

TƯ DUY PHẢN BIỆN TRONG THỜI ĐẠI AI: TÁC ĐỘNG CỦA AI ĐỐI VỚI PHÁT TRIỂN TƯ DUY PHẢN BIỆN

Lê Thị Thu Hà, Nguyễn Đức Tài
Trường Đại học Giáo dục, ĐHQGHN

Tóm tắt: Khi các công cụ trí tuệ nhân tạo (AI) tạo sinh như Chat GPT ngày càng được tích hợp vào giáo dục đại học, ảnh hưởng của chúng đến sự phát triển nhận thức của sinh viên – đặc biệt là tư duy phản biện – đang thu hút sự quan tâm ngày càng lớn. Nghiên cứu trình bày một tổng quan tài liệu dạng tường thuật, tổng hợp và trình bày tổng quan các nghiên cứu được công bố trong giai đoạn 2020–2025, xoay quanh mối quan hệ giữa việc sử dụng AI và năng lực tư duy phản biện của sinh viên đại học trên toàn thế giới. Tổng quan chỉ ra một ảnh hưởng hai chiều. Một mặt, các ứng dụng AI có thể thúc đẩy tư duy phản biện thông qua phản hồi cá nhân hóa; cung cấp góc nhìn đa dạng và cải thiện khả năng tiếp cận thông tin; làm giảm gánh nặng trí tuệ ngắn hạn (Chunpeng Zhai, 2024). Mặt khác, sự phụ thuộc quá mức vào nội dung do AI tạo ra hoặc thiếu sự định hướng có thể làm giảm nỗ lực nhận thức, dẫn đến “sự lười biếng siêu nhận thức” và suy yếu khả năng suy luận độc lập và sáng tạo (Nguyễn Thiên Thanh Trúc, 2024; Fan, 2025; Lee, 2025). Kết quả tổng quan cho thấy vai trò của thiết kế sư phạm trong việc trung gian hóa tác động nhận thức của AI – cho thấy rằng AI mang lại lợi ích lớn nhất khi được tích hợp trong bối cảnh học tập chủ động và có tính phản tư.

Từ khóa: Trí tuệ nhân tạo, tư duy phản biện, sinh viên

CRITICAL THINKING IN THE AGE OF AI: THE IMPACT OF AI ON THE DEVELOPMENT OF CRITICAL THINKING

Le Thi Thu Ha, Nguyen Duc Tai
University of Education, VNU

Abstract: As generative artificial intelligence (AI) tools such as ChatGPT are increasingly integrated into higher education, their impact on students' cognitive development—particularly critical thinking—has drawn growing attention. This study presents a narrative literature review, synthesizing and summarizing research published between 2020 and 2025 on the relationship between AI use and the critical thinking abilities of university students worldwide. The review reveals a dual impact. On the one hand, AI applications can foster critical thinking through personalized feedback, offering diverse perspectives and improving access to information, while also reducing short-term cognitive load (Chunpeng Zhai, 2024). On the other hand, excessive reliance on AI-generated content or lack of proper guidance may diminish cognitive effort, leading to “metacognitive laziness” and weakening independent as well as creative reasoning skills (Nguyen Thien Thanh Truc, 2024; Fan, 2025; Lee, 2025). The findings also highlight the mediating role of pedagogical design in shaping AI's cognitive impact—suggesting that AI provides the greatest benefits when integrated within active and reflective learning contexts.

