

QUEM PENSA QUANDO A IA PENSA? A EROÇÃO DO CONHECIMENTO REAL ENTRE UNIVERSITÁRIOS NO USO DA INTELIGÊNCIA ARTIFICIAL GENERATIVA

WHO THINKS WHEN AI THINKS? THE EROSION OF REAL KNOWLEDGE AMONG UNIVERSITY STUDENTS IN THE USE OF GENERATIVE ARTIFICIAL INTELLIGENCE

¿QUIÉN PIENSA CUANDO LA IA PIENSA? LA EROSIÓN DEL CONOCIMIENTO REAL ENTRE LOS UNIVERSITARIOS EN EL USO DE LA INTELIGENCIA ARTIFICIAL GENERATIVA

Nathacha Ferreira Barbosa¹
Diêgo do Nascimento Morais²
Luiza Mércia Freire Corrêa³
Viviane D'Barsoles Gonçalves Werutsky⁴
Thiago Herbert Santos Oliveira⁵
Saul Teixeira de Araújo⁶

RESUMO: Esse artigo buscou investigar de que modo o uso intensivo da inteligência artificial generativa por estudantes universitários se relaciona com a erosão do conhecimento real, isto é, o esvaziamento progressivo das competências cognitivas que o ensino superior pretende formar. A metodologia consistiu em revisão bibliográfica de natureza qualitativa e exploratória, com levantamento de estudos empíricos publicados entre 2021 e 2026 em bases científicas, privilegiando pesquisas com mensuração neural, comportamental e de desempenho acadêmico. Os resultados convergem para um padrão consistente: a adoção da inteligência artificial entre universitários ultrapassa 90% em alguns países e 70% no Brasil, ao passo que estudos de eletroencefalografia, levantamentos de larga escala e análises de desempenho associam o uso dependente da ferramenta à redução da conectividade neural, ao descarregamento cognitivo e à queda em indicadores de pensamento crítico e de rendimento. Conclui-se que o problema não reside na ferramenta, mas no modo de uso: a inteligência artificial empregada como substituta do esforço cognitivo, e não como apoio, converte-se em vetor de erosão epistêmica, o que exige redesenho avaliativo e formação em literacia crítica.

Palavras-chave: Inteligência Artificial. Erosão Epistêmica. Pensamento Crítico.

¹Discente do curso de Letras na Unifatecie.

²Discente do curso de Filosofia na Unifatecie.

³Pesquisadora.

⁴Pesquisadora.

⁵Pesquisador.

⁶Pesquisador.

ABSTRACT: This article aimed to investigate how the intensive use of generative artificial intelligence by university students relates to the erosion of real knowledge, that is, the progressive emptying of the cognitive competencies that higher education aims to develop. The methodology consisted of a qualitative and exploratory bibliographic review, surveying empirical studies published between 2021 and 2026 in scientific databases, prioritizing research with neural, behavioral, and academic performance measurements. The results converge toward a consistent pattern: artificial intelligence adoption among university students exceeds 90% in some countries and 70% in Brazil, while electroencephalography studies, large-scale surveys, and performance analyses associate dependent use of the tool with reduced neural connectivity, cognitive offloading, and declines in critical thinking and achievement indicators. It is concluded that the problem lies not in the tool, but in the mode of use: artificial intelligence employed as a substitute for cognitive effort, rather than as support, becomes a vector of epistemic erosion, which requires assessment redesign and training in critical literacy.

Keywords: Artificial Intelligence. Epistemic Erosion. Critical Thinking.

RESUMEN: Este artículo buscó investigar de qué modo el uso intensivo de la inteligencia artificial generativa por estudiantes universitarios se relaciona con la erosión del conocimiento real, es decir, el vaciamiento progresivo de las competencias cognitivas que la educación superior pretende formar. La metodología consistió en una revisión bibliográfica de naturaleza cualitativa y exploratoria, con relevamiento de estudios empíricos publicados entre 2021 y 2026 en bases científicas, privilegiando investigaciones con medición neural, conductual y de rendimiento académico. Los resultados convergen hacia un patrón consistente: la adopción de la inteligencia artificial entre universitarios supera el 90% en algunos países y el 70% en Brasil, mientras que estudios de electroencefalografía, encuestas de gran escala y análisis de desempeño asocian el uso dependiente de la herramienta con la reducción de la conectividad neural, la descarga cognitiva y la caída en indicadores de pensamiento crítico y de rendimiento. Se concluye que el problema no reside en la herramienta, sino en el modo de uso.

Palabras clave: Inteligencia Artificial. Erosión Epistémica. Pensamiento Crítico.

INTRODUÇÃO

Em pouco mais de dois anos, a inteligência artificial generativa deixou de ser uma curiosidade tecnológica para se tornar parte da rotina cognitiva de quase todos os estudantes universitários. O lançamento público do ChatGPT, em novembro de 2022, inaugurou uma transição de hábitos cuja velocidade não encontra paralelo recente na história da educação. Onde antes havia a biblioteca, o caderno e a longa hesitação diante de uma página em branco, hoje há um assistente disponível a qualquer hora, capaz de redigir, resumir, argumentar e resolver em segundos aquilo que costumava demandar horas de esforço deliberado. A pergunta que dá título a este trabalho, quem pensa quando a inteligência artificial pensa, não é retórica: ela aponta para uma inquietação concreta sobre o que acontece com o conhecimento real do estudante quando o ato de pensar é, de forma crescente, delegado a uma máquina.

A dimensão do fenômeno é inequívoca. No Reino Unido, a proporção de estudantes universitários que declararam usar alguma ferramenta de inteligência artificial generativa saltou de 66% para 92% no intervalo de apenas doze meses, enquanto o uso específico para tarefas avaliativas avançou de 53% para 88% no mesmo período (Freeman , 2025). No Brasil, levantamento conduzido em 2024 indicou que sete em cada dez universitários já incorporavam a inteligência artificial à sua rotina de estudos, com aumento de dezoito pontos percentuais no uso frequente em relação ao ano anterior (AGÊNCIA BRASIL, 2024). Não se trata, portanto, de um comportamento marginal ou restrito a uma vanguarda tecnológica, mas de uma prática que se naturalizou no cotidiano acadêmico antes mesmo que as instituições pudessem compreendê-la.

O entusiasmo com os ganhos de eficiência, no entanto, convive com um desconforto crescente. Pesquisas recentes em neurociência, psicologia cognitiva e educação têm documentado que o uso dependente da inteligência artificial se associa a um fenômeno conhecido como descarregamento cognitivo, isto é, a transferência sistemática de tarefas mentais para sistemas externos, com a consequente redução do esforço cognitivo próprio (Gerlich, 2025). O risco é que aquilo que se ganha em conveniência se perca em capacidade, ou seja, que o estudante, ao terceirizar o pensar, deixe de desenvolver justamente as competências que a universidade existe para formar.

3

O problema ganha contornos mais nítidos quando se observa a distância entre o que a ferramenta entrega e o que o estudante efetivamente aprende. Um texto gerado por inteligência artificial pode ser correto, fluente e bem estruturado, e ainda assim não corresponder a nenhum ganho cognitivo para quem o submete como seu. Diferentemente de um livro lido ou de um problema resolvido com esforço, que deixam marcas duradouras na memória e na capacidade de raciocínio, o conteúdo gerado por máquina pode atravessar o estudante sem nele se fixar. Forma-se, assim, uma aparência de competência que não se sustenta quando a ferramenta é retirada, e é justamente essa dissociação entre o produto e o aprendizado que o presente artigo se propõe a examinar à luz das evidências disponíveis.

