

ỨNG DỤNG TRÍ TUỆ NHÂN TẠO TRONG DẠY HỌC MÔN TIẾNG ANH: NGHIÊN CỨU THỰC TRẠNG TẠI MỘT SỐ TRƯỜNG TRUNG HỌC PHỔ THÔNG CÔNG LẬP THÀNH PHỐ HÀ NỘI

ARTIFICIAL INTELLIGENCE APPLICATION IN ENGLISH LANGUAGE TEACHING: A SURVEY AT SELECTED
PUBLIC HIGH SCHOOLS IN HANOI

Hà Văn Dũng¹,
Phạm Bảo Anh^{2,+}

¹Viện Chiến lược và Chính sách Giáo dục - Bộ Giáo dục và Đào tạo;

²Trường Trung học phổ thông Việt Đức, thành phố Hà Nội

+Tác giả liên hệ • Email: phambaoanh3110@gmail.com

Article history

Received: 07/6/2026

Accepted: 24/8/2026

Published: 20/9/2026

Keywords

Artificial intelligence,
English language teaching,
English teachers, high school
students

ABSTRACT

The rapid development of artificial intelligence (AI), particularly generative AI, is expanding opportunities to support foreign language teaching and learning. However, evidence on the extent of AI use in Vietnamese public high schools remains limited, especially studies that simultaneously examine both teachers and students. This study employed a cross-sectional quantitative survey with descriptive and comparative components, drawing on data from 62 English teachers and 375 students to examine the current use of AI-enabled tools in English language teaching and learning at five public high schools in Hanoi. Data were analysed using descriptive statistics, Cronbach's α , McDonald's ω , Welch's ANOVA, and Games-Howell post hoc tests. The findings indicate that AI-enabled tools were used relatively widely, with ChatGPT, Google Translate, and Gemini among the most frequently selected tools. Teachers reported high levels of awareness, self-assessed competence, facilitating conditions, and perceived effectiveness, whereas their level of AI use in teaching activities remained moderate. Among students, self-assessed competence and conditions for AI use received the highest mean scores, while difficulties, risks, and support needs were rated at a moderate level. Frequency of AI use was associated with awareness, self-assessed competence, and perceived effectiveness; teachers with fewer years of teaching experience tended to use AI more frequently and in more diverse ways. The findings suggest that greater attention should be paid to the quality of AI use, information verification competence, academic integrity, and pedagogically appropriate AI integration.

1. Mở đầu

Sự phát triển nhanh chóng của trí tuệ nhân tạo (AI), đặc biệt là AI tạo sinh, đang làm thay đổi cách thức con người tiếp cận, tạo lập và xử lý thông tin. Trong giáo dục, AI có thể hỗ trợ thiết kế học liệu, xây dựng nhiệm vụ, cung cấp phản hồi và cá nhân hóa học tập. Tuy nhiên, giá trị giáo dục của AI phụ thuộc không chỉ vào khả năng kỹ thuật của công cụ mà còn vào mục tiêu sử dụng, năng lực người dùng và cách thức tích hợp vào dạy học. Vì vậy, khung năng lực AI của UNESCO dành cho GV nhấn mạnh tư duy lấy con người làm trung tâm, đạo đức AI, nền tảng và ứng dụng AI, sự phạm AI và phát triển nghề nghiệp; khung dành cho HS chú trọng đánh giá phê phán, sử dụng AI an toàn, có trách nhiệm và duy trì quyền tự chủ của người học (Miao và Cukurova, 2024; Miao và cộng sự, 2024).

Trong dạy học môn Tiếng Anh, AI có nhiều khả năng ứng dụng do đặc trưng của môn học gắn với xử lý ngôn ngữ, giao tiếp, phản hồi và luyện tập cá nhân. GV có thể sử dụng AI để tạo học liệu, xây dựng câu hỏi, thiết kế tình huống giao tiếp và hỗ trợ phản hồi; HS có thể dùng AI để học từ vựng, luyện ngữ pháp, viết, hội thoại, phát âm, dịch thuật và chuẩn bị thuyết trình. Các nghiên cứu tổng quan gần đây cho thấy, AI có thể hỗ trợ cá nhân hóa, phản hồi viết, luyện phát âm, tương tác ngôn ngữ và tự học, đồng thời liên quan đến động cơ, thái độ, mức độ tham gia và lo âu của người học ngoại ngữ (AlTwijri và Alghizzi, 2024; Crompton và cộng sự, 2024; Peña-Acuña và Corga Fernandes Durão, 2024). Tuy nhiên, sử dụng AI không đồng nghĩa với tích hợp AI có hiệu quả. GV vẫn cần lựa chọn công cụ theo mục tiêu ngôn ngữ, kiểm chứng đầu ra và thiết kế nhiệm vụ bảo đảm HS tham gia tự duy. Nghiên cứu

của An và cộng sự (2023) cho thấy ý định sử dụng AI của GV Tiếng Anh trung học có liên quan đến kì vọng hiệu quả, điều kiện hỗ trợ, tri thức công nghệ AI, tri thức sư phạm với AI và năng lực tích hợp AI; điều này cho thấy nhận thức tích cực chưa đủ để tạo ra hành vi sử dụng thường xuyên và có định hướng sư phạm (An và cộng sự, 2023). Bên cạnh cơ hội, AI tạo sinh đặt ra rủi ro về độ chính xác, thiên lệch, quyền riêng tư, đạo đức học thuật và sự phụ thuộc của người học. Mô hình ngôn ngữ lớn có thể tạo ra nội dung trôi chảy nhưng thiếu chính xác hoặc không minh bạch về nguồn; trong học tập, AI có thể hỗ trợ gợi ý và phản hồi nhưng cũng có thể bị sử dụng để làm thay nhiệm vụ. Vì vậy, UNESCO khuyến nghị cách tiếp cận lấy con người làm trung tâm, bảo vệ quyền tự chủ, dữ liệu cá nhân và bảo đảm sự phù hợp về độ tuổi, đạo đức và mục tiêu sư phạm khi sử dụng AI tạo sinh trong giáo dục (Kasneci và cộng sự, 2023; Miao và Holmes, 2023).

