

Reseña de

Your Brain on ChatGPT: Accumulation of Cognitive Debt when Using an AI Assistant for Essay Writing Task

Por David Carmona Barrales

Ficha técnica:

Título: *Your Brain on ChatGPT: Accumulation of Cognitive Debt when Using an AI Assistant for Essay Writing Task.*

Autor: Nataliya Kosmyna, Eugene Hauptmann, Ye Tong Yuan, Jessica Situ, Xian-Hao Liao, Ashly Vivian Beresnitzky, Iris Braunstein y Pattie Maes.

Público objetivo: Académico y de investigación. Específicamente, investigadores en los campos de la neurociencia cognitiva, la interacción humano-computadora (HCI), la psicología de la educación y la tecnología educativa.

Tipo: Artículo de investigación científica (Paper).

Institución: Principalmente, el MIT Media Lab, con colaboraciones del MIT, Wellesley College y Massachusetts College of Art and Design (MassArt).

Fecha de Publicación: 2025.

Extensión: 206 páginas (incluye apéndices extensos).

Enlace al documento: <https://arxiv.org/abs/2506.08872>

Contexto:

Desde la popularización masiva de los Modelos Lingüísticos Grandes (LLM), como ChatGPT, a finales de 2022, ha existido una tensión palpable entre su potencial como asistentes democratizadores del conocimiento y la preocupación por sus efectos en el desarrollo del pensamiento crítico, la creatividad y la autonomía intelectual.

Los autores, un equipo interdisciplinario liderado por Nataliya Kosmyna desde el prestigioso MIT Media Lab, una institución reconocida por su investigación de vanguardia en la intersección de la tecnología y la experiencia humana, abordan esta controversia no desde la especulación teórica, sino a través de una investigación empírica. El estudio se posiciona como una de las primeras investigaciones longitudinales y multimodales que busca medir objetivamente el "coste cognitivo" del uso de la IA, respondiendo a una necesidad urgente de la comunidad educativa y científica por obtener datos concretos sobre este fenómeno transformador.

Exposición y Reconstrucción del Contenido:

La tesis principal del documento es que el uso de asistentes de IA como ChatGPT para tareas cognitivamente complejas, como la escritura de ensayos, conduce a la acumulación de una "**deuda cognitiva**". Esta deuda se manifiesta en una reducción medible del compromiso neuronal, un deterioro en la codificación de la memoria y un menor sentido de autoría intelectual en comparación con métodos más tradicionales como el uso de motores de búsqueda o el trabajo sin herramientas externas.

Para sustentar esta afirmación, los autores diseñaron un estudio longitudinal de cuatro meses con 54 participantes de universidades de élite de Boston. Estos fueron divididos en tres grupos:

1. **Grupo LLM:** Utilizaban exclusivamente ChatGPT para escribir sus ensayos.
2. **Grupo Search Engine:** Podían usar cualquier motor de búsqueda, pero no LLMs.
3. **Grupo Brain-only:** No podían utilizar ninguna herramienta externa, dependiendo únicamente de su propio conocimiento.

A lo largo de tres sesiones, los participantes escribieron ensayos de 20 minutos sobre temas estandarizados (tipo SAT¹). Una cuarta sesión opcional invirtió los roles de los grupos LLM y Brain-only para observar los efectos del cambio de herramienta. La investigación se basó en un enfoque multimodal que triangulaba datos de tres fuentes:

- **Análisis de Neuroimagen (EEG):** Se midió la actividad cerebral para evaluar la conectividad neuronal y la carga cognitiva a través de la métrica dDTF (Dynamic Directed Transfer Function). Los resultados mostraron una jerarquía clara: el grupo Brain-only exhibió la conectividad neuronal más fuerte y extensa en todas las bandas de frecuencia (alfa, beta, theta, delta), seguido por el grupo Search Engine, y finalmente el grupo LLM, que mostró la menor actividad neuronal.
- **Análisis Lingüístico (NLP):** Se analizaron los ensayos resultantes en busca de patrones. Los textos del grupo LLM mostraron una notable homogeneidad lingüística (en n-gramas, entidades nombradas y ontologías), sugiriendo una convergencia hacia un estilo "estándar" dictado por la IA, en contraste con la mayor variabilidad y originalidad del grupo Brain-only .
- **Análisis Cualitativo (Entrevistas):** Las entrevistas post-sesión revelaron diferencias conductuales significativas. La más notable fue la "*capacidad de citar*": los miembros del grupo LLM tuvieron una dificultad abrumadora para recordar y citar frases de los ensayos que acababan de escribir, un fallo que no se observó en los otros grupos. Asimismo, su "*sentido de propiedad*" o autoría sobre el texto era consistentemente más bajo.

¹ Se refiere al examen SAT (Scholastic Assessment Test) de Estados Unidos, que evalúa la capacidad de estudiantes para analizar un texto argumentativo y redactar una respuesta, también argumental.