Para nomear esse risco, este artigo recorre a uma analogia provinda da própria ciência da computação. Em 2024, pesquisadores demonstraram na revista *Nature* que modelos de inteligência artificial treinados recursivamente sobre dados gerados por outros modelos sofrem um processo de degradação progressiva e irreversível, batizado de colapso de modelo, ou *model collapse* (Shumailov I, *et al.*, 2024). Privado do contato com a riqueza e a variabilidade dos dados

humanos originais, o modelo passa a se realimentar de suas próprias saídas, perde diversidade, amplifica os próprios erros e, geração após geração, produz resultados cada vez mais homogêneos e pobres. A hipótese heurística que organiza este trabalho é a de que algo estruturalmente análogo pode ocorrer no plano humano: o estudante que substitui o esforço de pensar pela ingestão passiva de conteúdo gerado por máquina tende a empobrecer seu próprio repertório cognitivo, num processo que aqui se denomina erosão epistêmica.

Cabe, desde já, uma ressalva metodológica importante: a analogia com o colapso de modelo é heurística, e não literal. O cérebro humano não é uma rede neural artificial, e o aprendizado não se reduz a um problema de distribuição estatística de dados. A força do paralelo está em iluminar um mecanismo, a degradação que advém de cortar o vínculo com a fonte original de variabilidade e esforço, e não em afirmar uma identidade entre os dois sistemas. É com essa cautela que se mobiliza o conceito ao longo do texto.

Convém ainda situar o fenômeno em perspectiva histórica, para evitar tanto o alarmismo quanto a ingenuidade. A humanidade já atravessou outras transições em que ferramentas externas reconfiguraram a cognição: a invenção da escrita, que Platão, pela voz de Sócrates no diálogo Fedro, temia que enfraquecesse a memória; a difusão da imprensa; a calculadora eletrônica; e, mais recentemente, os motores de busca, que transferiram para a nuvem boa parte daquilo que antes guardávamos na própria mente. Em cada uma dessas passagens, o ganho de alcance veio acompanhado da perda de certas capacidades, e o saldo dependeu menos da ferramenta em si do que do modo como as instituições educativas souberam incorporá-la. A inteligência artificial generativa, contudo, apresenta uma diferença qualitativa que percorre todo este artigo: enquanto as tecnologias anteriores externalizavam sobretudo o armazenamento e a recuperação da informação, a inteligência artificial externaliza a própria operação de pensar, ou seja, a síntese, a argumentação e a produção de sentido. É essa novidade que torna o presente debate mais agudo do que seus antecessores.

Diante desse cenário, o presente artigo tem por objetivo geral investigar a relação entre o uso intensivo da inteligência artificial generativa por estudantes universitários e a erosão do conhecimento real. Como objetivos específicos, busca-se: caracterizar a magnitude e os padrões de adoção da inteligência artificial no ensino superior; sistematizar as evidências empíricas recentes sobre os efeitos cognitivos do uso dependente da ferramenta; e discutir as condições sob as quais ela atua como apoio ao aprendizado ou, ao contrário, como vetor de seu esvaziamento. A lacuna que se pretende abordar é a escassez de sínteses que articulem, de modo

crítico, as evidências neurais, comportamentais e de desempenho dispersas em uma literatura ainda muito recente, conectando-as a um marco conceitual capaz de orientar a prática pedagógica.

Justifica-se a relevância do tema por sua dupla urgência. De um lado, a velocidade da adoção tornou o fenômeno incontornável, pois não cabe mais perguntar se os estudantes usarão a inteligência artificial, mas como a usarão e com que consequências. De outro, a janela para uma resposta formativa é estreita, uma vez que os hábitos cognitivos que se consolidam nos primeiros anos de contato com a ferramenta tendem a se cristalizar. Compreender agora os mecanismos da erosão epistêmica é condição para preservar aquilo que o ensino superior tem de mais valioso, que não é a transmissão de conteúdos, realizada pela máquina com desenvoltura, mas a formação de sujeitos capazes de pensar por conta própria.

Cabe esclarecer, por fim, a estrutura argumentativa que organiza o restante do texto. Após esta introdução, apresenta-se um referencial teórico que articula três conceitos fundamentais, quais sejam o descarregamento cognitivo, a teoria da carga cognitiva e o colapso de modelo, os quais fornecem as lentes pelas quais as evidências serão interpretadas. Em seguida, descreve-se o percurso metodológico da revisão. A seção de resultados e discussão, apresentada de forma integrada, percorre os três eixos de evidência, que são a escala da adoção, os efeitos neurais e cognitivos e os efeitos sobre o desempenho, antes de retomar a hipótese da erosão epistêmica e de discutir o uso da inteligência artificial como apoio legítimo ao pensar. As considerações finais sintetizam os achados, explicitam as limitações e apontam caminhos para a prática e para a investigação futura.

REFERENCIAL TEÓRICO

Descarregamento cognitivo e o efeito Google

O conceito que melhor ancora teoricamente a discussão sobre os efeitos da inteligência artificial na cognição é o de descarregamento cognitivo, definido por Risko e Gilbert (2016) como o uso de uma ação física ou mental para alterar as exigências de processamento de informação de uma tarefa, de modo a reduzir a demanda cognitiva. Anotar um número de telefone, programar um alarme ou recorrer a uma calculadora são exemplos cotidianos e, em princípio, adaptativos dessa estratégia. O ser humano sempre estendeu sua mente para o ambiente, e a escrita talvez seja o mais antigo e poderoso desses dispositivos de externalização.

O problema não é, portanto, o descarregamento em si, mas seus efeitos colaterais sobre a memória e a aprendizagem quando se torna sistemático. A pesquisa seminal de Sparrow e colaboradores (2011) cunhou a expressão efeito Google para descrever um fenômeno revelador: quando as pessoas esperam que uma informação permaneça disponível externamente, elas a memorizam pior, mas recordam melhor onde encontrá-la. A mente, racionalmente econômica, deixa de codificar aquilo que pode ser recuperado a qualquer momento. Aplicado ao estudo, o princípio é direto, pois, se o estudante espera acessar a resposta na internet mais tarde, tende a não se engajar em níveis profundos de codificação ao estudar, o que compromete a retenção de longo prazo do material.

O ponto decisivo, sublinhado pela literatura mais recente, é que a inteligência artificial generativa representa um salto qualitativo nesse processo. Se os motores de busca descarregavam o armazenamento e a recuperação de informações, e os celulares a memória de tarefas, a inteligência artificial generativa descarrega uma camada mais profunda, que é a própria síntese e o raciocínio. O que se externaliza já não é o que se guarda, mas o que se realiza, ou seja, não o conteúdo do pensamento, mas o ato de compor uma frase, articular um argumento ou estruturar uma ideia. É essa migração do armazenamento para a performance cognitiva que distingue o momento atual e que justifica a preocupação específica deste artigo.

6

Essa discussão dialoga com a tese da mente estendida, proposta por Clark e Chalmers, segundo a qual objetos externos ao corpo podem ser considerados parte da mente quando desempenham um papel ativo nos processos cognitivos. Sob essa perspectiva, integrar uma ferramenta ao próprio pensamento não é, por si só, um problema, e pode até ampliar as capacidades humanas. A questão crítica, contudo, reside na natureza dessa integração. Quando a ferramenta complementa e potencializa um processo cognitivo que permanece ativo no sujeito, a extensão é genuína e benéfica. Quando, ao contrário, ela substitui o processo, esvaziando o sujeito de sua participação, o que se tem não é uma mente estendida, mas uma mente terceirizada, dependente de um suporte sobre o qual não exerce domínio. A diferença entre uma e outra é precisamente a fronteira que separa o uso saudável da erosão epistêmica.