Tại Việt Nam, chuyển đổi số trong giáo dục đã được xác định là một định hướng quan trọng trong Quyết định số 131/QĐ-TTg ngày 25/01/2022 của Thủ tướng Chính phủ (2022). Trong lĩnh vực ngoại ngữ, Chiêm Thị Anh Thư và Nguyễn Huỳnh Trang (2024) nghiên cứu việc sử dụng Duolingo trong học từ vựng của HS THPT, Nguyễn Văn Minh (2025) phân tích tiềm năng của ChatGPT trong giảng dạy các kĩ năng tiếng Anh, Phạm và Dang (2025) khảo sát việc sử dụng AI trong dạy và học ngoại ngữ tại Việt Nam. Các nghiên cứu này cho thấy, AI ngày càng hiện diện trong giáo dục ngoại ngữ, nhưng chủ yếu tập trung vào một công cụ, một nhóm đối tượng hoặc bối cảnh giáo dục đại học.

Trong bài báo này, “ứng dụng AI trong dạy học môn Tiếng Anh” được hiểu theo nghĩa thao tác là việc GV và HS sử dụng các công cụ số có tích hợp hoặc khai thác chức năng dựa trên AI để hỗ trợ dạy và học tiếng Anh. Các công cụ được khảo sát gồm AI tạo sinh hoặc trợ lý đa chức năng như ChatGPT, Gemini và Microsoft Copilot; dịch máy như Google Translate; hỗ trợ viết như Grammarly và QuillBot; luyện phát âm như ELSA Speak; và hỗ trợ học ngôn ngữ như Duolingo. Nghiên cứu không xem các công cụ này là tương đương về công nghệ hoặc chức năng, mà khảo sát mức độ tự báo cáo về việc sử dụng chúng trong các hoạt động dạy và học.

Trên cơ sở đó, bài báo nhằm khảo sát thực trạng ứng dụng AI trong dạy và học môn Tiếng Anh tại một số trường THPT công lập ở Hà Nội, tập trung vào ba câu hỏi: (1) GV và HS nhận thức, đánh giá như thế nào về AI trong dạy và học môn Tiếng Anh?; (2) AI được sử dụng ở mức độ nào, thông qua những công cụ và hình thức nào?; (3) Các phương diện nhận thức, năng lực tự đánh giá, hình thức sử dụng, hiệu quả cảm nhận và nhu cầu hỗ trợ khác nhau như thế nào giữa các nhóm GV theo thâm niên và giữa các nhóm GV, HS theo tần suất sử dụng AI?. Bài báo kì vọng bổ sung bằng chứng thực tiễn và cung cấp căn cứ cho hoạt động bồi dưỡng, hướng dẫn sử dụng AI an toàn, có trách nhiệm và phù hợp với mục tiêu môn học.

2. Phương pháp nghiên cứu

2.1. Thiết kế nghiên cứu

Nghiên cứu sử dụng thiết kế khảo sát định lượng cắt ngang theo hướng mô tả và so sánh. Trong thiết kế cắt ngang, dữ liệu được thu thập từ một quần thể hoặc mẫu nghiên cứu tại một thời điểm hoặc trong một khoảng thời gian xác định nhằm mô tả đặc điểm của các biến và xem xét sự khác biệt hoặc mối liên hệ giữa các nhóm tại thời điểm khảo sát (Setia, 2016). Thiết kế này phù hợp với mục tiêu của nghiên cứu hiện tại vì nghiên cứu tập trung mô tả thực trạng nhận thức, năng lực, điều kiện, mức độ và hình thức sử dụng AI trong dạy và học môn Tiếng Anh, đồng thời so sánh các nhóm GV và HS theo một số đặc điểm được xác định trước. Do dữ liệu được thu thập theo thiết kế cắt ngang, các kết quả so sánh chỉ phản ánh sự khác biệt hoặc mối liên hệ giữa các nhóm tại thời điểm khảo sát, không được sử dụng để suy luận quan hệ nhân quả.

Dữ liệu được thu thập bằng hai phiếu hỏi có cấu trúc dành riêng cho GV và HS. Nghiên cứu không sử dụng phỏng vấn, quan sát lớp học hoặc câu hỏi mở. Vì vậy, các kết quả phản ánh nhận thức và đánh giá tự báo cáo của người tham gia, không phải phép đo trực tiếp về năng lực sử dụng AI hoặc kết quả học tập môn Tiếng Anh.

2.2. Bối cảnh, khách thể và mẫu nghiên cứu

Nghiên cứu được thực hiện tại 05 trường THPT công lập trên địa bàn TP. Hà Nội, gồm: THPT Thăng Long, THPT Trần Nhân Tông, THPT Trần Phú, THPT Việt Đức và THPT Đoàn Kết - Hai Bà Trưng. Việc lựa chọn các trường được thực hiện theo phương pháp có chủ đích, phù hợp với phạm vi và điều kiện triển khai nghiên cứu, đồng thời cho phép tiếp cận hai nhóm khách thể có liên quan trực tiếp đến quá trình ứng dụng AI trong dạy và học môn Tiếng Anh là GV Tiếng Anh và HS ở cả ba khối lớp 10, 11 và 12. Việc thu thập dữ liệu tại nhiều trường thay vì chỉ tại một đơn vị nhằm tạo sự đa dạng nhất định về bối cảnh trường học trong phạm vi khảo sát và hạn chế khả năng kết quả chỉ phản ánh đặc điểm riêng của một trường. Tuy nhiên, do các trường không được lựa chọn bằng phương

pháp xác suất để đại diện cho toàn bộ các trường THPT công lập tại Hà Nội, kết quả nghiên cứu chỉ được diễn giải trong phạm vi các trường được khảo sát và không được suy rộng cho toàn bộ thành phố.

