlo con ellas es circular: el modelo se explicaría a sí mismo y no aportaría nada nuevo.

Por eso se optó por explicar el índice usando variables que no forman parte de su cálculo, la frecuencia de uso, el tipo de carrera y el tipo de universidad, lo que sí permite responder una pregunta con sentido: ¿qué factores del contexto del estudiante se asocian con una mayor deuda cognitiva?

A lo largo de esta sección aparecen algunos términos estadísticos que conviene aclarar en lenguaje sencillo, sin entrar en su tratamiento matemático. Un coeficiente (β) indica cuánto sube o baja el índice de deuda cognitiva cuando un factor aumenta: si es positivo, ese factor se asocia a más deuda; si es negativo, a menos.

El valor p estima la probabilidad de que una relación observada sea fruto del azar, de modo que cuanto más chico es, mayor confianza ofrece (por convención se considera sólida cuando es menor a 0,05). Y el R^2 expresa qué porcentaje de las diferencias entre estudiantes logra explicar el modelo. Importa subrayar que el modelo no prueba causas: describe asociaciones que el marco teórico permite interpretar.

Modelo de regresión lineal con variables externas

El modelo de regresión explica el índice de deuda cognitiva (de 0 a 100) a partir de cuatro factores: con qué frecuencia el estudiante usa IA para tareas académicas, con qué frecuencia la usa en su vida en general, si su carrera es de perfil discursivo o humanístico frente a técnico o cuantitativo, y si estudia en una

universidad privada o pública. Los dos últimos factores se construyeron siguiendo los criterios de las secciones 8.3 y 8.4.

En conjunto, estos cuatro factores explican alrededor del 50,6 % de las diferencias de deuda cognitiva entre los estudiantes de la muestra (R^2 = 0,506). Es un resultado considerable: significa que con solo cuatro datos de contexto, sin mirar las respuestas que componen el índice, ya se captura la mitad del fenómeno. Las pruebas de validación, pensadas para descartar que el modelo esté simplemente memorizando esta muestra, confirman que el ajuste se sostiene de manera razonable. En promedio, el modelo se equivoca unos siete puntos al estimar el índice de un estudiante individual, una precisión propia de una herramienta orientativa y no de un diagnóstico individual.

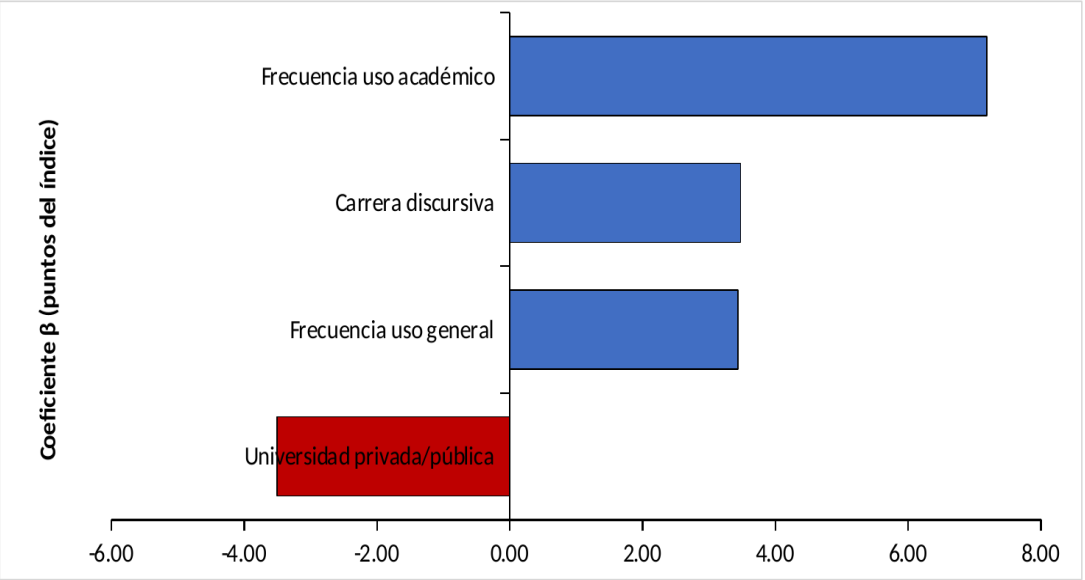


Figura 11. Coeficientes (β) de la regresión lineal múltiple. Cada barra indica cuántos puntos del índice (escala 0–100) sube o baja la deuda cognitiva cuando ese factor aumenta un nivel, manteniendo los demás constantes. Las barras hacia la derecha (azules) se asocian a más deuda; la barra hacia la izquierda (roja, "universidad privada") a menos. Cuanto más larga la barra, más fuerte el efecto. En conjunto, los cuatro factores explican el 50,6 % de las diferencias de deuda entre estudiantes (R^2 validación cruzada = 0,417).

Conviene aclarar cómo leer el gráfico: cada factor recibe un número (su coeficiente) que dice cuánto mueve la deuda cognitiva; positivo significa que la sube y negativo que la baja. La Figura 11 muestra el peso de cada factor. El más

fuerte es, con claridad, la frecuencia de uso académico de IA (β = +7,13): cada salto de un nivel de frecuencia se asocia, en promedio, a unos siete puntos más de deuda cognitiva. La frecuencia de uso general fuera de la facultad aparece en segundo lugar (β = +3,44) y suma su efecto de manera independiente. Ambos resultados confirman, ahora con un modelo que considera varios factores a la vez, lo que el marco teórico anticipaba: la deuda cognitiva no se vincula tanto con la IA en sí como con la intensidad y la habitualidad con que se delega en ella el esfuerzo intelectual. Cuanto más se vuelve la consulta a la IA una reacción automática, en los términos del Sistema 1 de Kahneman, menos se ejercita el pensamiento deliberado del Sistema 2 que la educación universitaria busca desarrollar.