Teoria da carga cognitiva e a fricção produtiva da aprendizagem

Para compreender por que a supressão do esforço pode ser prejudicial, é útil recorrer à teoria da carga cognitiva, formulada originalmente por Sweller. Segundo essa perspectiva, a aprendizagem significativa exige um esforço mental que, longe de ser um obstáculo a eliminar,

é o próprio motor da construção de esquemas mentais duradouros na memória de longo prazo. O conceito de dificuldades desejáveis, desenvolvido na psicologia da aprendizagem, vai na mesma direção, pois certos obstáculos, como o esforço de recuperar uma informação, a luta para formular uma explicação e a hesitação diante de um problema mal estruturado, produzem aprendizagem mais robusta justamente porque são difíceis.

Disso decorre um corolário central para a tese aqui defendida: quando a inteligência artificial elimina a fricção produtiva da tarefa, ela pode estar suprimindo não um incômodo dispensável, mas o próprio mecanismo pelo qual o conhecimento se consolida. A facilidade com que a ferramenta entrega um texto pronto poupa o estudante do desconforto da página em branco e, com ele, da elaboração mental que essa página exigiria. O esforço economizado só se converte em ganho líquido se for redirecionado para tarefas cognitivas igualmente exigentes; do contrário, evapora, e com ele a aprendizagem que dele dependeria.

O colapso de modelo: a metáfora e seus limites

O terceiro pilar conceitual deste trabalho provém da ciência da computação. Em estudo publicado na Nature, Shumailov e colaboradores (2024) demonstraram que modelos generativos treinados de forma recursiva sobre dados produzidos por outros modelos sofrem o que denominaram colapso de modelo, ou seja, uma degradação progressiva em que o modelo, privado do contato com a riqueza dos dados humanos originais, ajusta-se cada vez mais às regularidades empobrecidas de sua própria produção. As primeiras gerações perdem a variabilidade das caudas da distribuição, isto é, os casos raros, as exceções e a diversidade, e as gerações seguintes podem entrar em colapso completo, produzindo saídas homogêneas, repetitivas e, no limite, sem sentido.

O mecanismo é hoje compreendido como uma quase inevitabilidade estatística, pois, ao gerar conteúdo, o modelo não recupera exemplos da memória, mas agrega padrões aprendidos para produzir a saída mais provável, o que reduz intrinsecamente a variância. Quando essa saída empobrecida realimenta o treinamento, o efeito se acumula. A imagem que daí emerge, a de um sistema que se degrada por se alimentar de si mesmo, em vez de se nutrir da fonte original, oferece uma metáfora potente para pensar o aprendizado humano sob dependência da inteligência artificial.

Convém, no entanto, demarcar com rigor os limites dessa transposição, sob pena de incorrer em determinismo. O cérebro humano não é uma rede neural artificial; a aprendizagem

não se reduz a um ajuste de distribuição estatística; e o estudante, diferentemente do modelo, possui agência sobre aquilo de que se alimenta. A analogia é heurística e serve para iluminar um mecanismo, que é a degradação advinda de romper o vínculo com a fonte de variabilidade e esforço, e não para postular uma identidade entre os dois sistemas. É essa, aliás, a diferença que abre espaço para a intervenção pedagógica, como se discutirá adiante. A metáfora alerta para um risco, mas não decreta um destino.

Aprendizagem profunda, aprendizagem superficial e a metacognição

Um quarto aporte conceitual, oriundo da pedagogia e da psicologia educacional, completa o quadro teórico e ajuda a precisar o que se perde na erosão epistêmica. A distinção entre aprendizagem profunda e aprendizagem superficial, consolidada na literatura sobre abordagens de estudo, opõe duas posturas diante do conhecimento. Na aprendizagem superficial, o estudante busca cumprir a tarefa com o mínimo de esforço, memorizando de forma mecânica e tratando o conteúdo como um conjunto de fragmentos desconexos a serem reproduzidos. Na aprendizagem profunda, ao contrário, ele procura compreender o sentido, relacionar o novo com o já sabido, questionar e integrar, com vistas à construção de um entendimento durável. A primeira produz respostas; a segunda produz conhecimento.

8

O uso substitutivo da inteligência artificial alinha-se de modo quase perfeito à aprendizagem superficial, pois oferece o atalho ideal para cumprir a tarefa sem compreender o conteúdo. Quando o objetivo se reduz a entregar o trabalho, e a ferramenta o entrega pronto, desaparece o incentivo para o esforço de integração que caracteriza a aprendizagem profunda. O perigo é que essa superficialidade, antes limitada por uma barreira de esforço, encontra agora um caminho sem atrito, o que pode torná-la a opção padrão para um número crescente de estudantes.

Conectado a essa distinção está o conceito de metacognição, entendido como a capacidade de monitorar e regular os próprios processos de pensamento. A literatura sobre os efeitos da inteligência artificial tem alertado para um fenômeno descrito como preguiça metacognitiva, no qual o estudante, ao delegar a tarefa, também abdica do monitoramento que normalmente exerceria sobre o próprio raciocínio. Sem esse monitoramento, perde-se não apenas a competência específica, mas a própria consciência de sua perda, o que torna a erosão particularmente insidiosa. O estudante não percebe que sabe menos, porque a ferramenta lhe

devolve, a cada interação, a sensação de competência. É essa ilusão de domínio, descolada do domínio efetivo, que talvez represente o efeito mais perverso do uso acrítico da tecnologia.

MÉTODOS

Este estudo caracteriza-se como uma revisão bibliográfica de natureza qualitativa, exploratória e descritiva, voltada à sistematização e à interpretação crítica de evidências empíricas sobre os efeitos cognitivos do uso da inteligência artificial generativa por estudantes do ensino superior. A opção pela revisão narrativa, e não sistemática em sentido estrito, justifica-se pela natureza ainda emergente e heterogênea do campo, no qual coexistem desenhos metodológicos muito distintos, que vão de experimentos com eletroencefalografia a levantamentos de larga escala, e que dificultam a padronização exigida por uma metanálise, mas se beneficiam de uma leitura integradora.

O levantamento foi conduzido nos meses iniciais de 2026 nas bases Scopus, Web of Science, PubMed, SciELO, arXiv e Google Acadêmico, além de repositórios institucionais como o do Instituto de Tecnologia de Massachusetts (MIT) e da Microsoft Research. Foram empregados descritores em português, inglês e espanhol, combinados por operadores booleanos, entre os quais inteligência artificial, pensamento crítico, descarregamento cognitivo, generative AI, cognitive offloading, academic performance e model collapse.

Adotaram-se como critérios de inclusão: publicações compreendidas entre 2021 e 2026, de modo a contemplar o período imediatamente anterior e posterior à popularização da inteligência artificial generativa; estudos empíricos com mensuração de variáveis cognitivas, comportamentais ou de desempenho acadêmico, bem como revisões e estudos teóricos de referência; e trabalhos voltados a estudantes ou a usuários adultos em contextos de aprendizagem e trabalho do conhecimento. Foram excluídos materiais de divulgação sem lastro empírico, textos meramente opinativos e publicações cujo foco se restringisse a aspectos técnicos da inteligência artificial sem relação com a cognição humana. Privilegiaram-se, na seleção final, os estudos com maior rigor metodológico e maior repercussão no debate acadêmico.

A análise do material seguiu a lógica da análise de conteúdo temática, organizando os achados em torno de três eixos: a magnitude da adoção da inteligência artificial no ensino superior; os efeitos neurais e cognitivos documentados; e os efeitos sobre o desempenho e a integridade acadêmica. Por se tratar de pesquisa de revisão, fundamentada exclusivamente em

dados secundários de domínio público, sem coleta direta junto a seres humanos, o estudo dispensa apreciação por Comitê de Ética em Pesquisa, nos termos da regulamentação aplicável a essa modalidade investigativa.

RESULTADOS E DISCUSSÃO

Por conveniência expositiva, e em conformidade com a possibilidade prevista nas normas editoriais, optou-se por apresentar de forma integrada os resultados e a discussão, organizando-os nos três eixos temáticos delineados na metodologia. Essa escolha permite que cada conjunto de evidências seja imediatamente confrontado com a literatura e com a hipótese da erosão epistêmica.