Análisis crítico:

Fortalezas:

- **Metodología robusta:** La combinación de EEG, NLP y entrevistas cualitativas proporciona una visión holística y convincente del fenómeno. En lugar de depender de una única métrica, los autores triangulan sus hallazgos, mostrando cómo los patrones de conectividad neuronal (EEG) se correlacionan con las características del texto producido (NLP) y con la experiencia subjetiva y la memoria del participante (entrevistas).
- **Diseño longitudinal con grupo de control invertido:** El estudio se desarrolla a lo largo de cuatro meses con múltiples sesiones. La cuarta sesión, dota al diseño de mayor valor metodológico. Demuestra que los participantes que habían dependido de la IA (LLM-to-Brain) mostraron una conectividad neuronal más débil y un peor rendimiento en memoria cuando se les privó de la herramienta, sugiriendo que el "músculo" cognitivo no se había ejercitado de la misma manera.
- **Relevancia y oportunidad:** La investigación aborda una de las preguntas más apremiantes de nuestro tiempo de una manera empírica y rigurosa. En un mar de opiniones y especulaciones, este trabajo aporta datos neurofisiológicos y conductuales concretos, convirtiéndose en una referencia indispensable para educadores, legisladores y desarrolladores de tecnología.
- **Sofisticación analítica:** El uso de técnicas avanzadas como la dDTF para analizar la direccionalidad del flujo de información en el cerebro o el PaCMAPI para visualizar clusters en los embeddings de los textos demuestra un altísimo nivel de competencia técnica, lo que refuerza la fiabilidad de los resultados.

Aspectos mejorables:

A pesar de sus notables fortalezas, el estudio presenta importantes limitaciones y áreas que merecen un escrutinio crítico profundo.

- **Generalizabilidad de la muestra:** La muestra de 54 participantes, aunque adecuada para un estudio con EEG, está compuesta exclusivamente por estudiantes de universidades de élite de Boston (MIT, Harvard, etc.). Este es un grupo demográfico con un capital cognitivo y educativo excepcionalmente alto. Es altamente cuestionable si los resultados se pueden extrapolar a la población estudiantil general, que presenta una diversidad mucho mayor en habilidades, antecedentes socioeconómicos y estilos de aprendizaje. El estudio podría estar midiendo el efecto de la IA en un "cerebro optimizado", y los efectos podrían ser drásticamente diferentes



(quizás incluso más pronunciados o, alternativamente, beneficiosos en ciertos aspectos) en otros contextos demográficos.

- *Artificialidad de la tarea y definición de "uso de IA"*: La tarea asignada —un ensayo de 20 minutos tipo SAT— es un escenario de alta presión y corto plazo que incentiva un uso superficial y extractivo de la IA (búsqueda de respuestas rápidas, casi un "copiar y pegar" mejorado). Esto no representa la totalidad de cómo se pueden integrar las herramientas de IA en procesos de aprendizaje más largos y reflexivos (ej. investigación para un trabajo exhaustivo, brainstorming para elaborar listas de chequeo, reestructuración de borradores). El estudio, por su diseño, penaliza el uso de la IA al constreñirlo a un paradigma que maximiza sus debilidades (generación pasiva de contenido) y minimiza sus fortalezas (asistencia iterativa y colaborativa).
- *Conceptualización de la "deuda cognitiva"*: El término "*deuda cognitiva*" es una metáfora poderosa pero conceptualmente difusa. El estudio la define operativamente por una menor conectividad neuronal y una pobre capacidad de citación. Sin embargo, no logra establecer de manera concluyente si este efecto es una atrofia cognitiva a largo plazo o simplemente un cambio de estrategia adaptativo y temporal. Me pregunto: *¿Es el cerebro volviéndose "más perezoso" o está externalizando eficientemente una función (la composición de texto a corto plazo) para, potencialmente, liberar recursos para otras tareas de nivel superior que el experimento no midió?* El documento asume una interpretación de "déficit", pero no explora a fondo interpretaciones alternativas basadas en la cognición extendida o la descarga cognitiva eficiente.
- *Interpretación de los datos de EEG*: Si bien el análisis es técnicamente bueno, la traducción de patrones de conectividad de bandas de frecuencia específicas (ej. alfa, theta) a constructos psicológicos complejos como "creatividad", "pensamiento crítico" o "compromiso profundo" es inherentemente una inferencia interpretativa. La literatura apoya estas conexiones, pero no son leyes deterministas. Una menor conectividad en el grupo LLM podría, desde otra perspectiva, interpretarse como una "eficiencia neuronal" para esa tarea específica (lograr un buen resultado con menos esfuerzo), aunque los datos conductuales (memoria, propiedad) del propio estudio argumentan en contra de esta visión optimista. Sin embargo, esta ambigüedad inherente a la neuroimagen funcional debería ser reconocida de forma más explícita.

Síntesis y Conclusión:

Su principal valor reside en ser uno de los primeros estudios en ofrecer evidencia empírica de que la dependencia de herramientas de IA para la escritura puede tener un coste cognitivo tangible. Sin embargo, la conclusión central de que *la aparente facilidad y eficiencia que proporcionan los LLMs puede enmascarar una reducción del esfuerzo*

neuronal, un peor anclaje del conocimiento en la memoria y una erosión de la capacidad intelectual es apresurada, dadas las importantes limitaciones que se reflejan en el apartado de *aspectos mejorables* de esta reseña, sobre todo en lo tocante a las inferencias interpretativas de los resultados.

A pesar de sus limitaciones, los resultados no deben ser ignorados y recomiendo la lectura crítica de este artículo a cualquier persona involucrada en la educación, la aplicación de la IA en el trabajo, el desarrollo de políticas tecnológicas o la investigación en neurociencias. La reflexión sobre los resultados del estudio. Más allá de las cuestionables inferencias de sus autores, puede ser de gran valor en muchos campos de aplicación de la IA. En el propio artículo no se aboga por una prohibición de la IA, sino que ilumina la necesidad de un enfoque pedagógico crítico que enseñe *cómo y cuándo* usar estas herramientas de manera que aumenten, y no sustituyan, el esfuerzo cognitivo fundamental para el aprendizaje.

Referencia del artículo (APA 7^a Ed.):

Kosmyna, N. et al. (2025). *Your brain on ChatGPT: Accumulation of cognitive debt when using an AI assistant for essay writing task*. [Preprint]. arXiv. <https://arxiv.org/abs/2506.08872>