El perfil disciplinar de la carrera también pesa (β = +3,78), aunque su valor p (0,054) queda justo en el límite de lo que se considera estadísticamente sólido. La dirección, de todos modos, es coherente con lo hallado en la sección 8.3: las carreras discursivas presentan, en promedio, casi cuatro puntos más de deuda cognitiva que las técnicas, incluso comparando estudiantes con la misma frecuencia de uso.

El dato más interesante es el del tipo de universidad (β = -3,40). El signo es negativo, lo que indicaría que, una vez descontado el efecto de la frecuencia de uso y del perfil de la carrera, los estudiantes de universidades privadas tendrían en realidad algo menos de deuda cognitiva que los de públicas: justo lo contrario de lo que sugería la comparación simple de la sección 8.4. La explicación es sencilla y vale la pena detenerse en ella. Aquella comparación atribuía a la universidad una diferencia que en realidad venía del tipo de carreras que cada grupo concentra. Las universidades privadas de la muestra reúnen más estudiantes de carreras discursivas, que son las más vulnerables. Cuando el modelo separa ambos efectos, queda claro que lo que importa no es si la universidad es pública o privada, sino el perfil de la carrera. Es un buen ejemplo de cómo una tercera variable puede crear una asociación engañosa entre otras dos.

Síntesis del modelado y limitaciones

La regresión deja una conclusión clara: cerca de la mitad de las diferencias de deuda cognitiva entre estudiantes puede explicarse solo con datos de contexto, frecuencia de uso y perfil de la carrera, sin recurrir a los ítems del cuestionario. Esto confirma que esos factores son determinantes reales del fenómeno y no coincidencias. A su vez, la validación del instrumento (sección 8.6) mostró qué preguntas discriminan mejor entre niveles de deuda: la primera reacción de consultar la IA ante una tarea compleja, la frecuencia de uso académico y la percepción de que el rendimiento bajaría sin IA encabezan la correlación ítem-total. Esa información abre una línea aplicada concreta, el diseño de una versión breve del instrumento, que se retoma en la sección 10.1.

Las limitaciones deben quedar claras. La muestra, de 114 respuestas, superó la estimación inicial pero sigue siendo modesta para fines predictivos exigentes. Al no ser un muestreo aleatorio, los resultados describen a esta muestra y no pueden extrapolarse sin más al conjunto de los estudiantes argentinos. Varios subgrupos de carreras y universidades quedaron con muy pocos casos, lo que impide análisis finos a ese nivel. Y, al tratarse de un estudio de un único momento, no permite ver cómo evoluciona la deuda cognitiva en un mismo estudiante a lo largo del tiempo, algo que solo un diseño longitudinal podría capturar.

Con esas limitaciones a la vista, el modelado cumple lo que se propuso: mostrar que el fenómeno puede estudiarse con herramientas accesibles y que existen relaciones medibles entre el contexto del estudiante y su deuda cognitiva. No es un punto de llegada sino un piso: con muestras de varios cientos o miles de respuestas, técnicas más avanzadas como *Random Forest* (Breiman, 2001) darían estimaciones más estables manteniendo esta misma línea de trabajo. En lo personal, lo que más me quedó de esta etapa fue comprobar con números algo que venía percibiendo de manera difusa: que el factor decisivo no es tanto la IA como herramienta, sino la frecuencia con que uno deja de poner en juego el propio esfuerzo.

9. Conclusiones

La presente investigación se propuso operacionalizar el concepto de deuda cognitiva en el contexto de la educación superior argentina, diseñar un instrumento de medición preliminarmente validado y caracterizar el fenómeno en una muestra de estudiantes universitarios de Rosario y Buenos Aires. El trabajo partió de una pregunta simple pero de difícil respuesta empírica: ¿en qué medida el uso intensivo de herramientas de Inteligencia Artificial generativa está afectando los procesos cognitivos de los estudiantes universitarios? Las 114 respuestas obtenidas, el instrumento construido y los análisis realizados permiten ofrecer ahora respuestas provisorias y metodológicamente sustentadas a esa pregunta.

9.1 Síntesis de hallazgos

El primer hallazgo central es que la deuda cognitiva es un fenómeno empíricamente medible. El instrumento de trece ítems diseñado para este estudio alcanzó un coeficiente alfa de Cronbach de 0,719, superando el umbral convencional de 0,70 que la investigación psicométrica considera indicativo de consistencia interna aceptable. Los ítems discriminan entre participantes de manera coherente, con correlaciones ítem-total que van desde 0,33 hasta 0,62 para la mayoría de los reactivos. Esto significa que el constructo no es una abstracción teórica vacía: tiene manifestaciones conductuales y perceptuales observables que un cuestionario estructurado puede capturar con suficiente fiabilidad para una investigación exploratoria.

El segundo hallazgo es que la deuda cognitiva está presente en la muestra con una distribución que invita a la reflexión. El índice medio de 42,6 sobre 100 ubica a la muestra en el rango bajo-medio del espectro. El cincuenta por ciento de los estudiantes se concentra en la categoría media, el 36,8% en la categoría baja y el 8,8% en la categoría alta. No se registró ningún caso en el nivel muy alto, lo que sugiere que la deuda cognitiva en su forma más extrema es poco frecuente en la población estudiada. Sin embargo, la distribución tampoco ofrece tranquilidad: casi seis de cada diez estudiantes presentan niveles medios o superiores, lo que indica que el fenómeno está extendido y no se limita a casos marginales.