A naturalização da inteligência artificial no ensino superior

O primeiro achado a se destacar é a velocidade sem precedentes da adoção. Os números, longe de descreverem uma tendência incipiente, revelam uma transformação já consolidada. Levantamento conduzido pelo Higher Education Policy Institute (HEPI) com 1.041 estudantes de graduação registrou que a parcela de universitários usando ferramentas de inteligência artificial passou de 66% para 92% em apenas um ano, ao passo que o uso para tarefas avaliativas subiu de 53% para 88% (Freeman, 2025). O autor do relatório observa que mudanças de comportamento dessa magnitude em doze meses são praticamente inéditas na pesquisa educacional, o que sinaliza que a ferramenta se incorporou ao cotidiano acadêmico antes que as instituições tivessem tempo de respondê-la.

Esse padrão se repete em diferentes contextos nacionais. Pesquisa global do Digital Education Council, abrangendo mais de 3.800 estudantes em dezesseis países, apontou que 86% dos universitários usam inteligência artificial em seus estudos, sendo o ChatGPT a ferramenta mais citada, por cerca de dois terços dos respondentes (DIGITAL EDUCATION COUNCIL, 2024). No Brasil, o quadro não é diferente: dados divulgados pela Agência Brasil em 2024 indicaram que sete em cada dez estudantes universitários e futuros universitários já recorriam à ferramenta na rotina de estudos, com aumento de dezoito pontos percentuais no uso frequente em relação a 2023 (AGÊNCIA BRASIL, 2024). A Tabela 1 sintetiza os principais indicadores de adoção identificados na literatura.

Tabela 1 Indicadores de adoção da inteligência artificial generativa por estudantes, 2024 a 2025.

Indicador	Contexto e fonte	Valor
Uso de qualquer ferramenta de IA	Reino Unido. Freeman (2025)	92%
Uso de IA em tarefas avaliativas	Reino Unido. Freeman (2025)	88%
Uso de IA nos estudos	Global. Digital Education Council (2024)	86%
Uso de IA na rotina de estudos	Brasil. Agência Brasil (2024)	70%
ChatGPT como ferramenta preferida	Global. Digital Education Council (2024)	66%
Estudantes sem literacia suficiente em IA	Global. Digital Education Council (2024)	58%

Fonte: elaborada pelos autores, 2026; dados extraídos de Freeman (2025), Digital Education Council (2024) e Agência Brasil (2024).

Há, contudo, um dado especialmente revelador no levantamento global: embora 86% dos estudantes usem inteligência artificial, 58% admitem não possuir conhecimento e habilidades suficientes para empregá-la de modo adequado, e 48% sentem-se despreparados para um mercado de trabalho mediado por essas tecnologias (DIGITAL EDUCATION COUNCIL, 2024). Configura-se, assim, um paradoxo central: a adoção é massiva, mas a literacia é rasa. Usa-se intensamente uma ferramenta que não se compreende criticamente, e é precisamente nessa fenda, entre o uso onipresente e o domínio frágil, que a erosão epistêmica encontra terreno fértil.

Importa, ainda, qualificar não apenas quanto, mas como os estudantes usam a inteligência artificial, pois é o tipo de tarefa delegada que determina o impacto cognitivo. Os levantamentos disponíveis indicam que os usos mais frequentes incluem explicar conceitos, resumir artigos, sugerir ideias de pesquisa e, de modo crescente, gerar texto. Essa distribuição é instrutiva, pois alguns desses usos situam-se mais próximos do apoio legítimo, como pedir o esclarecimento de um conceito que será depois trabalhado, enquanto outros, como a geração de texto submetido como próprio, deslizam para a substituição. O mesmo levantamento que documenta a adoção massiva revela, portanto, uma zona cinzenta de práticas cuja classificação como apoio ou substituição depende de detalhes de uso que raramente são objeto de reflexão por parte do estudante. É essa indistinção prática, e não apenas a frequência do uso, que torna o fenômeno difícil de governar e que recomenda cautela na interpretação dos próprios números de adoção.

Evidências neurais do esvaziamento cognitivo

Se a escala da adoção delinea o problema, são os estudos de base neurocientífica que lhe conferem materialidade. O trabalho de maior repercussão nesse campo foi conduzido por pesquisadores do MIT Media Lab e divulgado em 2025 sob o título *Your Brain on ChatGPT*.

Cinquenta e quatro participantes foram divididos em três grupos, um que escrevia redações com auxílio de um modelo de linguagem, outro que usava motor de busca e um terceiro que escrevia apenas com os próprios recursos cognitivos, e tiveram sua atividade cerebral monitorada por eletroencefalografia ao longo de quatro sessões distribuídas por vários meses (Kosmyna, *et al.*, 2025).

Os resultados foram tão consistentes quanto inquietantes. A conectividade cerebral mostrou-se escalonada de forma inversa ao grau de auxílio externo: o grupo que escrevia sem ferramentas exibiu as redes neurais mais fortes e distribuídas; o grupo do motor de busca, um engajamento intermediário; e o grupo do modelo de linguagem, a conectividade mais fraca de todas (KOSMYNA N, *et al.*, 2025). Em outras palavras, quanto mais a tarefa cognitiva era delegada à máquina, menos o cérebro se engajava. O achado de maior eloquência, porém, foi de ordem comportamental: 83% dos usuários do modelo de linguagem foram incapazes de citar uma única frase das redações que acabavam de escrever, ao passo que praticamente a totalidade dos demais grupos recordava o próprio trabalho com facilidade (ARMITAGE R, 2025). O texto havia sido produzido, mas não havia sido aprendido, pois passara pelas mãos do estudante sem deixar rastro na memória.

Talvez o aspecto mais preocupante do estudo resida na quarta sessão, quando os grupos foram trocados. Os participantes que haviam dependido do modelo de linguagem e passaram a escrever sem auxílio mantiveram uma conectividade neural reduzida e um subengajamento das ondas alfa e beta, o que sugere que o hábito de delegar deixara um rastro que não se desfez de imediato (Kosmyna, *et al.*, 2025). É justamente esse resíduo que os autores nomearam de dívida cognitiva, ou seja, o custo, pago posteriormente, de não ter exercitado o pensamento no momento devido. A imagem é poderosa e dialoga diretamente com a hipótese deste artigo, pois, assim como o modelo que se realimenta de dados sintéticos perde a capacidade de gerar diversidade, o estudante que se acostuma a não pensar parece perder, ao menos transitoriamente, a prontidão para fazê-lo.

É necessário, por honestidade intelectual, registrar as limitações desse estudo, reconhecidas pelos próprios autores e por comentadores. A amostra é pequena, a técnica de eletroencefalografia não acessa estruturas cerebrais profundas, e o trabalho, à época de sua divulgação, ainda não havia passado por revisão por pares (Stanković, *et al.*, 2026). Trata-se, portanto, de um achado direcionalmente importante, que demanda replicação em maior escala,

e não de um veredito clínico definitivo. Ainda assim, sua convergência com outras linhas de evidência confere-lhe peso que nenhuma limitação isolada anula.

As implicações educacionais desse conjunto de achados são profundas e merecem ser explicitadas. Se a conectividade neural escalona inversamente ao grau de auxílio externo, então a escrita assistida por inteligência artificial, embora produza textos de qualidade aparente superior, pode estar formando estudantes que não internalizam o conteúdo que assinam. O produto melhora enquanto o processo se esvazia, e essa dissociação entre o artefato e a aprendizagem é precisamente o que torna o fenômeno difícil de detectar pelos instrumentos avaliativos tradicionais, que medem o produto. Um professor que avalia apenas a redação entregue não tem como perceber que ela não deixou rastro na memória de quem a produziu. O estudo do MIT, ao revelar essa fenda entre o texto e o aprendiz, oferece um argumento empírico poderoso a favor de avaliações centradas no processo, e não no produto, tema ao qual se retornará nas considerações finais.