Mẫu nghiên cứu thực tế gồm 62 GV Tiếng Anh (tất cả GV của 05 trường) và 375 HS, với tổng số 437 người tham gia. GV là những người đang trực tiếp giảng dạy môn Tiếng Anh tại 05 trường khảo sát. Mẫu HS được phân bổ tương đối đồng đều giữa các trường và ba khối lớp 10, 11, 12. Số HS tại mỗi trường dao động từ 71 đến 77; số HS ở mỗi khối dao động từ 120 đến 130. Người tham gia được lựa chọn theo phương pháp chọn mẫu có chủ đích kết hợp phân tầng theo trường và khối lớp. Tiêu chí lựa chọn gồm đang giảng dạy hoặc học tập tại một trong 05 trường thuộc phạm vi nghiên cứu và đồng ý tham gia khảo sát. Công thức Cochran (1977) được sử dụng như một căn cứ tham khảo để định hướng quy mô mẫu HS trong điều kiện chưa xác định chính xác tham số của tổng thể. Với mức tin cậy 95%, $p=0,5$ và sai số ước lượng 0,05, quy mô mẫu tham khảo đối với tổng thể lớn là khoảng 384 người. Mẫu thực tế gồm 375 HS. Tuy nhiên, do các trường và người tham gia không được lựa chọn hoàn toàn bằng phương pháp xác suất, giá trị 384 chỉ được sử dụng để tham khảo về quy mô mẫu, không phải căn cứ khẳng định tính đại diện của mẫu đối với toàn bộ HS THPT tại Hà Nội.

2.3. Khung lý thuyết xây dựng công cụ khảo sát

Công cụ khảo sát được phát triển trên cơ sở tích hợp ba hướng tiếp cận chính và các nghiên cứu thực nghiệm liên quan: Thứ nhất, khung năng lực AI của UNESCO dành cho GV và HS cung cấp cơ sở để xác định các nội dung về nhận thức, khả năng sử dụng, kiểm chứng đầu ra, đạo đức, trách nhiệm và duy trì quyền tự chủ khi sử dụng AI (Miao và Cukurova, 2024; Miao và cộng sự, 2024); Thứ hai, mô hình UTAUT gợi ý xem xét đồng thời kì vọng hiệu quả, mức độ dễ sử dụng và điều kiện hỗ trợ, qua đó phân biệt nhận thức tích cực với hành vi sử dụng thực tế (Venkatesh và cộng sự, 2003; An và cộng sự, 2023); Thứ ba, khung TPACK nhấn mạnh rằng năng lực công nghệ chỉ có ý nghĩa sư phạm khi được kết nối với nội dung môn học và phương pháp dạy học (Mishra và Koehler, 2006).

Các nghiên cứu về AI trong dạy học ngoại ngữ cho thấy công cụ có thể được sử dụng cho tạo nội dung, phản hồi, luyện phát âm, hỗ trợ viết, hội thoại, dịch thuật và cá nhân hóa học tập. Đồng thời, hiệu quả phụ thuộc vào mục tiêu, thiết kế nhiệm vụ và mức độ tham gia chủ động của người học (AITwijri và Alghizzi, 2024; Crompton và cộng sự, 2024; Peña-Acuña và Corra Fernandes Durão, 2024). Các nghiên cứu và hướng dẫn về AI tạo sinh cũng chỉ ra những vấn đề như thông tin thiếu chính xác, thiên lệch, quyền riêng tư, gian lận học thuật và phụ thuộc vào công cụ. Do đó, hiệu quả cảm nhận cần được khảo sát đồng thời với khả năng kiểm chứng, nhận thức rủi ro và nhu cầu hỗ trợ (Kasneji và cộng sự, 2023; Miao và Holmes, 2023).

Từ các cơ sở trên, nghiên cứu xác định 06 miền nội dung lý thuyết và nội dung vận hành hóa công cụ khảo sát (bảng 1):

Bảng 1. Cơ sở lý thuyết và nội dung vận hành hóa công cụ khảo sát

Miền nội dung	Cơ sở lý thuyết	Chỉ báo đối với GV	Chỉ báo đối với HS
Nhận thức và thái độ	UNESCO; UTAUT	Vai trò, lợi ích, giới hạn và sự sẵn sàng sử dụng AI	Lợi ích, giới hạn và sự sẵn sàng sử dụng AI để học tiếng Anh
Năng lực và điều kiện	UNESCO; UTAUT	Lựa chọn công cụ, viết câu lệnh, kiểm chứng; thiết bị, Internet và hỗ trợ	Tra cứu, luyện tập, kiểm tra đầu ra; thiết bị và Internet
Mức độ sử dụng	UTAUT; nghiên cứu AI trong ELT	Tần suất chuẩn bị bài, tạo học liệu, phản hồi và đánh giá	Tần suất học từ vựng, ngữ pháp, viết, nói, phát âm và dịch
Hình thức và mục đích	TPACK; nghiên cứu AI trong ELT	Thiết kế bài dạy, câu hỏi, tình huống giao tiếp và đánh giá	Luyện viết, hội thoại, phát âm, dịch thuật và thuyết trình
Hiệu quả cảm nhận	UTAUT; nghiên cứu AI trong giáo dục	Tiết kiệm thời gian, chất lượng học liệu, cá nhân hóa và tổ chức dạy học	Hứng thú, tự học, phản hồi và cải thiện kỹ năng
Khó khăn, rủi ro và hỗ trợ	UNESCO; nghiên cứu về GenAI	Kiểm chứng, kiểm soát sản phẩm, hướng dẫn và bồi dưỡng	Phụ thuộc, thông tin sai, gian lận, bảo mật và nhu cầu hướng dẫn