El tercer hallazgo, y el más robusto desde el punto de vista estadístico, es la asociación entre frecuencia de uso académico de IA e índice de deuda cognitiva. La correlación de $r = 0{,}638$ ($p < 0{,}001$) es la más fuerte reportada en este estudio, y se acompaña de un gradiente perfectamente monotónico: los estudiantes que nunca usan IA presentan un índice medio de 26,9 puntos, mientras que quienes la usan varias veces al día alcanzan 54,5 puntos. Esta diferencia de casi 28 puntos a lo largo del espectro de frecuencia es conceptualmente coherente con la hipótesis de que la exposición sostenida a herramientas que automatizan procesos cognitivos genera efectos acumulativos sobre la autonomía intelectual. También es consistente con los hallazgos de Kosmyna et al. (2025) sobre efectos residuales del uso prolongado de IA y con la advertencia de Laban et al. (2025) sobre degradación del desempeño en uso conversacional intensivo.

El cuarto hallazgo es que la conducta de revisión opera como un factor protector real y cuantificable. Al restringir el análisis a los 51 estudiantes de alta frecuencia de uso, los de mayor riesgo, y dividirlos por su nivel de revisión crítica, se observa que quienes revisan intensamente presentan un índice de 48,5 puntos frente a 54,6 de quienes no lo hacen, una diferencia de 6,1 puntos estadísticamente significativa ($p = 0{,}043$). Este resultado tiene implicaciones pedagógicas directas: la revisión no neutraliza el efecto de la alta frecuencia, pero lo atenúa de manera mensurable incluso en el grupo más expuesto. Dicho de otro modo, no es solo cuánto se usa la IA lo que determina la deuda cognitiva; también, y mucho, cómo se usa.

El quinto hallazgo es la diferencia significativa entre carreras discursivas y técnicas: 45,1 puntos versus 37,1 puntos ($t = 3{,}34$, $p = 0{,}001$). Esta diferencia, no anticipada en las hipótesis originales, sugiere que la vulnerabilidad a la deuda cognitiva no es uniforme sino que depende del tipo de operación intelectual que cada disciplina demanda. Las carreras discursivas, cuya producción académica se basa en la escritura argumentativa, la síntesis conceptual y la elaboración textual, son precisamente aquellas donde los LLMs ofrecen el mayor nivel de automatización. Las carreras técnicas, en cambio, exigen la resolución de problemas cuantitativos y la aplicación procedimental de conocimientos que los

modelos de lenguaje replican con menor fidelidad, lo que obliga al estudiante a mayor involucramiento cognitivo real. Este hallazgo conecta directamente con la Taxonomía de Bloom: la IA automatiza con especial eficacia los niveles de comprensión y síntesis textual, que son precisamente los que las carreras discursivas más demandan.

El sexto hallazgo proviene del modelado exploratorio. La regresión lineal múltiple con cuatro variables externas, frecuencia de uso académico, frecuencia de uso general, perfil disciplinar y tipo de universidad, explica el 50,6% de la varianza del índice ($R^2 = 0,506$), con un R^2 de validación cruzada de 0,417. Este resultado, lejos de ser modesto, es llamativo: con apenas cuatro variables simples y sin recurrir a los ítems del cuestionario es posible predecir más de la mitad de la variabilidad del fenómeno. Esto confirma que la deuda cognitiva no es una experiencia aleatoria ni idiosincrática, sino que está estructuralmente asociada a variables contextuales identificables y, en principio, intervenibles.

9.2 Alcances y limitaciones

Los resultados deben leerse con las limitaciones propias de un estudio exploratorio. La muestra, si bien superó la estimación inicial, no es probabilística ni representativa del conjunto de estudiantes universitarios argentinos. El instrumento, aunque internamente consistente, requiere validación psicométrica más exhaustiva, análisis factorial confirmatorio, validez de criterio, fiabilidad test-retest, antes de poder ser utilizado con fines diagnósticos más allá del contexto académico exploratorio. El diseño transversal no permite distinguir si la alta frecuencia de uso causa la deuda cognitiva o si perfiles cognitivos previos condicionan tanto el uso como la deuda. Estas limitaciones no invalidan los hallazgos, pero definen con precisión el tipo de afirmación que este trabajo puede sostener: una primera evidencia cuantitativa sobre la existencia y distribución del fenómeno, con fundamentos suficientes para orientar investigaciones de mayor escala.

El ítem sobre el uso de la IA para generar un primer borrador sobre el que se trabaja posteriormente merece mención especial. Su correlación ítem-total

negativa (r = −0,31) y el hecho de que su eliminación eleve el alfa de 0,719 a 0,777 revelan que la conducta que describe es cualitativamente distinta del resto del constructo. Usar la IA como punto de partida para una elaboración posterior no es equivalente a usarla como sustituto del proceso cognitivo: es, más bien, una forma de co-producción que puede ser protectora o perjudicial dependiendo del grado de involucramiento real en la fase de reelaboración. Esta ambivalencia conceptual, detectada por los propios datos, es en sí misma un resultado sustantivo que enriquece la comprensión teórica del fenómeno.