Keywords: Artificial Intelligence, Critical thinking, ChatGPT, Students

Nhận bài: 18/07/2025

Phản biện: 16/08/2025

Duyệt đăng: 21/08/2025

I. ĐẶT VẤN ĐỀ

Trí tuệ nhân tạo (AI) lần đầu được giới thiệu bởi John McCarthy năm 1956 như một lĩnh vực nghiên cứu nhằm mô phỏng trí thông minh con người trong máy móc. Trải qua quá trình phát triển, AI đã được ứng dụng rộng rãi trong nhiều lĩnh vực, đặc biệt là giáo dục. Khác với các công cụ truyền thống, AI tạo sinh có khả năng tự động tạo ra dữ liệu mới dưới dạng văn bản, hình ảnh, âm nhạc, với thành phần chính là các mô hình ngôn ngữ lớn. Trong giáo dục, trí tuệ nhân tạo trong giáo dục (AIED) hình thành từ thập niên 1970, tập trung vào việc phát triển phần mềm hỗ trợ dạy - học, cá nhân hóa quá trình học tập và cung cấp phản hồi kịp thời (Beverly Park Woolf, 2015). Tuy nhiên, AI cũng đặt ra những thách thức về tính chính xác và ảnh hưởng tiêu cực đến thói quen học tập (Đinh Thị Mỹ Hạnh & Trần Văn Hưng, 2021).

Trong số các công cụ AI, ChatGPT nổi bật với tốc độ lan tỏa nhanh chóng sau khi ra mắt tháng 11/2022, đạt 180,5 triệu người dùng hàng tháng tính đến cuối năm 2023 (TechCrunch, 2023). ChatGPT ngày càng được nâng cấp, hỗ trợ đa phương thức và phục vụ nhiều nhu cầu sáng tạo. Trong bối cảnh giáo dục đại học, công cụ này mở ra cơ hội phát triển tư duy phản biện – kỹ năng cốt lõi trong phân tích, đánh giá và giải quyết vấn đề (Halpern, 2010; Emily R Lai, 2011). Tổng quan này góp phần làm rõ vai trò của AI trong giáo dục, đồng thời định hướng cho giảng viên và nhà hoạch định chính sách trong việc tích hợp công nghệ một cách có trách nhiệm và hiệu quả.

II. NỘI DUNG NGHIÊN CỨU

2.1. Phương pháp nghiên cứu

Phương pháp nghiên cứu lý luận được sử dụng

để phân tích và tổng hợp các tài liệu nghiên cứu về Chat GPT và tư duy phản biện. Các cơ sở dữ liệu đã sử dụng bao gồm: PubMed, ResearchGate, Google Scholar, ScienceDirect, Semantic Scholar.

Nghiên cứu tổng hợp các kết quả từ các bài tập chí khoa học bằng tiếng Anh và tiếng Việt,

Bảng: Ưu điểm và nhược điểm khi sử dụng AI được tổng hợp từ các tài liệu

Ưu điểm	Nhược điểm
Công cụ có khả năng giao tiếp và hoàn thành tác vụ bằng văn bản cũng như tìm kiếm và tổng hợp thông tin, dịch thuật hoặc hỗ trợ các công việc khác dựa trên yêu cầu người dùng (Ngo, 2023)	Phụ thuộc quá nhiều vào Chat GPT có thể dẫn đến nguy cơ hạn chế khả năng học tập và tư duy (Kasneji và cộng sự, 2023)
Người học tăng hiệu quả khả năng tự học (Kasneji và cộng sự, 2023), việc học tập được cá nhân hóa	Giảm động lực học tập do khả năng ghi nhớ bị suy yếu (Bai và cộng sự, 2023)
Phát triển kỹ năng đọc và viết (Duong và cộng sự, 2023)	Tính không kiểm chứng trong tri thức do AI cung cấp có thể gây ra các vấn đề về liên chính học thuật
Hỗ trợ người học trong hoạt động nhóm (Saxena & Doleck, 2023)	Khả năng tổng hợp được lượng thông tin lớn của Chat GPT cũng tạo ra sự quá tải nhận thức do người dùng cần phải xử lý thông tin nhiều hơn bởi sự nghi ngờ về chất lượng và nguồn gốc thông tin mà công cụ này đem lại (Bai và cộng sự, 2023)
Cung cấp những phản hồi thường xuyên và nhanh chóng	Làm tăng sự ỷ lại, lười nhác của người học/người sử dụng
AI có thể giúp cải thiện hiệu suất ngắn hạn (Yizhou Fan, 2025)	Rủi ro về tính bảo mật và riêng tư của dữ liệu hay thông tin người dùng cung cấp