06 miền nội dung trên thuộc tầng lý thuyết. Khi phân tích dữ liệu, các biến được tổ chức thành 05 thang đo đối với mỗi nhóm khách thể theo cấu trúc nội dung của phiếu. Ở GV, mức độ sử dụng và hình thức, mục đích được phân tích thành hai thang riêng. Ở HS, hai miền này được kết hợp thành một thang, còn khó khăn, rủi ro và nhu cầu hỗ trợ được hình thành thành một thang riêng. Các khung lý thuyết và quá trình tham khảo chuyên gia chủ yếu hỗ trợ giá trị nội dung của công cụ; còn Cronbach's α , McDonald's ω và tương quan biến - tổng được sử dụng để đánh giá độ tin

cây nội tại của các thang đo trong mẫu nghiên cứu. Nghiên cứu chưa thực hiện phân tích nhân tố khám phá hoặc khẳng định; do đó, chưa đưa ra kết luận về giá trị cấu trúc của mô hình đo lường.

2.4. Công cụ thu thập dữ liệu

Nghiên cứu sử dụng phiếu khảo sát có cấu trúc, một phiếu dành cho GV Tiếng Anh và một phiếu dành cho HS THPT. Mỗi phiếu gồm 02 phần, cụ thể: Phần thứ nhất của phiếu GV thu thập thông tin về đơn vị công tác, giới tính, thâm niên, trình độ đào tạo, mức độ sử dụng và các công cụ AI đã sử dụng; Phần thứ hai gồm 32 biến quan sát, được phân tích thành năm thang đo: nhận thức và thái độ; năng lực và điều kiện; mức độ sử dụng; hình thức và mục đích; hiệu quả cảm nhận. Phiếu khảo sát được thiết kế trên Google Forms và gửi tới GV theo link: <https://forms.gle/FiZTipPLoD4Deuao6>. Phần thứ nhất của phiếu HS thu thập thông tin về trường, giới tính, khối lớp, điều kiện thiết bị và Internet, mức độ sử dụng và các công cụ AI; Phần thứ hai gồm 32 biến quan sát, được phân tích thành năm thang đo: nhận thức và thái độ; năng lực và điều kiện; mức độ và hình thức sử dụng; hiệu quả cảm nhận; khó khăn, rủi ro và nhu cầu hỗ trợ. Phiếu khảo sát được thiết kế trên Google Forms và gửi đến HS theo link: <https://forms.gle/kz6FftR9GrrFaeqc7>.

Việc tham gia khảo sát được thực hiện trên cơ sở tự nguyện và ẩn danh; phiếu hỏi không thu thập thông tin định danh cá nhân và dữ liệu chỉ được sử dụng cho mục đích nghiên cứu. Tuy nhiên, nghiên cứu không thực hiện thủ tục xin phép chính thức của nhà trường hoặc thu thập sự đồng thuận riêng của cha mẹ/người giám hộ đối với từng HS tham gia. Đây là một hạn chế trong quy trình đạo đức nghiên cứu đối với người tham gia chưa thành niên và sẽ được khắc phục trong các nghiên cứu tiếp theo.

Các biến quan sát sử dụng thang Likert 5 mức. Với các phát biểu về mức độ đồng ý, điểm số được quy ước từ 1 - hoàn toàn không đồng ý đến 5 - hoàn toàn đồng ý. Với các phát biểu về tần suất, điểm số được quy ước từ 1 - không bao giờ đến 5 - rất thường xuyên. Trước khi khảo sát chính thức, phiếu hỏi được tham khảo ý kiến của 14 GV Tiếng Anh tại một trường THPT - nơi tác giả đang công tác. Các GV góp ý về mức độ rõ ràng của cách diễn đạt, sự phù hợp của nội dung với bối cảnh dạy và học tiếng Anh ở THPT và khả năng hiểu của người trả lời. Trên cơ sở góp ý, một số mục hỏi được rà soát và điều chỉnh cách diễn đạt trước khi triển khai khảo sát chính thức. Nghiên cứu không tiến hành khảo sát thử độc lập.

2.5. Quy trình thu thập và phân tích dữ liệu

Quy trình nghiên cứu gồm 04 bước: (1) Tổng quan tài liệu và xây dựng khung lý thuyết; (2) Xác định thành tố, xây dựng chỉ báo và thiết kế phiếu; (3) Tham khảo ý kiến GV có kinh nghiệm/chuyên môn phù hợp và tiến hành khảo sát; (4) Làm sạch, mã hóa và phân tích dữ liệu bằng phần mềm JASP. Đây là phần mềm thống kê có giao diện đồ họa, hỗ trợ các phân tích mô tả, kiểm định so sánh và nhiều mô hình thống kê thông dụng (Love và cộng sự, 2019).

Tần suất và tỉ lệ phần trăm được sử dụng để mô tả đặc điểm mẫu, mức độ sử dụng và các công cụ AI. Điểm trung bình (ĐTB), độ lệch chuẩn (SD), giá trị nhỏ nhất và lớn nhất được sử dụng để mô tả các thang đo. Điểm thang đo được tính bằng trung bình cộng của các biến quan sát trong cùng nhóm. Độ tin cậy nội tại được kiểm tra bằng Cronbach's α , McDonald's ω và tương quan biến - tổng. Đối với phiếu GV, GV26 được đảo chiều trước khi tính điểm. Những biến có tương quan biến - tổng thấp được xem xét đồng thời dựa trên sự thay đổi hệ số tin cậy khi loại biến và ý nghĩa nội dung của biến.

Sự khác biệt giữa các nhóm được kiểm định bằng Welch ANOVA. Khi kết quả tổng thể có ý nghĩa thống kê, hậu kiểm Games-Howell được sử dụng để xác định các cặp nhóm khác biệt. Quy mô ảnh hưởng được báo cáo bằng ω^2 , với mức ý nghĩa thống kê ($p < 0,05$).