Vale ainda explorar a metáfora econômica que dá nome ao conceito central do estudo, pois a expressão dívida cognitiva é mais precisa do que à primeira vista pode parecer. Toda dívida envolve um benefício imediato pago com juros no futuro, e é exatamente essa a estrutura do uso substitutivo da inteligência artificial. No presente, o estudante colhe a vantagem da tarefa concluída com pouco esforço; no futuro, paga o custo acumulado da competência que não desenvolveu e da memória que não formou. O agravante é que, tal como nas dívidas financeiras mal administradas, os juros são compostos, ou seja, cada delegação que dispensa o esforço torna o esforço seguinte mais penoso, num ciclo que reforça a dependência. É nesse acúmulo progressivo que a analogia com o colapso de modelo encontra sua expressão mais exata, pois em ambos os casos a degradação não resulta de um evento isolado, mas da repetição de um processo que, geração após geração ou tarefa após tarefa, afasta o sistema da fonte original de riqueza e o aprisiona em suas próprias limitações.

O descarregamento cognitivo e a queda do pensamento crítico

A segunda linha de evidência provém da psicologia cognitiva e da pesquisa por levantamento. O estudo de Gerlich (2025), publicado no periódico *Societies*, talvez seja o mais citado nesse domínio. Combinando questionários e entrevistas com 666 participantes de diferentes faixas etárias e níveis de escolaridade, o autor encontrou uma correlação negativa significativa entre o uso frequente de ferramentas de inteligência artificial e as habilidades de

pensamento crítico, mediada pelo aumento do descarregamento cognitivo (Gerlich, 2025). O mecanismo é descrito como uma espécie de preguiça cognitiva, isto é, uma diminuição da disposição para o pensamento profundo e reflexivo, que se instala quando o esforço mental é sistematicamente terceirizado.

Dois aspectos desse estudo merecem destaque para os fins deste artigo. O primeiro é que os participantes mais jovens exibiram simultaneamente maior dependência da inteligência artificial e menores escores de pensamento crítico, avaliados pela escala Halpern (Gerlich, 2025). O autor sugere, com a devida cautela, que os chamados nativos digitais, por nunca terem conhecido um mundo sem essas ferramentas, podem recorrer a elas de modo mais reflexo e menos deliberado. O segundo é que o nível de escolaridade funcionou como fator de proteção, pois hábitos consolidados de reflexão pareciam contrabalançar a conveniência digital. Esse achado é decisivo, uma vez que indica que a erosão não é destino inevitável, mas resultado de um modo de uso que a educação pode, em princípio, reorientar.

Convém sublinhar a ressalva que o próprio Gerlich faz: o estudo evidencia correlações, e não causalidade direta. É plausível que pessoas com menor inclinação ao pensamento crítico recorram mais à inteligência artificial, e não apenas o contrário. A relação, portanto, é provavelmente bidirecional e mediada por múltiplos fatores. Ainda assim, a força e a consistência da associação, replicada em outros estudos, tornam difícil descartá-la como mero artefato.

Corroborando esse quadro, pesquisa conduzida pela Microsoft Research em parceria com a Carnegie Mellon University, apresentada na conferência CHI 2025, examinou 319 trabalhadores do conhecimento e 936 exemplos reais de uso de inteligência artificial. O achado central é particularmente instrutivo: quanto maior a confiança depositada na ferramenta, menor o engajamento em pensamento crítico, com coeficiente de regressão negativo e estatisticamente robusto (Lee, *et al.*, 2025). Inversamente, quanto maior a autoconfiança do indivíduo nas próprias capacidades, mais ele se dispunha a avaliar criticamente as respostas geradas. O estudo também observou que usuários de inteligência artificial tendiam a produzir um conjunto menos diverso de resultados para uma mesma tarefa, uma homogeneização que ecoa, de modo perturbador, a perda de variabilidade característica do colapso de modelo.

Embora conduzido com trabalhadores e não com estudantes, esse estudo oferece uma lição diretamente transponível para o contexto acadêmico. Os autores observam que a natureza do pensamento crítico se desloca, sob o uso da inteligência artificial, da produção de conteúdo

para a verificação e a integração das respostas geradas. Em vez de redigir, o usuário passa a conferir; em vez de criar, a editar. Essa transição não é, em si, problemática, e pode até ser desejável em certos contextos profissionais. O risco específico para o ensino superior, contudo, é que a verificação competente das respostas de uma máquina pressupõe justamente o domínio que a delegação sistemática impede de formar. Em outras palavras, para usar a inteligência artificial criticamente é preciso saber aquilo que a ferramenta dispensa de aprender, o que configura um círculo que apenas a formação deliberada pode romper. O estudante que nunca redigiu um argumento por conta própria carece dos critérios para avaliar se o argumento que a máquina lhe oferece é bom.

A convergência entre esses estudos permite formular o mecanismo central da erosão epistêmica com alguma precisão. Ele não opera por um golpe súbito, mas por um deslocamento gradual: a inteligência artificial reduz o custo imediato de produzir, mas, ao fazê-lo, suprime as fricções produtivas, ou seja, a hesitação diante do problema, a busca pela palavra exata e o esforço de articular um argumento, que são precisamente os momentos em que o aprendizado se consolida. Como observam diversos autores, o descarregamento cognitivo só é benigno quando o esforço mental liberado é redirecionado para tarefas igualmente exigentes; quando simplesmente desaparece, o resultado é a aprendizagem superficial e a atrofia das competências (Favero, *et al.*, 2025).

Um avanço metodológico recente ajuda a responder à principal fragilidade dessa literatura, que é seu caráter predominantemente correlacional. Boa parte das evidências discutidas até aqui, incluindo os estudos de Gerlich e de Lee e colaboradores, baseia-se em levantamentos e entrevistas, que estabelecem associações, mas não permitem inferir causalidade. Em resposta a essa lacuna, ensaios controlados randomizados de larga escala têm começado a oferecer evidência causal direta. O próprio desenho experimental de Kosmyna e colaboradores (2025), ao reatribuir os participantes a condições distintas na quarta sessão, aproxima-se dessa lógica e documenta que aqueles que se habituaram à ferramenta desempenham-se pior quando ela é retirada do que desempenhariam caso nunca a tivessem utilizado. Esse tipo de achado fortalece a hipótese de que não se trata apenas de pessoas menos reflexivas recorrendo mais à inteligência artificial, mas de a ferramenta, em certos modos de uso, efetivamente reduzir o engajamento cognitivo.

Do conhecimento ao desempenho: o que se perde na prática

Resta examinar se esses efeitos cognitivos se traduzem em resultados acadêmicos mensuráveis. A evidência, embora mais heterogênea, aponta numa direção preocupante quando o uso é dependente e desestruturado. Revisão sistemática de estudos publicados entre 2023 e 2025 reuniu pesquisas que examinaram indicadores como média de notas, escores em provas, qualidade da escrita e precisão de código. Entre os achados, registrou-se que estudantes que usavam o ChatGPT com frequência apresentavam médias notavelmente mais baixas, sobretudo aqueles já propensos à procrastinação ou submetidos a pressão de tempo; em outro estudo da mesma revisão, alunos que recorreram à inteligência artificial na preparação para provas pontuaram, em média, cerca de sete pontos abaixo dos que não a utilizaram (Abbas, *et al.*, 2025).