9.3 Propuesta de mitigación

Un estudio que diagnostica un problema queda incompleto si no avanza hacia posibles respuestas. Si bien el diseño y la evaluación de intervenciones exceden el alcance de esta tesis, los hallazgos obtenidos permiten formular una propuesta de mitigación fundamentada en la evidencia, y no en recomendaciones genéricas. Dos resultados del estudio orientan especialmente esta propuesta: que la conducta de revisión crítica opera como factor protector incluso entre los usuarios más intensivos, y que las carreras discursivas presentan mayor vulnerabilidad que las técnicas. La propuesta se organiza en tres niveles complementarios que comparten un mismo principio rector, alineado con la literatura reciente sobre IA y aprendizaje: el objetivo no es reducir el acceso a la IA: es preservar el involucramiento cognitivo del estudiante en su uso.

Primer nivel: los docentes y las prácticas de aula

El hallazgo más accionable del estudio es que la revisión crítica atenúa la deuda cognitiva. De allí se desprende que la intervención más eficaz no consiste en prohibir o restringir el uso de IA, una estrategia de dudosa viabilidad y escaso respaldo. Consiste en transformar la forma en que se usa. La literatura sobre evaluación en la era de la IA converge en un desplazamiento que se aplica de lleno a este hallazgo: pasar de una evaluación centrada únicamente en el producto final a una que también valore el proceso. Este enfoque, denominado modelo de evaluación de proceso y producto, propone calificar tanto el trabajo entregado como la calidad de los pasos que el estudiante siguió para construirlo, incluida la

forma en que interactuó con la IA. Revisiones sistemáticas recientes sobre evaluación en entornos con IA describen este enfoque como modelo de proceso y producto, y lo señalan como una de las estrategias más prometedoras para preservar la integridad del aprendizaje (Ncube et al., 2025).

De este principio se derivan ejemplos concretos aplicables al aula. Un primer ejemplo es solicitar que, junto al trabajo final, el estudiante entregue una breve reflexión sobre qué le aportó la IA, qué modificó del texto generado y qué errores o sesgos detectó en él. Un segundo ejemplo, particularmente alineado con el factor protector identificado en este estudio, consiste en diseñar actividades donde la IA produzca deliberadamente una respuesta y la tarea del estudiante sea evaluarla, corregirla o refutarla. Esta clase de consigna convierte la revisión crítica, que el estudio identificó como protectora, en el objeto mismo de la evaluación. La propuesta es coherente, además, con la observación de que en la era de la IA el nivel de evaluar tiende a desplazar al de crear como operación cognitiva de mayor orden en la taxonomía de Bloom (Lubbe et al., 2025), un marco ya presente en esta tesis.

Segundo nivel: las instituciones y la política educativa

Dado que el fenómeno está distribuido de manera no uniforme, las instituciones podrían focalizar esfuerzos antes que aplicar medidas homogéneas. Como el estudio mostró mayor vulnerabilidad en las carreras discursivas, se propone priorizar en ellas la formación en uso crítico de IA. A nivel de lineamiento general, la evidencia internacional sugiere que las universidades deberían adoptar políticas institucionales explícitas sobre el uso de IA y no dejarlo librado a la decisión individual de cada docente, lo que genera inconsistencia y desorientación en los estudiantes. Una encuesta de UNESCO (2025) sobre educación superior halló que dos tercios de las instituciones ya cuentan con orientaciones sobre IA o las están desarrollando, lo que confirma que se trata de una tendencia ya consolidada.

Un ejemplo concreto de implementación es incorporar en los primeros años un espacio de alfabetización en IA. Es importante precisar qué se entiende aquí por alfabetización, porque la propuesta no se limita a enseñar a operar la herramienta. Siguiendo el Marco de Competencias en IA para Estudiantes de UNESCO (2024),

la alfabetización en IA debe habilitar al estudiante no solo para usar estos sistemas, sino para cuestionarlos, comprender sus límites y sesgos, y dimensionar los riesgos cognitivos de la delegación excesiva. Un segundo aporte concreto que esta tesis ofrece al nivel institucional es el propio instrumento desarrollado, que podría utilizarse como herramienta de monitoreo periódico para que cada institución observe la evolución de la deuda cognitiva en su población y evalúe el efecto de las medidas que adopte.

Tercer nivel: los estudiantes y la autorregulación

Finalmente, buena parte de la mitigación depende de la conciencia y los hábitos del propio estudiante. El estudio sugiere que el problema no es la frecuencia de uso en sí; es la ausencia de involucramiento cognitivo posterior. La investigación reciente ha dado un nombre preciso a este riesgo: la pereza metacognitiva, definida como la tendencia del estudiante a delegar en la IA los procesos de planificación, monitoreo del error y reflexión estratégica que sostienen el aprendizaje autorregulado, en lugar de asumir el esfuerzo cognitivo correspondiente (Fan, Y., et al., 2025). Este concepto es prácticamente una descripción operativa de los mecanismos que componen la deuda cognitiva medida en esta tesis, lo que refuerza la pertinencia de intervenir en este nivel.

Como lineamiento, se propone promover en los estudiantes una relación deliberada con la IA, en la que cada delegación sea una decisión consciente y no un reflejo automático. De allí surgen ejemplos concretos de autorregulación: intentar resolver una tarea por cuenta propia antes de consultar la IA, de modo que la herramienta verifique o complemente un razonamiento ya iniciado en lugar de reemplazarlo; utilizar la IA para contrastar el propio pensamiento y no para sustituirlo; y reservar deliberadamente ciertas tareas, especialmente las de escritura y síntesis, para realizarlas sin asistencia, a modo de entrenamiento de las capacidades que la delegación tiende a erosionar. Estas prácticas apuntan a preservar los procesos de planificación, monitoreo y revisión que la pereza metacognitiva amenaza.