Đối với HS, nhóm "chưa từng sử dụng AI" chỉ có một trường hợp nên được giữ trong thống kê mô tả nhưng không đưa vào phân tích ANOVA. Phân tích so sánh theo mức độ sử dụng AI được thực hiện với 374 HS thuộc 04 nhóm còn lại.

ĐTB được diễn giải theo các khoảng: 1,00-1,80: rất thấp; 1,81-2,60: thấp; 2,61-3,40: trung bình; 3,41-4,20: cao; 4,21-5,00: rất cao.

3. Kết quả nghiên cứu

3.1. Thực trạng ứng dụng AI trong giảng dạy của giáo viên

3.1.1. Đặc điểm mẫu và các công cụ AI được sử dụng

Mẫu GV gồm 62 người, chủ yếu là nữ (88,7%) và có trình độ thạc sĩ (69,4%). Khoảng một nửa GV (51,6%) cho biết sử dụng AI thường xuyên hoặc rất thường xuyên. ChatGPT, Google Translate và Gemini là ba công cụ được lựa chọn nhiều nhất.

3.1.2. Độ tin cậy và thống kê mô tả

Bảng 2. Độ tin cậy và thống kê mô tả các thang đo GV

Thang đo	Biến	Cronbach's α	McDonald's ω	Tương quan biến - tổng	M	SD	Mức
Nhận thức và thái độ	GV1-GV6	0,849	0,847	0,540-0,721	3,960	0,576	Cao
Năng lực và điều kiện	GV7-GV12	0,921	0,922	0,718-0,822	4,110	0,581	Cao
Mức độ sử dụng	GV13-GV20	0,883	0,895	0,202-0,765	3,224	0,761	Trung bình
Hình thức và mục đích	GV21-GV25	0,886	0,887	0,667-0,779	3,790	0,642	Cao
Hiệu quả cảm nhận	GV26-GV32	0,812	0,830	0,292-0,789	4,111	0,495	Cao

Các thang đo GV có độ tin cậy nội tại tốt, với Cronbach's α dao động từ 0,812 đến 0,921 và McDonald's ω từ 0,830 đến 0,922. Phần lớn tương quan biến - tổng đạt mức phù hợp. Riêng GV13 có tương quan biến - tổng thấp hơn các biến còn lại (0,202), nhưng được giữ lại do việc loại biến không làm cải thiện đáng kể độ tin cậy chung và nội dung biến vẫn phù hợp với phạm vi của thang đo. GV26 được đảo chiều trước khi tính điểm tổng hợp. Về thống kê mô tả, hiệu quả cảm nhận có ĐTB cao nhất ($M=4,111$; $SD=0,495$), gần tương đương với năng lực và điều kiện ($M=4,110$; $SD=0,581$). Nhận thức và thái độ ($M=3,960$; $SD=0,576$) và hình thức, mục đích sử dụng ($M=3,790$; $SD=0,642$) đều ở mức cao. Trong khi đó, mức độ sử dụng có ĐTB thấp nhất ($M=3,224$; $SD=0,761$), thuộc mức trung bình. Kết quả cho thấy, có sự khác biệt giữa đánh giá tích cực của GV đối với AI và mức độ sử dụng thực tế được tự báo cáo trong các hoạt động dạy học.

3.1.3. Sự khác biệt theo thâm niên giảng dạy

Bảng 3. So sánh các thang đo GV theo thâm niên

Thang đo	Dưới 5 năm M (SD)	5≤10 năm M (SD)	10≤20 năm M (SD)	≥20 năm M (SD)	Welch (F)	p	ω^2
Nhận thức và thái độ	4,148 (0,903)	4,098 (0,503)	3,848 (0,510)	3,688 (0,383)	2,118	0,128	0,032
Năng lực và điều kiện	4,352 (0,823)	4,220 (0,647)	4,007 (0,428)	3,833 (0,321)	1,971	0,148	0,032
Mức độ sử dụng	4,014 (0,743)	3,352 (0,673)	3,071 (0,581)	2,422 (0,571)	8,646	<0,001	0,292
Hình thức và mục đích	4,356 (0,817)	3,945 (0,483)	3,583 (0,575)	3,325 (0,465)	5,375	0,007	0,202
Hiệu quả cảm nhận	3,937 (0,682)	4,201 (0,431)	4,130 (0,524)	4,000 (0,324)	0,793	0,511	<0,001

Không ghi nhận sự khác biệt có ý nghĩa thống kê giữa các nhóm thâm niên đối với nhận thức và thái độ, năng lực và điều kiện, hoặc hiệu quả cảm nhận ($p>0,05$). Ngược lại, mức độ sử dụng khác biệt có ý nghĩa theo thâm niên, $F(3;20,86)=8,646$, $p<0,001$, $\omega^2=0,292$. Hậu kiểm Games-Howell cho thấy, GV dưới 5 năm kinh nghiệm có điểm cao hơn nhóm 10≤20 năm ($MD=0,943$; $p=0,023$) và nhóm từ 20 năm trở lên ($MD=1,592$; $p<0,001$); nhóm 5≤10 năm cũng có điểm cao hơn nhóm từ 20 năm trở lên ($MD=0,930$; $p=0,009$). Hình thức và mục đích sử dụng AI cũng khác biệt theo thâm niên với $F(3;20,55)=5,375$, $p=0,007$, $\omega^2=0,202$. Nhóm dưới 5 năm có điểm cao hơn nhóm từ 20 năm trở lên ($MD=1,031$; $p=0,029$), và nhóm 5≤10 năm cao hơn nhóm từ 20 năm trở lên ($MD=0,621$; $p=0,031$). Nhìn chung, sự khác biệt theo thâm niên tập trung vào hành vi và phạm vi sử dụng AI, trong khi không ghi nhận khác biệt rõ về nhận thức, năng lực tự đánh giá và hiệu quả cảm nhận.