Cơ chế của AI là tự học từ các thông tin con người nạp vào chúng. Ngoài ra, chúng không ngừng bổ sung kiến thức cho bản thân trên môi trường internet bất chấp việc các nguồn tri thức đó có được kiểm chứng hay không. Tri thức AI tổng hợp mang tính toàn cầu, nó có thể đúng ở quốc gia, khu vực này nhưng lại không đúng ở quốc gia, khu vực, bối cảnh khác, chưa kể đến các yếu tố khác như chính trị, kinh tế hay văn hóa xã hội đặc thù ở mỗi khu vực địa lý. Bởi vậy, các thông tin do AI cung cấp mang tính rủi ro cao khi sử dụng mà không có sự tìm hiểu hay phân tích lại. Chẳng hạn như thông tin Chat GPT đưa ra có thể là thông tin sai lệch, thông tin chúng tự tạo ra khiến cho việc học tập của sinh viên bị lệch lạc. Open AI biết về hạn chế này, họ cũng cho biết rằng “Chat GPT đôi khi biết những câu trả lời nghe có vẻ hợp lý nhưng lại không chính xác hoặc vô nghĩa.” Người dùng có thể không phân biệt được đúng hay sai, khiến cho quá trình học tập và làm việc trở nên rủi ro hơn nếu phụ thuộc hoàn toàn vào chúng.

2.2.2 Trí tuệ nhân tạo trong giáo dục

Các hệ thống trí tuệ nhân tạo (AI) được huấn

bao gồm cả nghiên cứu định lượng và định tính. Về mặt thời gian, các nghiên cứu được tổng hợp trong giai đoạn từ 2020 đến 2025.

2.2. Kết quả và thảo luận

2.2.1 Ưu điểm và hạn chế khi dùng AI/Chat GPT tổng hợp từ các tài liệu

luyện với mục tiêu cung cấp câu trả lời hữu ích, chính xác và không gây xúc phạm, song ba mục tiêu đôi khi mâu thuẫn nhau (Alexander, 2022; Urban, 2024). Trong thực tiễn, ChatGPT có thể tạo ra phản hồi nghe hợp lý nhưng chỉ đúng một phần nhằm thỏa mãn kỳ vọng người dùng (Bang, 2023). Công cụ giúp học sinh nhanh chóng tiếp cận thông tin, cải thiện kết quả học tập trong ngắn hạn (Rahman & Watanobe, 2023). Tuy nhiên, sự tiện lợi cũng tiềm ẩn rủi ro: học sinh dễ phụ thuộc vào AI, hạn chế nỗ lực tư duy độc lập, từ đó suy giảm kỹ năng tư duy phản biện và giải quyết vấn đề (Hao & Yu, 2023). Đặc biệt, khi AI cung cấp thông tin sai, người học có thể tiếp nhận mà không kiểm chứng (Chunpeng Zhai, 2024).

Mặc dù vậy, AI không phải bản chất tiêu cực; vấn đề nằm ở cách sử dụng. GenAI thay đổi bản chất nỗ lực tư duy phản biện theo ba hướng: từ tìm kiếm sang xác minh thông tin; từ giải quyết vấn đề sang tích hợp phản hồi AI; và từ thực hiện sang giám sát nhiệm vụ. Người dùng vẫn phải đầu tư công sức vào xác minh, điều chỉnh và đánh giá phản hồi để đảm bảo tiêu chuẩn học tập.

AI cũng mở ra cơ hội cá nhân hóa học tập thông qua phân tích phong cách, sở thích và trình độ người học, từ đó thiết kế nội dung và phản hồi phù hợp. Cách tiếp cận này nâng cao hiệu quả, nhưng tiềm ẩn nguy cơ gia tăng bất bình đẳng khi học sinh có điều kiện kinh tế - xã hội hạn chế khó tiếp cận công cụ chất lượng cao (Hao & Yu, 2023; Sullivan, 2023). Ngoài ra, việc truyền thông tập trung vào ChatGPT như công cụ gian lận đã ảnh hưởng đến nhận thức xã hội và thái độ học tập của sinh viên, trong khi những lợi ích học thuật của AI chưa được nhấn mạnh đúng mức.