Los tres niveles comparten el principio común enunciado al inicio y que se desprende del conjunto de la investigación: la deuda cognitiva no se combate

alejando a los estudiantes de la IA: se combate enseñándoles a usarla sin dejar de pensar. La revisión crítica, que el estudio identificó como el factor protector más claro, es a la vez la competencia que estas propuestas buscan cultivar en los tres planos. Hay que aclarar que estas medidas son recomendaciones fundadas en los hallazgos y en la literatura, pero que este trabajo no las puso a prueba empíricamente; contrastar su eficacia mediante diseños cuasiexperimentales es una de las líneas de investigación futura que se detallan más adelante.

9.4 Contribución del trabajo

Esta tesis aporta dos cosas. La primera es un instrumento de medición de la deuda cognitiva con validez preliminar, listo para replicarse y refinarse en estudios de mayor escala. La segunda es evidencia empírica sobre un fenómeno que la literatura teórica venía describiendo con creciente preocupación, pero que hasta ahora carecía de una operacionalización sistemática en el contexto educativo argentino. El trabajo establece que la deuda cognitiva es medible, que está distribuida de manera no uniforme en la población estudiantil, que responde a variables contextuales identificables y que puede ser atenuada mediante intervenciones sobre la conducta de revisión. Esas cuatro afirmaciones, juntas, ofrecen una base empírica suficiente para fundamentar tanto investigaciones futuras como orientaciones pedagógicas concretas.

Para cerrar, una proyección que se desprende del marco teórico y que hay que formular con cautela. Si la paradoja de la confianza en la IA introducida en la sección 5.1 operara de manera sostenida, la relación observada aquí entre intensidad de uso y deuda cognitiva no se debilitaría con el avance tecnológico; al contrario, se reforzaría. A medida que los modelos se vuelven más capaces y confiables, la confianza depositada en ellos y la creencia en su superioridad podrían aumentar, y con ellas la disposición a delegar el esfuerzo cognitivo propio. Extrapolada en el tiempo, esta dinámica podría derivar en una dependencia cada vez más profunda y difícil de revertir. Importa subrayar que esto es una hipótesis interpretativa y no un hallazgo del presente trabajo: el diseño transversal empleado describe un momento puntual y no permite medir trayectorias en el tiempo.

Justamente por eso, contrastar esta proyección mediante estudios longitudinales constituye una de las líneas de investigación futura más relevantes que abre esta tesis. Lo que el presente estudio sí establece es el punto de partida: existe una asociación medible y significativa entre el uso de IA y la deuda cognitiva, sobre la cual esa dinámica futura, de confirmarse, operaría.

10. Próximos pasos para una investigación futura

Esta tesis fue diseñada como una prueba de concepto: su valor reside menos en la precisión predictiva de sus modelos que en la plataforma metodológica que construye. Los resultados obtenidos abren un programa de investigación cuya extensión natural se describe a continuación en cuatro ejes.

10.1 Ampliación y mejora del instrumento

El paso metodológico más inmediato y necesario es la expansión de la muestra a varios cientos de participantes, obtenidos mediante un muestreo estratificado que garantice representatividad por institución, disciplina, año cursado y región geográfica. Con una muestra de ese tamaño se vuelve viable el análisis factorial exploratorio y confirmatorio, que permitirá evaluar si las seis dimensiones teóricas propuestas emergen efectivamente como factores diferenciados en los datos, o si el instrumento requiere reorganización dimensional.

La reformulación del ítem sobre el uso de la IA como borrador es prioritaria. Una estrategia es desdoblarlo en dos ítems separados: uno que evalúe la delegación inicial (“Cuando tengo que escribir un trabajo, le pido a la IA que genere el primer texto”) y otro que evalúe el involucramiento posterior (“Cuando uso la IA para un primer borrador, reescribo sustancialmente lo que generó antes de entregarlo”). Esta separación permitiría capturar la dimensión protectora de la reelaboración sin contaminación del componente de delegación.

Una línea complementaria de mejora instrumental es el desarrollo de una versión breve del cuestionario, de cinco a siete ítems, orientada al tamizaje rápido en contextos institucionales. El análisis de correlación ítem-total (sección 8.6) ya

identificó qué preguntas concentran el mayor poder para discriminar entre niveles de deuda, la primera reacción de consultar la IA ante una tarea compleja, la frecuencia de uso académico y la percepción de que el rendimiento bajaría sin IA, lo que vuelve viable una versión abreviada sin una pérdida sustantiva de capacidad discriminativa.

10.2 Diseño longitudinal

El diseño transversal de esta investigación no permite responder si la deuda cognitiva aumenta a medida que el estudiante avanza en su trayectoria universitaria, ni si intervenciones pedagógicas específicas logran reducirla en el tiempo. Un diseño longitudinal con seguimiento de cohortes a lo largo de dos o tres años respondería ambas preguntas. Este diseño también permitiría distinguir con mayor precisión entre la relación causal (el uso de IA genera deuda cognitiva) y la relación de selección (perfiles cognitivos previos condicionan tanto el uso de IA como la tendencia a la dependencia), una distinción que el diseño transversal no puede resolver.

Un diseño cuasiexperimental, en el que grupos de estudiantes expuestos a intervenciones pedagógicas centradas en metacognición y revisión crítica se comparan con grupos de control, sería especialmente valioso. Si el hallazgo de esta tesis sobre el factor protector de la revisión se confirma en condiciones controladas, se dispondría de evidencia causal sólida para fundamentar programas de formación específicos a escala institucional.