3.1.4. Sự khác biệt theo mức độ sử dụng AI

Bảng 4. So sánh các thang đo GV theo mức độ sử dụng AI

Thang đo	Không thường xuyên M (SD)	Thỉnh thoảng M (SD)	Thường xuyên M (SD)	Rất thường xuyên M (SD)	Welch (F)	p	ω^2
Nhận thức và thái độ	3,556 (0,499)	3,639 (0,543)	4,186 (0,360)	4,750 (0,253)	20,980	<0,001	0,407
Năng lực và điều kiện	3,847 (0,524)	3,750 (0,611)	4,359 (0,394)	4,639 (0,427)	7,993	0,001	0,274
Mức độ sử dụng	2,990 (0,684)	2,889 (0,745)	3,332 (0,647)	4,229 (0,496)	9,058	<0,001	0,215
Hình thức và mục đích	3,367 (0,489)	3,489 (0,583)	3,969 (0,476)	4,767 (0,388)	16,730	<0,001	0,377
Hiệu quả cảm nhận	3,857 (0,360)	3,897 (0,451)	4,286 (0,396)	4,500 (0,755)	5,091	0,010	0,178

Tần suất sử dụng AI có liên hệ với sự khác biệt ở các thang nhận thức và thái độ, năng lực và điều kiện, hình thức và mục đích, và hiệu quả cảm nhận của GV ($p < 0,05$). Nhìn chung, các nhóm sử dụng thường xuyên hoặc rất thường xuyên có điểm nhận thức, năng lực tự đánh giá và hiệu quả cảm nhận cao hơn các nhóm sử dụng ít hơn. Quy mô ảnh hưởng lớn nhất ghi nhận ở nhận thức và thái độ ($\omega^2 = 0,407$) và hình thức, mục đích sử dụng ($\omega^2 = 0,377$), tiếp theo là năng lực và điều kiện ($\omega^2 = 0,274$) và hiệu quả cảm nhận ($\omega^2 = 0,178$).

3.2. Thực trạng ứng dụng AI trong học tập của học sinh

3.2.1. Đặc điểm mẫu và các công cụ AI được sử dụng

Trong 375 HS, ba khối lớp được phân bố tương đối đồng đều. Có 60,5% HS có thiết bị riêng và Internet ổn định; 39,5% còn lại có ít nhất một hạn chế về thiết bị hoặc kết nối. Có 51,2% HS cho biết sử dụng AI thường xuyên hoặc rất thường xuyên. Google Translate, ChatGPT và Gemini là các công cụ phổ biến nhất.

3.2.2. Độ tin cậy và thống kê mô tả

Bảng 5. Độ tin cậy và thống kê mô tả các thang đo HS

Thang đo	Biến	Cronbach's α	McDonald's ω	Tương quan biến - tổng	M	SD	Mức
Nhận thức và thái độ	HS1-HS6	0,718	0,722	0,377-0,530	3,458	0,718	Cao
Năng lực và điều kiện	HS7-HS12	0,731	0,738	0,312-0,560	3,607	0,721	Cao
Mức độ và hình thức sử dụng	HS13-HS20	0,696	0,692	0,311-0,466	3,440	0,653	Cao
Hiệu quả cảm nhận	HS21-HS25	0,657	0,664	0,338-0,473	3,460	0,728	Cao
Khó khăn, rủi ro và nhu cầu hỗ trợ	HS26-HS32	0,662	0,661	0,246-0,431	3,291	0,643	Trung bình

02 thang đo nhận thức và thái độ, năng lực và điều kiện có hệ số tin cậy trên 0,70. 03 thang còn lại có độ tin cậy ở mức chấp nhận được đối với nghiên cứu khám phá. HS32 có tương quan biến - tổng thấp nhất (0,246), tuy nhiên việc loại biến không cải thiện đáng kể Cronbach's α nên biến được giữ lại.

Về thống kê mô tả, năng lực và điều kiện có ĐTB cao nhất ($M = 3,607$; $SD = 0,721$). Hiệu quả cảm nhận ($M = 3,460$; $SD = 0,728$), nhận thức và thái độ ($M = 3,458$; $SD = 0,718$), mức độ và hình thức sử dụng ($M = 3,440$; $SD = 0,653$) đều ở mức cao nhưng nằm gần ngưỡng phân loại. Khó khăn, rủi ro và nhu cầu hỗ trợ có ĐTB thấp nhất ($M = 3,291$; $SD = 0,643$), thuộc mức trung bình.

3.2.3. Sự khác biệt theo mức độ sử dụng AI

Bảng 6. So sánh các thang đo HS theo mức độ sử dụng AI

Thang đo	Không thường xuyên M (SD)	Thỉnh thoảng M (SD)	Thường xuyên M (SD)	Rất thường xuyên M (SD)	Welch (F)	p	ω^2
Nhận thức và thái độ	3,077 (0,727)	3,338 (0,629)	3,626 (0,693)	3,766 (0,661)	14,28	<0,001	0,102
Năng lực và điều kiện	3,305 (0,677)	3,409 (0,681)	3,858 (0,703)	3,798 (0,623)	14,94	<0,001	0,102
Mức độ và hình thức sử dụng	3,148 (0,542)	3,259 (0,656)	3,617 (0,614)	3,726 (0,621)	17,12	<0,001	0,109
Hiệu quả cảm nhận	3,095 (0,720)	3,366 (0,732)	3,625 (0,688)	3,684 (0,640)	11,06	<0,001	0,075
Khó khăn, rủi ro và nhu cầu hỗ trợ	3,097 (0,653)	3,313 (0,718)	3,308 (0,608)	3,426 (0,505)	3,373	0,020	0,016

Tần suất sử dụng AI có liên hệ với sự khác biệt có ý nghĩa ở nhận thức và thái độ, năng lực và điều kiện, hiệu quả cảm nhận, cũng như khó khăn, rủi ro và nhu cầu hỗ trợ ($p < 0,05$). Ở thang đo nhận thức và thái độ, năng lực và điều kiện, và hiệu quả cảm nhận, HS sử dụng AI thường xuyên hoặc rất thường xuyên nhìn chung có ĐTB cao hơn các nhóm sử dụng ít hơn. Không ghi nhận khác biệt rõ giữa nhóm thường xuyên và rất thường xuyên ở các thang này. Đối với khó khăn, rủi ro và nhu cầu hỗ trợ, nhóm sử dụng rất thường xuyên có điểm cao hơn nhóm đã từng sử dụng nhưng không thường xuyên ($MD = 0,330$; $p = 0,010$). Tuy nhiên, quy mô ảnh hưởng nhỏ ($\omega^2 = 0,016$).