Nếu được sử dụng đúng cách, AI có thể nâng cao nghiên cứu học thuật, phát triển tư duy phản biện, giải quyết vấn đề và kỹ năng nghiên cứu. GenAI mang lại nhiều lợi ích: hỗ trợ cá nhân hóa, tiết kiệm thời gian tìm kiếm thông tin, cung cấp phản hồi tức thì, và tạo điều kiện để người học tập trung vào phân tích, ứng dụng tri thức (Haman, 2023; Fan, 2025).

2.2.3. Trí tuệ nhân tạo và tư duy phản biện

Tư duy phản biện (Critical Thinking), được định nghĩa là khả năng phân tích, đánh giá và tổng hợp thông tin để đưa ra quyết định hợp lý, là một kỹ năng nhận thức cơ bản cần thiết cho thành công trong học tập, năng lực chuyên môn và quyền công dân có hiểu biết. Ba sự thay đổi chính trong bản chất nỗ lực tư duy phản biện: (1) Chuyển từ tìm kiếm thông tin sang xác minh thông tin; (2) Chuyển từ giải quyết vấn đề sang tích hợp phản hồi từ AI; (3) Chuyển từ thực hiện nhiệm vụ sang giám sát nhiệm vụ, bao gồm việc diễn đạt ý định thành câu lệnh, điều hướng phản hồi của AI và đánh giá kết quả so với tiêu chuẩn công việc. Nhìn chung, dù GenAI giúp giảm tải nhiều khâu (tìm kiếm thông tin, tạo nội dung, phản hồi cá nhân hóa), người dùng vẫn phải đầu tư đáng kể vào việc xác minh, chỉnh sửa, điều hướng và đánh giá chất lượng phản hồi của AI để đảm bảo đáp ứng tiêu chuẩn công việc.

Một nghiên cứu (Acosta-Enriquez & cs, 2025) trên sinh viên đại học từ nhiều ngành học khác nhau dựa vào công cụ khảo sát gồm bảng hỏi tần suất và mục đích sử dụng AI tạo sinh, thang đo căng thẳng học tập và thang đo tư duy phản biện cho thấy hiệu quả của AI phụ thuộc vào điểm số tư duy phản biện ban đầu của sinh viên. Căng thẳng học tập và tư duy phản biện đều là biến trung gian quan trọng, tác động mạnh mẽ mối quan hệ giữa AI và kết quả học tập. Tác giả đưa ra kết luận rằng AI tạo sinh có tiềm năng nâng cao kết quả học tập

nếu được tích hợp vào chiến lược học tập có chủ đích và rèn luyện tư duy phản biện. Một nghiên cứu khác cũng đưa ra kết quả tư duy phản biện có mối quan nghịch đáng kể với hành vi phụ thuộc vào AI (Hou & cs, 2025). Tư duy phản biện chính là yếu tố bảo vệ trước nguy cơ lệ thuộc vào AI tạo sinh.

Kết quả chính từ nghiên cứu (Eriona Cela & cs, 2024) cho thấy không có sự khác biệt đáng kể về kỹ năng tư duy phản biện giữa những sinh viên đã từng tiếp xúc với các công cụ AI và những sinh viên chưa từng tiếp xúc. Tuy nhiên, có một mối tương quan nghịch có ý nghĩa thống kê giữa việc phụ thuộc vào các công cụ AI cho bài tập và kỹ năng giải quyết vấn đề của sinh viên, cho thấy rằng việc phụ thuộc quá mức vào AI có thể cản trở sự phát triển khả năng giải quyết vấn đề độc lập. Công cụ AI có thể hỗ trợ cho tư duy phản biện, miễn là việc sử dụng chúng được định hướng theo các mục đích minh bạch và đạo đức (Nuria Chaparro- Banegas & cs, 2024).