10.3 Extensión a nuevas poblaciones y contextos

La muestra de esta tesis está compuesta casi exclusivamente por estudiantes de grado de universidades argentinas. Investigaciones futuras deberían extender el instrumento a estudiantes de posgrado, donde la profundidad del uso de IA y las exigencias de producción autónoma son distintas; a docentes y tutores, cuya percepción del fenómeno en sus estudiantes constituye una fuente de validez externa del instrumento; y a contextos latinoamericanos más amplios, lo que

permitiría evaluar si las diferencias entre perfiles institucionales y disciplinares observadas en Argentina se replican en otros sistemas educativos de la región.

La incorporación de variables adicionales podría enriquecer sustancialmente el modelo explicativo: el año cursado, la experiencia previa con tecnología, el rendimiento académico autoreportado, la frecuencia de lectura autónoma y las estrategias de estudio habituales son candidatos naturales. Variables institucionales como la política de la universidad respecto al uso de IA, la formación de los docentes en herramientas de IA y la existencia de criterios de evaluación explícitos sobre el tema podrían capturarse mediante preguntas adicionales al cuestionario o mediante datos administrativos institucionales.

10.4 Modelado avanzado

Con una muestra de varios cientos de observaciones, la transición desde la regresión lineal hacia métodos de ensamble como *Random Forest* o Gradient Boosting (Breiman, 2001) se vuelve metodológicamente viable y estadísticamente justificada. Estos modelos, construidos sobre la misma base conceptual y los mismos criterios de codificación establecidos en esta tesis, producirían estimaciones más robustas de la importancia de variables y mejorarían sustancialmente la capacidad predictiva.

Una segunda línea de modelado avanzado implica el análisis de redes de ítems, que representaría las correlaciones entre dimensiones como una red de nodos y permitiría identificar cuáles son los ítems más centrales, con mayor influencia sobre el resto del constructo, y cuáles son más periféricos. Esta aproximación, desarrollada en psicometría de redes (Borsboom y Cramer, 2013), complementa al análisis factorial y puede revelar la arquitectura interna del constructo de una manera que el alfa de Cronbach no permite.

10.5 Diferenciación entre modelos y proveedores de IA

Una limitación del presente trabajo, asumida de manera deliberada, es que trata a la Inteligencia Artificial generativa como una categoría única. El instrumento pregunta por la frecuencia y la forma de uso de la IA sin distinguir entre los

distintos productos disponibles, como ChatGPT, Claude, Gemini o Copilot, ni entre sus diferentes versiones. Esta decisión es coherente con el foco de la investigación, que está puesto en el uso humano y en la dependencia que ese uso genera, y no en las características técnicas de cada sistema. A los fines de medir la deuda cognitiva, lo relevante es la conducta de delegación del estudiante, con independencia de qué herramienta utilice para delegar.

No obstante, agrupar a todos los modelos bajo una misma etiqueta es, necesariamente, una simplificación. Los distintos productos presentan estilos de respuesta, niveles de calidad y formas de interacción notablemente diferentes entre sí, e incluso un mismo producto varía de manera sensible entre versiones. Una línea de investigación futura especialmente prometedora consiste en diferenciar el análisis según el modelo y el proveedor utilizados, para evaluar si la deuda cognitiva varía en función de la herramienta concreta y no solo de la intensidad de uso.

Esta diferenciación es relevante porque existe evidencia de que la calidad del modelo podría intensificar la dependencia en lugar de atenuarla. La literatura describe un fenómeno conocido como paradoja de la confianza en la IA: a medida que los sistemas se vuelven más capaces y sus respuestas más fluidas y convincentes, los usuarios tienden a confiar más en ellos y a verificar menos, lo que vuelve los errores más difíciles de detectar. Una amplia revisión sobre la sobredependencia en modelos de lenguaje resume esta paradoja al señalar que, cuanto más capaz es un modelo, más difícil resulta advertir cuándo se equivoca, porque su fluidez y su tono seguro proyectan una autoridad que desalienta la evaluación crítica (Ibrahim et al., 2025). Trasladada al plano de la deuda cognitiva, esta paradoja sugiere una hipótesis contraintuitiva que merece ser investigada: cuanto mejor es el modelo, mayor podría ser la dependencia que genera, ya que un asistente más competente reduce con más eficacia la necesidad percibida de esfuerzo propio.

Un diseño futuro podría incorporar al instrumento preguntas sobre cuál es el producto de uso principal del estudiante y con qué frecuencia utiliza cada uno, y luego comparar los niveles de deuda cognitiva entre subgrupos. Esto permitiría

contrastar empíricamente la hipótesis anterior y avanzar hacia una comprensión más matizada del fenómeno, en la que no toda la IA generativa se considere equivalente, sino que se reconozca que distintas herramientas pueden inducir distintos grados y formas de dependencia cognitiva.

11. Percepción personal

Lo que más me llamó la atención durante el proceso no fue la utilidad de las herramientas, eso era esperable y se confirmó ampliamente, sino la velocidad con la que cambia la relación con el esfuerzo cognitivo cuando se tiene acceso a ellas de manera sostenida. Al realizar una tarea compleja, sin lugar a dudas está claro que requiere un alto esfuerzo cognitivo, representa tiempo y esfuerzo. Con el tiempo, y a medida que la asistencia tecnológica se volvió más habitual, la tolerancia a esa lentitud se fue reduciendo. El umbral de lo que se percibe como "demasiado esfuerzo" se movió, y eso ocurrió de manera imperceptible, sin una decisión consciente al respecto.