3.3. Thảo luận

Kết quả cho thấy, các công cụ có yếu tố AI đã được sử dụng tương đối phổ biến trong dạy và học môn Tiếng Anh tại các trường khảo sát, trong đó ChatGPT, Google Translate và Gemini được lựa chọn nhiều. Tuy nhiên, các công cụ này có chức năng khác nhau; vì vậy, mức độ phổ biến trong nghiên cứu nên được hiểu là mức độ tiếp cận và sử dụng các công cụ có tích hợp AI theo nghĩa rộng, không đồng nghĩa với việc AI tạo sinh đã được tích hợp sâu

vào dạy học. Phát hiện này phù hợp với xu hướng gia tăng sử dụng AI trong giáo dục ngoại ngữ (Crompton và cộng sự, 2024; Phạm và Dang, 2025).

Ở GV, nhận thức, năng lực tự đánh giá, điều kiện và hiệu quả cảm nhận đều ở mức cao, trong khi mức độ sử dụng trong hoạt động dạy học chỉ ở mức trung bình. Khoảng cách này cho thấy, nhận thức tích cực chưa nhất thiết chuyển thành hành vi sử dụng thường xuyên. Việc tích hợp công nghệ còn phụ thuộc vào điều kiện hỗ trợ và khả năng kết nối công cụ với mục tiêu nội dung, phương pháp và đánh giá (Venkatesh và cộng sự, 2003; Mishra và Koehler, 2006). Thâm niên không tạo ra khác biệt rõ về nhận thức hoặc hiệu quả cảm nhận, nhưng GV có thâm niên thấp hơn có xu hướng sử dụng AI thường xuyên và đa dạng hơn; nguyên nhân của khác biệt này cần được tiếp tục kiểm chứng.

Ở HS, điều kiện tiếp cận nhìn chung khá thuận lợi nhưng chưa đồng đều, khi một bộ phận vẫn gặp hạn chế về thiết bị hoặc kết nối Internet. Vì vậy, việc tổ chức dạy học có sử dụng AI cần tính đến khả năng tiếp cận của người học. Đồng thời, các điểm số về nhận thức, năng lực và hiệu quả cảm nhận chỉ phản ánh tự đánh giá của HS, không phải bằng chứng trực tiếp về sự cải thiện năng lực tiếng Anh.

Tần suất sử dụng AI có liên hệ với nhận thức, năng lực tự đánh giá và hiệu quả cảm nhận ở cả GV và HS. Tuy nhiên, mối liên hệ này có thể diễn ra theo hai chiều: người có thái độ và năng lực tích cực hơn có thể chủ động sử dụng AI nhiều hơn, trong khi sử dụng thường xuyên cũng có thể làm tăng trải nghiệm và sự tự tin. Do nghiên cứu sử dụng thiết kế cắt ngang, nên không thể xác định chiều tác động hoặc suy luận quan hệ nhân quả.

Những phát hiện trên cho thấy việc sử dụng AI trong dạy học tiếng Anh cần tập trung vào chất lượng hơn là tần suất. Nhà trường cần có hướng dẫn về kiểm chứng thông tin, bảo vệ dữ liệu và đạo đức học thuật; GV cần được bồi dưỡng năng lực sử dụng/tích hợp AI theo mục tiêu sư phạm; HS cần được hướng dẫn sử dụng AI để hỗ trợ tư duy thay vì thay thế hoạt động học tập. Do dữ liệu nghiên cứu là dữ liệu tự báo cáo, kết quả chỉ phản ánh nhận thức và hành vi của người tham gia, không phải bằng chứng trực tiếp về chất lượng tích hợp AI hoặc tác động đến kết quả học tập.

4. Kết luận và bình luận

Kết quả nghiên cứu cho thấy, các công cụ có yếu tố AI đã được GV và HS tại 05 trường THPT công lập được khảo sát sử dụng tương đối phổ biến trong dạy và học môn Tiếng Anh. Ở GV, nhận thức, năng lực tự đánh giá, điều kiện sử dụng và hiệu quả cảm nhận ở mức cao, trong khi mức độ sử dụng trong các hoạt động dạy học ở mức trung bình. Ở HS, năng lực tự đánh giá và điều kiện sử dụng có ĐTB cao hơn các phương diện khác; đồng thời vẫn tồn tại những khó khăn, rủi ro và nhu cầu hỗ trợ liên quan đến quá trình sử dụng AI. Tần suất sử dụng AI có liên hệ với một số phương diện nhận thức, năng lực tự đánh giá và hiệu quả cảm nhận, còn GV có thâm niên thấp hơn có xu hướng sử dụng AI thường xuyên và đa dạng hơn. Tuy nhiên, các kết quả trên chỉ phản ánh nhận thức, tự đánh giá và hành vi tự báo cáo của người tham gia tại thời điểm khảo sát. Bài báo không quan sát trực tiếp việc sử dụng AI trong lớp học, không đánh giá chất lượng sản phẩm dạy học hoặc học tập có sử dụng AI và không đo tác động của AI đến năng lực tiếng Anh. Vì vậy, các phát hiện không được diễn giải như bằng chứng về hiệu quả thực tế hay quan hệ nhân quả. Một hạn chế khác của nghiên cứu là chưa thực hiện đầy đủ thủ tục cho phép và đồng thuận đối với người tham gia chưa thành niên; do đó, các nghiên cứu tiếp theo cần tuân thủ chặt chẽ hơn các yêu cầu đạo đức áp dụng cho nghiên cứu với HS. Trong phạm vi này, nghiên cứu gợi ý cần chú trọng phát triển năng lực kiểm chứng thông tin, đạo đức học thuật và khả năng tích hợp AI phù hợp với mục tiêu sư phạm. Các nghiên cứu tiếp theo sẽ kết hợp dữ liệu quan sát, phỏng vấn, sản phẩm học tập và thiết kế can thiệp để đánh giá sâu hơn chất lượng và tác động của việc sử dụng AI trong dạy học tiếng Anh.