Estudo realizado em uma universidade do Zimbábue com 248 estudantes de graduação oferece um dos dados mais concretos a respeito. Os pesquisadores identificaram que 32,7% dos participantes apresentavam padrões aditivos de uso da inteligência artificial generativa, com média de 18,3 interações diárias e tentativas frustradas de redução relatadas por 65,8% deles. Mais relevante para os fins deste artigo, a análise estatística revelou correlações negativas significativas entre a severidade da dependência e o desempenho acadêmico: os estudantes fortemente dependentes exibiam um déficit médio de 0,41 ponto na média geral em comparação com os colegas não dependentes (Chiwaridzo, 2025). A Tabela 2 reúne os principais indicadores quantitativos discutidos nesta seção.

16

Tabela 2 Síntese de evidências empíricas sobre efeitos cognitivos e de desempenho do uso dependente da inteligência artificial.

Estudo	Amostra	Principal achado
Kosmyna <i>et al.</i> (2025). MIT, eletroencefalografia	54	83% dos usuários do modelo de linguagem não recordavam o próprio texto; menor conectividade neural no grupo dependente.
Gerlich (2025). Societies	666	Correlação negativa entre uso de IA e pensamento crítico, mediada por descarregamento cognitivo; jovens mais afetados.
Lee <i>et al.</i> (2025). Microsoft e Carnegie Mellon	319	Maior confiança na IA associada a menor engajamento em pensamento crítico; homogeneização dos resultados.
Wecks <i>et al.</i> (2024), apud Abbas <i>et al.</i> (2025)	Diversas	Estudantes que usaram IA na preparação para provas pontuaram cerca de 7 pontos abaixo dos demais.
Chiwaridzo (2025). Zimbábue	248	32,7% com padrão aditivo; déficit médio de 0,41 na média geral entre dependentes.

Fonte: elaborada pelos autores, 2026.

É imprescindível, todavia, resistir à tentação de uma leitura unilateral. A própria literatura adverte que os efeitos da inteligência artificial sobre a aprendizagem não são uniformemente negativos. Metanálise de 51 estudos experimentais identificou que o uso do ChatGPT pode exercer efeito positivo amplo sobre o desempenho de aprendizagem, quando empregado de forma estruturada e pedagogicamente orientada (Wang, Fan, 2025). Experimentos com tutores baseados em inteligência artificial demonstraram que estudantes podem aprender mais em menos tempo, em comparação com a aula expositiva tradicional. O fator decisivo, repetidamente, não é a presença da ferramenta, mas o modo de seu emprego, ou seja, a diferença entre o estudante que usa a inteligência artificial para gerar perguntas, testar hipóteses e receber retorno sobre o próprio raciocínio, e aquele que a usa para evitar pensar.

Para além das notas, há um efeito mais difícil de mensurar, porém mais profundo, que a literatura tem designado como desqualificação, ou deskilling, isto é, a erosão gradual de competências que, por deixarem de ser exercitadas, atrofiam. O raciocínio é simétrico ao do treino físico, pois uma faculdade não exercitada definha. Quando o estudante delega rotineiramente à inteligência artificial a tarefa de estruturar um argumento, sintetizar uma leitura ou depurar um trecho de código, ele deixa de praticar precisamente as operações que a universidade deveria consolidar. O perigo, aqui, não é a nota de uma prova, mas a competência que não se forma e que, por não ter sido formada, tampouco será percebida como ausente. Trata-se de uma perda silenciosa, que não dispara alarmes imediatos e que só se revela quando o apoio é retirado, exatamente o que os experimentos com retirada da ferramenta começam a evidenciar.

Um aspecto particularmente sensível dessa desqualificação diz respeito à criatividade e à originalidade do pensamento. Quando uma turma inteira recorre à mesma ferramenta, treinada sobre os mesmos dados e otimizada para produzir a resposta estatisticamente mais provável, o resultado tende à convergência. Os trabalhos passam a se parecer, as ideias se padronizam e a diversidade de perspectivas que constitui a riqueza do debate acadêmico se reduz. É nesse ponto que a analogia com o colapso de modelo se mostra mais literal do que metafórica, pois a homogeneização observada nos estudos empíricos sobre uso de inteligência artificial reproduz, no plano coletivo da produção estudantil, exatamente o empobrecimento de variabilidade que degrada os modelos realimentados por suas próprias saídas. A universidade, que deveria ser o espaço da divergência produtiva e do pensamento singular, corre o risco de se

tornar um eco de respostas médias, caso não cultive deliberadamente o engajamento crítico que resiste a essa convergência.

O quadro, é justo reconhecer, comporta vozes dissonantes e ponderações necessárias. Alguns pesquisadores argumentam que o receio em torno da inteligência artificial repete, ponto por ponto, temores históricos que se mostraram exagerados, de Sócrates, que na Antiguidade desconfiava da escrita por enfraquecer a memória, às inquietações com a calculadora e com a própria internet. Sob essa ótica, a inteligência artificial seria apenas mais uma extensão da mente, e a redistribuição do esforço cognitivo, longe de empobrecedora, liberaria recursos para operações de ordem superior. Há, ademais, a crítica metodológica legítima de que muitos dos estudos alarmantes são correlacionais, baseados em autorrelato ou conduzidos em condições laboratoriais pouco representativas do uso real. Essas ressalvas são pertinentes e impedem qualquer conclusão catastrofista. O que a balança das evidências sugere, contudo, não é que a inteligência artificial torne os estudantes inelutavelmente mais limitados, mas que, sob um modo específico de uso, o substitutivo e desprovido de engajamento crítico, ela se associa a custos cognitivos reais e mensuráveis. É essa nuance, e não o alarme nem o entusiasmo, que deve orientar a resposta educacional.

Há ainda uma dimensão que conecta diretamente o desempenho à integridade acadêmica e que merece registro. Levantamentos recentes indicam que mais da metade dos estudantes considera o uso da inteligência artificial para concluir integralmente uma tarefa uma forma de fraude, mas a fronteira entre o auxílio legítimo e a substituição ilícita permanece difusa, tanto para discentes quanto para docentes. Essa ambiguidade normativa agrava o problema cognitivo, pois, na ausência de critérios claros, o estudante tende a maximizar a conveniência, deslizando do apoio para a substituição sem perceber que cruza uma fronteira pedagogicamente decisiva. A erosão epistêmica, nesse sentido, não é apenas um efeito individual, mas o produto de um vácuo institucional, e é por isso que sua superação depende menos da vontade isolada do aluno e mais do redesenho das regras e das avaliações que estruturam o trabalho acadêmico.

O contexto brasileiro e o risco da desigualdade epistêmica

Embora boa parte das evidências de maior robustez metodológica provenha de contextos do Norte global, há razões para supor que o fenômeno assuma contornos particularmente sensíveis no Brasil. A adoção, como visto, já alcança cerca de 70% dos universitários, mas se dá em um sistema de ensino superior marcado por profunda heterogeneidade de condições

pedagógicas e por desigualdades estruturais de capital cultural (AGÊNCIA BRASIL, 2024). Estudos nacionais recentes, conduzidos em instituições brasileiras, registram uma percepção predominantemente positiva dos estudantes quanto à inteligência artificial, com o ChatGPT como ferramenta mais utilizada, mas apontam, em contrapartida, preocupações persistentes da comunidade acadêmica com a aprendizagem real e o desenvolvimento de habilidades críticas (Reer, 2025).

O achado de Gerlich (2025) de que a escolaridade e os hábitos de reflexão funcionam como fatores de proteção contra a erosão tem, nesse contexto, uma implicação distributiva inquietante. Se o uso reflexivo da inteligência artificial depende de um repertório crítico previamente constituído, então os estudantes que chegam à universidade com menor capital cultural, frequentemente os de origem socioeconômica mais vulnerável, estão mais expostos ao uso substitutivo e, portanto, à erosão. A inteligência artificial, longe de ser o nivelador democrático que por vezes se anuncia, pode aprofundar uma desigualdade epistêmica: aos que já sabem pensar, ela serve de apoio; aos que ainda estão aprendendo, pode servir de muleta que impede o aprendizado. Para um país que vê na educação superior um vetor de mobilidade social, o alerta é grave e merece investigação empírica específica.