Ese desplazamiento del umbral es, creo, el núcleo experiencial de la deuda cognitiva. No se trata de que la persona deje de poder pensar: en términos de capacidad, las funciones cognitivas permanecen intactas. Se trata de algo más sutil y más difícil de observar desde dentro: la disposición a activarlas autónomamente, sin asistencia tecnológica, se erosiona. No de golpe, sino por acumulación de pequeñas decisiones de delegación que individualmente parecen inocuas y colectivamente reconfiguran los hábitos de trabajo intelectual.

Lo que más me sorprendió de los datos es que el factor protector de la revisión resultó estadísticamente significativo incluso dentro del grupo de usuarios más intensivos. Esto sugiere que la variable crítica no es tanto la frecuencia de uso como la calidad del involucramiento posterior. Eso es, al mismo tiempo, una buena y una mala noticia. La buena es que el uso frecuente de IA no condena inevitablemente a la dependencia cognitiva: la revisión crítica puede compensarlo en parte. La mala es que revisar críticamente requiere exactamente el tipo de

esfuerzo que el uso sin revisión tiende a erosionar. Hay un círculo que puede cerrarse en sentido negativo con notable facilidad.

La deuda cognitiva, al final, no es solo un riesgo para los estudiantes universitarios: es un riesgo para cualquiera que use herramientas cognitivas potentes sin reflexionar sobre qué procesos está delegando y qué procesos está preservando. Este trabajo fue, en ese sentido, también un ejercicio de autoobservación. No tengo la certeza de haber salido de él sin deuda propia. Pero sí tengo la certeza de ser más consciente de cuándo la estoy acumulando.

Bibliografía

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). Taxonomy of educational objectives: Handbook I. Cognitive domain. David McKay.

Borsboom, D., & Cramer, A. O. J. (2013). Network analysis: An integrative approach to the structure of psychopathology. Annual Review of Clinical Psychology, 9, 91-121.

Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5-32.

Brynjolfsson, E., & McAfee, A. (2024). The coming AI economic revolution. Foreign Affairs. https://www.foreignaffairs.com/world/coming-ai-economic-revolution

Casillas Alvarado, M., Ramírez Martinell, A., & López-Olivera López, R. (2024). Periods of technological change in higher education. Journal of Comparative & International Higher Education, 16(2), 136-151. https://doi.org/10.32674/jcihe.v16i2.5920

Clark, A., & Chalmers, D. (1998). The extended mind. Analysis, 58(1), 7-19.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. Psychometrika, 16(3), 297-334.

Fan, L., Deng, K., & Liu, F. (2025). Educational impacts of generative artificial intelligence on learning and performance of engineering students in China. Scientific Reports, 15(1), Article 26521. https://doi.org/10.1038/s41598-025-06930-w

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

Feng, J., Yu, B., Tan, W. H., Dai, Z., & Li, Z. (2025). Key factors influencing educational technology adoption in higher education: A systematic review. PLOS Digital Health, 4(4), Article e0000764. https://doi.org/10.1371/journal.pdig.0000764

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry. American Psychologist, 34(10), 906-911.

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

Guzman, A. L., & Lewis, S. C. (2020). Artificial intelligence and communication: A Human-Machine Communication research agenda. New Media & Society, 22(1), 70-86. https://doi.org/10.1177/1461444819858691

Holmes, W., Bialik, M., & Fadel, C. (2019). Artificial intelligence in education: Promises and implications for teaching and learning. Center for Curriculum Redesign.

Ibrahim, L., Collins, K. M., Kim, S. S. Y., Reuel, A., Lamparth, M., Feng, K., Ahmad, L., Soni, P., El Kattan, A., Stein, M., Swaroop, S., Padmakumar, V., Sucholutsky, I., Strait, A., Yang, D., Liao, Q. V., & Bhatt, U. (2025). Measuring and mitigating overreliance is necessary for building human-compatible AI. arXiv. https://doi.org/10.48550/arXiv.2509.08010

Kahneman, D. (2011). Thinking, fast and slow. Farrar, Straus and Giroux.

Kosmyna, N., Hauptmann, E., Yuan, Y. T., Situ, J., Liao, X.-H., Beresnitzky, A. V., Braunstein, I., & Maes, P. (2025). Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing task. arXiv. https://doi.org/10.48550/arXiv.2506.08872

Laban, P., Hayashi, H., Zhou, Y., & Neville, J. (2025). LLMs get lost in multi-turn conversation. arXiv preprint arXiv:2505.06120. ICLR 2026 Outstanding Paper Award.

Lee, H.-P., Sarkar, A., Tankelevitch, L., Drosos, I., Rintel, S., Banks, R., & Wilson, N. (2025). 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 (CHI '25). ACM. https://doi.org/10.1145/3706598.3713778

Leviathan, Y., Kalman, M., & Matias, Y. (2025). Prompt repetition improves non-reasoning LLMs. arXiv preprint arXiv:2512.14982.

López, C., Balmaceda, T., Zeller, M., Peller, J., Aguerre, C., & Tagliazucchi, E. (2024). Ok, Pandora: Seis ensayos sobre inteligencia artificial. El Gato y La Caja.

Lubbe, A., Marais, E., & Kruger, D. (2025). Cultivating independent thinkers: The triad of artificial intelligence, Bloom's taxonomy and critical thinking in assessment pedagogy. Education and Information Technologies, 30(13), 17589-17622. https://doi.org/10.1007/s10639-025-13476-x

Luckin, R. (2018). Machine learning and human intelligence: The future of education for the 21st century. UCL Institute of Education Press.