Có một mối tương quan âm đáng kể giữa việc sử dụng thường xuyên công cụ AI và khả năng tư duy phản biện, được trung gian bởi việc giảm tải nhận thức. Những người tham gia trẻ tuổi thể hiện sự phụ thuộc cao hơn vào các công cụ AI và điểm tư duy phản biện thấp hơn so với những người tham gia lớn tuổi. Hơn nữa, trình độ học vấn cao hơn có liên quan đến kỹ năng tư duy phản biện tốt hơn, bất kể việc có sử dụng AI hay không. Việc quá phụ thuộc vào AI để phân tích có nguy cơ làm suy yếu sự phát triển của các kỹ năng phân tích của con người. Những cá nhân phụ thuộc quá nhiều vào AI để thực hiện các nhiệm vụ phân tích có thể trở nên kém thành thạo hơn trong việc tham gia vào các phân tích sâu sắc và độc lập. Sự phụ thuộc này có thể dẫn đến sự hiểu biết hời hợt về thông tin và làm giảm khả năng phân tích phản biện (Michael Gerlich, 2025).

Nhìn chung, tư duy phản biện bao gồm nhiều kỹ năng thành phần khác nhau. Kiến thức nền tảng là điều kiện cần nhưng chưa đủ để hình thành tư duy phản biện trong một lĩnh vực nhất định. Kết quả các nghiên cứu trên chỉ ra AI tạo sinh như Chat GPT có thể đóng vai trò kép: vừa là công cụ khuếch đại nhận thức, vừa tiềm ẩn các nguy cơ làm suy yếu các kỹ năng tư duy quan trọng (trong đó có tư duy phản biện). AI hỗ trợ hiệu quả các nhiệm vụ nhận thức bậc thấp như tìm kiếm, tóm tắt, hiểu khái niệm, giúp người học tiết kiệm thời gian và tăng năng suất. Thậm chí, nó còn có thể thúc đẩy kỹ năng tư duy phản biện. AI có thể tổng

hợp thông tin từ nhiều nguồn đa dạng giúp người dùng hình thành các quan điểm toàn diện, không phiên diện một chiều. Khả năng này hỗ trợ các quá trình lý luận và ra quyết định bằng cách trình bày thông tin một cách có tổ chức và dễ hiểu. Thay vì bị choáng ngợp bởi sự hỗn loạn của dữ liệu, người dùng có thể dựa vào AI để tạo ra các bản tóm tắt, biểu đồ hoặc phân tích có cấu trúc, cho phép họ tập trung vào việc diễn giải và áp dụng thông tin hơn là chỉ sắp xếp nó.

Mặc dù AI mang lại nhiều lợi ích, nhưng sự tích hợp sâu rộng của nó cũng tiềm ẩn những rủi ro đáng kể đối sự phát triển và duy trì tư duy phản biện của con người. Những rủi ro này chủ yếu xuất phát từ sự phụ thuộc vào công nghệ. Một trong những mối lo ngại chính là hiện tượng giảm tải nhận thức, nơi con người ủy thác các tác vụ nhận thức cho AI, có khả năng dẫn đến sự teo rút các kỹ năng tư duy phản biện. Sự phụ thuộc quá mức vào AI để giải quyết vấn đề có thể làm giảm nhu cầu cá nhân phải tự mình tham gia vào quá trình suy nghĩ sâu sắc. AI, đặc biệt là các mô hình tạo sinh, có khả năng tạo ra nội dung thuyết phục nhưng sai lệch hoặc gây hiểu lầm. Điều này đặt ra một thách thức nghiêm trọng đối với tư duy phản biện, vì người dùng phải đối mặt với một môi trường thông tin nơi việc phân biệt sự thật với sản phẩm của thuật toán ngày càng trở nên khó khăn. Trước bối cảnh đó, các trường học không thể cấm người học sử dụng AI mà cần hướng dẫn và định hướng sử dụng AI một cách có trách nhiệm.