A esse risco soma-se a fragilidade das respostas institucionais. Pesquisas internacionais indicam que, embora a percepção dos estudantes sobre o preparo das instituições para lidar com a inteligência artificial tenha melhorado, ainda persiste um descompasso considerável entre a velocidade da adoção e a maturidade das políticas de uso. No Brasil, onde a regulamentação do tema ainda engatinha e onde muitas instituições carecem de diretrizes claras, o vácuo normativo tende a ser preenchido por orientações informais, transmitidas oralmente e sem consistência entre docentes e cursos. Esse cenário de indefinição não é neutro, pois transfere ao estudante, individualmente, a tarefa de decidir o que é uso legítimo, justamente no momento em que ele ainda não dispõe do repertório crítico necessário para fazê-lo com discernimento. A construção de políticas institucionais explícitas, articuladas com a formação docente, configura-se, assim, como condição para que a ferramenta seja incorporada de modo a fortalecer, e não a esvaziar, a aprendizagem.

A erosão epistêmica e a analogia do colapso de modelo

Reunidas as evidências, é possível retomar e refinar a hipótese central. O conceito de erosão epistêmica, tal como aqui empregado, designa o processo pelo qual o conhecimento real

do estudante, entendido como a capacidade internalizada de analisar, avaliar, sintetizar e criar, se esvazia progressivamente quando o esforço cognitivo é terceirizado de modo sistemático. Não se trata de ignorância no sentido clássico, da ausência de informação, mas de algo mais sutil: a posse aparente de um produto intelectual que não foi efetivamente construído e que, por isso, não se sustenta sem a muleta que o gerou. O estudante do estudo do MIT que não conseguia citar a própria redação é a imagem emblemática dessa condição.

A analogia com o colapso de modelo ilumina esse processo em três planos. Primeiro, no plano da perda de variabilidade, pois, assim como o modelo realimentado por dados sintéticos converge para saídas homogêneas, o estudante que recorre invariavelmente à inteligência artificial tende a reproduzir um repertório padronizado, o que faz perder a singularidade do próprio pensamento, fenômeno já observado empiricamente na homogeneização dos resultados (Lee, *et al.*, 2025). Segundo, no plano da amplificação de erros, uma vez que o modelo colapsado reforça os próprios vieses, e o estudante que aceita acriticamente as respostas da inteligência artificial, incluindo suas conhecidas alucinações, incorpora distorções que não tem mais o instrumental para detectar. Terceiro, no plano da irreversibilidade relativa, já que a dívida cognitiva documentada no estudo do MIT sugere que o hábito de delegar deixa marcas que não se desfazem instantaneamente, tal como a degradação do modelo se acumula geração após geração.

20

Há, no entanto, uma diferença fundamental e esperançosa entre os dois sistemas. O modelo de inteligência artificial não tem agência sobre o próprio treinamento; o estudante, sim. A reversibilidade observada no estudo de Gerlich, em que a escolaridade e os hábitos de reflexão protegem contra a erosão, indica que o colapso humano não é uma fatalidade técnica, mas uma contingência pedagógica. É exatamente nesse ponto que a discussão deixa de ser diagnóstica para se tornar propositiva.

Essa diferença de agência tem uma consequência prática que convém sublinhar. Enquanto o colapso de um modelo de inteligência artificial só pode ser evitado por intervenção externa, no nível de seus desenvolvedores, a erosão epistêmica do estudante pode ser prevenida e até revertida pela própria comunidade educativa, desde que esta atue de modo deliberado. O paralelo, portanto, não conduz ao fatalismo, mas a um chamado à responsabilidade. Saber que o uso substitutivo da ferramenta tende a empobrecer o repertório cognitivo não é motivo para abandoná-la, e sim razão para empregá-la com discernimento. A história das tecnologias educacionais mostra que nenhuma ferramenta é, em si, salvadora ou destrutiva, e que seu efeito

depende inteiramente das mediações pedagógicas que a cercam. A inteligência artificial não foge a essa regra, e é nessa margem de escolha, que distingue o ser humano do modelo, que reside a esperança de um uso que fortaleça, em vez de esvaziar, o pensamento.

Da inteligência artificial como apoio do pensar

Se a erosão epistêmica não é destino, mas contingência, então a pergunta decisiva desloca-se do diagnóstico para a prescrição: sob que condições a mesma ferramenta que pode esvaziar o pensamento pode, ao contrário, fortalecê-lo? A literatura mais cuidadosa sugere que a resposta reside menos em proibir e mais em reconfigurar o modo de uso. O conceito pedagógico de andaime, originalmente associado à zona de desenvolvimento proximal de Vigotski, oferece a chave: o andaime é um apoio temporário, calibrado para sustentar o aprendiz naquilo que ainda não consegue fazer sozinho, mas projetado para ser progressivamente retirado à medida que a competência se internaliza. Uma inteligência artificial bem empregada funciona como esse apoio; uma inteligência artificial mal empregada torna-se um substrato permanente, sem o qual o edifício cognitivo não se sustenta.

Na prática, a distinção se traduz em modos de interação concretos. Solicitar à ferramenta que explique um conceito que será em seguida aplicado a um problema novo, pedir que ela formule contra-argumentos a uma tese que o estudante defenderá, requisitar que aponte lacunas em um raciocínio que o próprio aprendiz preencherá, todos esses usos mantêm o esforço cognitivo do lado humano e empregam a máquina como interlocutora que provoca, e não como oráculo que dispensa. Inversamente, pedir o texto pronto, a resposta acabada ou o problema resolvido transfere para a máquina justamente a operação que deveria formar o estudante. O mesmo aplicativo serve aos dois propósitos, e a diferença está na intenção pedagógica que organiza seu uso e na presença de uma postura crítica que verifica, contesta e se apropria do que a ferramenta oferece.

Daí decorre que a responsabilidade não pode ser depositada apenas sobre o estudante. Cabe ao docente desenhar tarefas que tornem o uso substitutivo inviável ou improdutivo, e que recompensem o engajamento genuíno; cabe à instituição construir uma cultura e uma política de uso que distingam claramente o apoio da fraude; e cabe ao próprio campo da pesquisa desenvolver instrumentos capazes de avaliar não o produto final, facilmente forjável por uma máquina, mas o processo de pensamento que o gerou. A literacia crítica em inteligência artificial, nesse sentido, não é um conteúdo a ser acrescentado ao currículo, mas uma

competência transversal que perpassa todas as disciplinas. Aprender a pensar com a inteligência artificial sem deixar que ela pense em nosso lugar talvez seja a habilidade formativa mais importante da década que se inicia.

Em termos concretos, o redesenho avaliativo pode assumir múltiplas formas, todas voltadas a tornar visível o processo cognitivo. Defesas orais nas quais o estudante deve explicar e justificar o que produziu reintroduzem a exigência de internalização; portfólios que documentam rascunhos, revisões e decisões tornam o percurso de elaboração tão avaliável quanto o resultado; tarefas que exigem a aplicação de um conceito a um contexto novo e específico resistem melhor à geração automática; e atividades metacognitivas, nas quais o aluno reflete criticamente sobre o próprio uso da ferramenta, transformam a inteligência artificial de ameaça em objeto de aprendizagem. Nenhuma dessas estratégias é, isoladamente, uma solução completa, mas em conjunto elas deslocam o eixo da avaliação do produto para o pensamento, que é exatamente o terreno em que a erosão epistêmica precisa ser combatida.