Monarch, R. M. (2021). Human-in-the-Loop Machine Learning: Active learning and annotation for human-centered AI. Manning Publications.

Ncube, P. D. N., Dzvapatsva, G., Matobobo, C., & Ranga, M. M. (2025). Redefining student assessment in AI-infused learning environments: A systematic review of challenges and strategies for academic integrity. AI and Ethics, 6, 68. https://doi.org/10.1007/s43681-025-00871-w

Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. Trends in Cognitive Sciences, 20(9), 676-688.

Selwyn, N. (2019). Should robots replace teachers? AI and the future of education. Polity Press.

Sparrow, B., Liu, J., & Wegner, D. M. (2011). Google effects on memory: Cognitive consequences of having information at our fingertips. Science, 333(6043), 776-778.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. Learning and Instruction, 4(4), 295-312.

UNESCO. (2021). AI and education: Guidance for policy-makers. UNESCO Publishing.

UNESCO. (2023). Technology in education: Global Education Monitoring Report 2023. UNESCO Publishing.

UNESCO. (2024). AI competency framework for students. UNESCO Publishing.

UNESCO. (2025). UNESCO survey: Two-thirds of higher education institutions have or are developing guidance on AI use. UNESCO.

Valverde-Berrocoso, J., Acevedo-Borrega, J., & Cerezo-Pizarro, M. (2022). Educational technology and student performance: A systematic review. Frontiers in Education, 7, Article 916502. https://doi.org/10.3389/feduc.2022.916502

Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. Harvard University Press.

Xu, J., Li, G., & Jiang, J. Y. (2025). AI self-preferencing in algorithmic hiring: Empirical evidence and insights. Manuscript submitted to Manufacturing & Service Operations Management. SSRN. https://ssrn.com/abstract=5417394

Anexo A: Cuestionario aplicado

A continuación se reproduce el instrumento tal como fue aplicado a los estudiantes a través de formulario digital. El cuestionario fue precedido por un consentimiento informado que explicaba el propósito del estudio, el carácter anónimo y voluntario de la participación, y el uso exclusivamente académico de las respuestas.

Pregunta 1. ¿En qué universidad estás cursando actualmente?

[Respuesta abierta]

Pregunta 2. En el ámbito académico: ¿con qué frecuencia usaste herramientas de IA para tareas o estudios universitarios durante el último mes?

Opciones: Nunca / Algunas veces por mes / Algunas veces por semana / Todos los días / Varias veces al día

Pregunta 3. Fuera del ámbito académico (trabajo, uso personal, entretenimiento, etc.): ¿con qué frecuencia usás herramientas de IA en tu vida cotidiana?

Opciones: Nunca / Algunas veces por mes / Algunas veces por semana / Todos los días / Varias veces al día

Las siguientes preguntas utilizan escala Likert de 5 puntos: 1 = Totalmente en desacuerdo, 2 = En desacuerdo, 3 = Ni de acuerdo ni en desacuerdo, 4 = De acuerdo, 5 = Totalmente de acuerdo.

Pregunta 4. Cuando me enfrento a una tarea académica compleja, mi primera reacción es consultar una IA antes de intentar resolverla por mi cuenta.

[Dimensión: Esfuerzo mental · Codificación: directa]

Pregunta 5. En vez de utilizar buscadores clásicos como Google, utilizo la IA como primera opción a la hora de resolver una actividad académica.

[Dimensión: Esfuerzo mental · Codificación: directa]

Pregunta 6. Siento que sin acceso a herramientas de IA mi rendimiento académico bajaría de manera significativa.

[Dimensión: Dependencia tecnológica · Codificación: directa]

Pregunta 7. Cuando una IA no me da una respuesta útil o no está disponible, me siento bloqueado para avanzar con la tarea.

[Dimensión: Dependencia tecnológica · Codificación: directa]

Pregunta 8. Cuando estudio o resuelvo trabajos prácticos sin usar IA, me lleva mucho más tiempo o lo siento más complicado.

[Dimensión: Dependencia tecnológica · Codificación: directa]

Pregunta 9. Me siento capaz de redactar un trabajo académico extenso (más de 1500 palabras) sin asistencia de IA.

[Dimensión: Autonomía cognitiva · Codificación: INVERSA (6 - valor)]

Pregunta 10. Cuando leo un texto académico difícil, prefiero pedirle a la IA un resumen antes que leerlo completo.

[Dimensión: Autonomía cognitiva · Codificación: directa]

Pregunta 11. Releo críticamente lo que la IA me genera buscando errores conceptuales, datos incorrectos o inconsistencias lógicas antes de usarlo.

[Dimensión: Conducta de revisión · Codificación: INVERSA (6 - valor)]

Pregunta 12. Cuando uso una respuesta generada por IA, suelo modificarla, reescribirla o complementarla con ideas propias antes de entregar el trabajo.

[Dimensión: Conducta de revisión · Codificación: INVERSA (6 - valor)]

Pregunta 13. Utilizo la IA para que genere un primer borrador del trabajo y luego trabajo sobre él.

[Dimensión: Conducta de revisión · Codificación: INVERSA (6 - valor)]

Pregunta 14. Después de varias horas usando IA para estudiar, me cuesta volver a concentrarme en tareas que requieran razonamiento propio.

[Dimensión: Fatiga cognitiva · Codificación: directa]