III. KẾT LUẬN

Nghiên cứu đã tổng hợp một số kết quả nghiên cứu mới trong những năm gần đây, bổ sung cho các luận điểm minh chứng tác động của AI đến tư

duy phản biện của học sinh, sinh viên. Bên cạnh đó, chúng tôi cũng ghi nhận một số giới hạn của nghiên cứu. Phạm vi nghiên cứu trải rộng từ học sinh đến sinh viên nhưng có xu hướng ở lứa tuổi sinh viên, nghiên cứu sâu hơn ở từng lứa tuổi sẽ hữu ích hơn cho lĩnh vực nghiên cứu về phương pháp giảng dạy ở từng cấp bậc học. Ngoài ra, nghiên cứu tập trung vào các kết quả nghiên cứu quốc tế, các nghiên cứu trong nước còn ít và chưa xem xét kỹ yếu tố văn hóa. Đây cũng là một khía cạnh cần quan tâm khi áp dụng các chính sách, phương pháp giảng dạy, phân tích mối liên hệ giữa tần suất sử dụng AI và tư duy phản biện ở học sinh, sinh viên.

Nhìn chung, AI thúc đẩy năng lực tư duy độc lập của con người khi phải xử lý lại các thông tin từ AI cung cấp do tính rủi ro học thuật, rèn luyện tư duy phản biện. Tuy nhiên, mặt còn lại, AI khiến người học trở nên lười tư duy chủ động, tăng sự ỷ lại và lười nhác, phụ thuộc vào AI, từ đó suy giảm tư duy phản biện. AI sẽ cung cấp cùng lúc khối lượng lớn thông tin bao gồm thông tin chưa được kiểm chứng khoa học, bởi vậy người học cần biết đánh giá tính xác thực của nguồn thông tin, luôn có thói quen nghi ngờ và tinh thần phản biện khi tương tác với AI. Cần phát triển kỹ năng đặt câu hỏi chi tiết, thông tin nhằm thu thập được thông tin chính xác, đầy đủ, phù hợp với bối cảnh câu hỏi. Việc đặt câu hỏi đúng và trúng giúp người học giảm tải việc phải xử lý các “thông tin rác” mà AI cung cấp. Điều này đòi hỏi sự tham gia của đội ngũ giảng viên với vai trò định hướng, chỉ dẫn cách làm chủ AI. Đó là lý do AI không thể thay thế giáo viên, AI chỉ thay thế được giáo viên không biết sử dụng AI trong thời đại công nghệ số toàn cầu.

TÀI LIỆU THAM KHẢO

- Lee, Y. F., Hwang, G. J., & Cheng, L. C. (2025). Impacts of a ChatGPT-supported concept mapping approach on students' database programming achievement and their problem-solving and critical thinking awareness. *Interactive Learning Environments*, 1–20.
- Beverly Park Woolf, “AI and Education: Celebrating 30 years of Marriage”, *Seventeenth International Conference on Artificial Intelligence in Education (AIED 2015 Workshop Proceedings)*, Vol 4, 2015, 85 (38-45), <http://users.sussex.ac.uk/~bend/aied2015/>
- Lee, H. P., Sarkar, A., Tankelevitch, L., Drosos, I., Rintel, S., Banks, R., & Wilson, N. (2025, April). The impact of generative AI on critical thinking: Self-reported reductions in cognitive effort and confidence effects from a survey of knowledge workers. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems* (pp. 1–22).
- Gerlich, M. (2025). AI tools in society: Impacts on cognitive offloading and the future of critical thinking. *Societies*, 15(1), 6.
- Wang, J., & Fan, W. (2025). The effect of ChatGPT on students' learning performance, learning perception, and higher-order thinking: Insights from a meta-analysis. *Humanities and Social Sciences Communications*, 12(1), 1–21.
- Fan, Y., Tang, L., Le, H., Shen, K., Tan, S., Zhao, Y., ... & 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.
- Urban, M., Děchtěrenko, F., Lukavský, J., Hrabalová, V., Svacha, F., Brom, C., & Urban, K. (2024). ChatGPT improves creative problem-solving performance in university students: An experimental study. *Computers & Education*, 215, 105031.