CONSIDERAÇÕES FINAIS

Este artigo propôs-se a investigar a relação entre o uso intensivo da inteligência artificial generativa por estudantes universitários e a erosão do conhecimento real, mobilizando a metáfora do colapso de modelo como chave heurística de leitura. A revisão das evidências mais recentes permite afirmar, com a prudência que o estado ainda incipiente do campo recomenda, que a preocupação tem fundamento empírico consistente. A adoção da inteligência artificial no ensino superior atingiu níveis de saturação, superando 90% em alguns países e 70% no Brasil, antes que se desenvolvesse a literacia crítica capaz de mediá-la, e estudos de natureza neural, cognitiva e de desempenho convergem para um padrão em que o uso dependente se associa a menor engajamento cerebral, descarregamento cognitivo, queda no pensamento crítico e prejuízo a indicadores acadêmicos.

A tese que organizou o trabalho, a de que terceirizar sistematicamente o pensar empobrece o repertório cognitivo, num processo análogo ao colapso dos modelos que se realimentam das próprias saídas, encontra respaldo, em particular, no achado do MIT de que 83% dos usuários de inteligência artificial não recordavam o próprio texto, na correlação documentada por Gerlich entre uso da ferramenta e declínio do pensamento crítico, e na homogeneização de resultados observada por Lee e colaboradores. Não se trata, contudo, de uma condenação da ferramenta. A mesma literatura demonstra que, empregada como apoio e não

como substituta, a inteligência artificial pode ampliar o aprendizado. A variável decisiva não é tecnológica, mas pedagógica, pois reside no modo de uso e na presença, ou ausência, de uma cultura de engajamento cognitivo.

Convém destacar que a contribuição central deste artigo não está em nenhum dado isolado, mas na convergência que a síntese revela. Tomados separadamente, cada estudo aqui discutido comporta limitações que autorizariam a cautela. Reunidos, porém, eles desenham um padrão coerente, no qual evidências neurais, cognitivas, de desempenho e de homogeneização apontam na mesma direção, partindo de métodos e contextos independentes. É essa triangulação que confere solidez à preocupação e que distingue o presente argumento de um pânico passageiro. Articular o que estava disperso, conectando a neurociência do MIT, a psicologia cognitiva de Gerlich, a pesquisa aplicada da Microsoft e os dados de desempenho de diferentes países a um marco conceitual unificado, é precisamente a tarefa que justifica uma revisão como esta e que pode orientar tanto a prática docente quanto a agenda de pesquisas futuras no contexto brasileiro.

Depreendem-se daí implicações práticas que extrapolam o diagnóstico. Às instituições, impõe-se a tarefa urgente de redesenhar avaliações que não possam ser simplesmente concluídas por uma máquina, deslocando o foco do produto acabado para o processo de pensamento, por meio de defesas orais, portfólios que evidenciem o percurso de elaboração e tarefas que exijam aplicação contextual e reflexão metacognitiva sobre o próprio uso da ferramenta. Cabe igualmente promover formação sistemática em literacia crítica de inteligência artificial, tanto para discentes quanto para docentes, capaz de ensinar não apenas a operar a ferramenta, mas a interrogá-la, a verificar suas respostas, a detectar suas alucinações e a distinguir o apoio da muleta. Aos estudantes, cabe a consciência de que a conveniência imediata pode cobrar, em dívida cognitiva, um preço diferido e por vezes invisível. E ao poder público, sobretudo no contexto brasileiro, impõe-se atenção ao risco de que a inteligência artificial aprofunde, em vez de reduzir, as desigualdades educacionais que ainda marcam o acesso ao conhecimento.

Com efeito, este estudo apresenta limitações que devem ser explicitadas. Por ser uma revisão narrativa, está sujeito à seletividade na escolha das fontes; a literatura analisada é muito recente e, em parte, ainda não revisada por pares; e a maioria dos estudos disponíveis é de natureza correlacional, o que impede afirmações causais definitivas. A própria analogia com o colapso de modelo, reitera-se, é heurística e não literal. Pesquisas futuras se beneficiariam de desenhos longitudinais e experimentais, conduzidos especificamente no contexto brasileiro,

capazes de mensurar a evolução das competências cognitivas dos estudantes ao longo do tempo e de testar intervenções pedagógicas voltadas ao uso reflexivo da ferramenta. A pergunta do título permanece, afinal, em aberto e incômoda: quem pensa quando a inteligência artificial pensa? A resposta que a universidade construir para ela definirá, em boa medida, o tipo de conhecimento e de pensador que será capaz de formar nas próximas décadas.

Cumprido encerrar com uma observação sobre o papel do professor, frequentemente relegado a segundo plano nos debates sobre tecnologia educacional. A inteligência artificial, por mais sofisticada que seja, não substitui a relação pedagógica que se estabelece entre quem ensina e quem aprende, nem a mediação humana que transforma informação em formação. Ao contrário, num cenário de adoção massiva da ferramenta, o docente torna-se mais necessário, e não menos, justamente porque cabe a ele cultivar aquilo que a máquina não oferece, quais sejam a provocação intelectual, o acompanhamento do percurso, o estímulo à dúvida produtiva e o reconhecimento do esforço genuíno. A erosão epistêmica não será revertida por proibições nem por detectores automáticos, mas por uma pedagogia que devolva ao estudante o protagonismo do próprio pensar. Nesse sentido, o desafio que a inteligência artificial coloca à universidade é menos tecnológico do que humano, e a resposta a ele dependerá, sobretudo, da capacidade das instituições de reafirmar o valor insubstituível do esforço de pensar.

REFERÊNCIAS

- ABBAS M, *et al.* The impact of generative AI tools on academic performance in higher education: a systematic review (2023 a 2025). *Journal of Educational Technology and Higher Education*, 2025; 2(1): 1-20.
- AGÊNCIA BRASIL. Sete a cada dez estudantes usam IA na rotina de estudos. Empresa Brasil de Comunicação, 2024.
- ARMITAGE R. Your brain on ChatGPT. *British Journal of General Practice*, 2025; 75(758): 410.
- CHIWARIDZO OT. Generative AI dependency: the emerging academic crisis and its impact on student performance, a case study of a university in Zimbabwe. *Cogent Education*, 2025; 12(1): 2549787.
- DIGITAL EDUCATION COUNCIL. Global AI Student Survey 2024. Digital Education Council, 2024.
- FAVERO L, *et al.* Cognitive offloading and the future of learning with generative AI. *Education and Information Technologies*, 2025; 30(4): 1-22.
- FREEMAN J. Student Generative AI Survey 2025. Higher Education Policy Institute, 2025.
- GERLICH M. AI tools in society: impacts on cognitive offloading and the future of critical thinking. *Societies*, 2025; 15(1): 6.
- KOSMYNA N, *et al.* Your brain on ChatGPT: accumulation of cognitive debt when using an AI assistant for essay writing task. *arXiv*, 2025; 2506.08872.

LEE HP, *et al.* The impact of generative AI on critical thinking: self-reported reductions in cognitive effort and confidence effects from a survey of knowledge workers. *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, 2025; 1121: 1-22.

REER. O impacto da inteligência artificial entre estudantes do ensino superior. *Revista Eletrônica da Estácio Recife*, 2025; 11(1): 1-15.

RISKO EF, GILBERT SJ. Cognitive offloading. *Trends in Cognitive Sciences*, 2016; 20(9): 676-688.

SHUMAILOV I, *et al.* AI models collapse when trained on recursively generated data. *Nature*, 2024; 631(8022): 755-759.

SPARROW B, *et al.* Google effects on memory: cognitive consequences of having information at our fingertips. *Science*, 2011; 333(6043): 776-778.

STANKOVIĆ M, *et al.* Comment on: your brain on ChatGPT, accumulation of cognitive debt when using an AI assistant for essay writing tasks. *arXiv*, 2026; 2601.00856.

WANG J, FAN W. The effect of ChatGPT on students' learning performance, learning perception, and higher-order thinking: a meta-analysis. *Humanities and Social Sciences Communications*, 2025; 12(1): 1-18.