Tuyên bố về vai trò của các tác giả: Hà Văn Dũng: Xây dựng ý tưởng nghiên cứu, Thiết kế phương pháp nghiên cứu, Giám sát, chỉ đạo quá trình nghiên cứu. Phạm Bảo Anh: Phân tích dữ liệu, Trực quan hoá dữ liệu, Viết bản thảo ban đầu. Hà Văn Dũng, Phạm Bảo Anh: Thực hiện nghiên cứu và thu thập dữ liệu, Chỉnh sửa và phản biện bản thảo.

Tuyên bố về GenAI và Quyền tác giả: Trong quá trình chuẩn bị bản thảo, các tác giả đã sử dụng ChatGPT để hỗ trợ tìm tài liệu, rà soát cấu trúc, diễn đạt học thuật và chuẩn hóa hình thức trình bày.

Tuyên bố về xung đột lợi ích: Các tác giả tuyên bố không có xung đột lợi ích.

Thông tin tài trợ: Nghiên cứu này không nhận được tài trợ từ bên ngoài.

Tài liệu tham khảo

AlTwijri, L. A., & Alghizzzi, T. M. (2024). Investigating the integration of artificial intelligence in English as foreign language classes for enhancing learners' affective factors: A systematic review. *Heliyon*, 10(10), e31053. <https://doi.org/10.1016/j.heliyon.2024.e31053>

- An, X., Chai, C. S., Li, Y., Zhou, Y., Shen, X., Zheng, C., & Chen, M. (2023). Modeling English teachers' behavioral intention to use artificial intelligence in middle schools. *Education and Information Technologies*, 28(5), 5187-5208. <https://doi.org/10.1007/s10639-022-11286-z>
- Chiêm Thị Anh Thư, Nguyễn Huỳnh Trang (2024). Khả năng ghi nhớ từ vựng và nhận thức của học sinh trung học phổ thông học tiếng Anh như ngoại ngữ khi sử dụng ứng dụng Duolingo: Nghiên cứu tại Đồng bằng sông Cửu Long. *Tạp chí Giáo dục*, 24(số đặc biệt 10), 296-302. <https://tcgd.tapchigiaoduc.edu.vn/index.php/tapchi/article/view/2627>
- Cochran, W. G. (1977). *Sampling techniques* (3rd ed.). John Wiley & Sons.
- Crompton, H., Edmett, A., Ichaporia, N., & Burke, D. (2024). AI and English language teaching: Affordances and challenges. *British Journal of Educational Technology*, 55(6), 2503-2529. <https://doi.org/10.1111/bjet.13460>
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., ... Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Love, J., Selker, R., Marsman, M., Jamil, T., Dropmann, D., Verhagen, J., Ly, A., Gronau, Q. F., Šmíra, M., Epskamp, S., Matzke, D., Wild, A., Knight, P., Rouder, J. N., Morey, R. D., & Wagenmakers, E.-J. (2019). JASP: Graphical statistical software for common statistical designs. *Journal of Statistical Software*, 88(2), 1-17. <https://doi.org/10.18637/jss.v088.i02>
- Miao, F., & Cukurova, M. (2024). *AI competency framework for teachers*. UNESCO.
- Miao, F., & Holmes, W. (2023). *Guidance for generative AI in education and research*. UNESCO.
- Miao, F., Shiohira, K., & Lao, N. (2024). *AI competency framework for students*. UNESCO.
- Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, 108(6), 1017-1054. <https://doi.org/10.1111/j.1467-9620.2006.00684.x>
- Nguyễn Văn Minh (2025). Tiềm năng và lợi ích của ChatGPT trong giảng dạy kỹ năng tiếng Anh của giảng viên. *Tạp chí Giáo dục*, 25(số đặc biệt 8), 300-305. <https://tcgd.tapchigiaoduc.edu.vn/index.php/tapchi/article/view/4323>
- Peña-Acuña, B., & Corga Fernandes Durão, R. (2024). Learning English as a second language with artificial intelligence for prospective teachers: A systematic review. *Frontiers in Education*, 9, 1490067. <https://doi.org/10.3389/feduc.2024.1490067>
- Pham, T. N., & Dang, T. X. (2025). An investigation into the application of artificial intelligence for language teaching and learning in Vietnam. *Vietnam Journal of Education*, 9(2), 265-283. <https://doi.org/10.52296/vje.2025.557>
- Setia, M. S. (2016). Methodology series module 3: Cross-sectional studies. *Indian Journal of Dermatology*, 61(3), 261-264. <https://doi.org/10.4103/0019-5154.182410>
- Thủ tướng Chính phủ (2022). *Quyết định số 131/QĐ-TTg ngày 25/01/2022 phê duyệt Đề án “Tăng cường ứng dụng công nghệ thông tin và chuyển đổi số trong giáo dục và đào tạo giai đoạn 2022-2025, định hướng đến năm 2030”*.
- Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly*, 27(3), 425-478. <https://doi.org/10.2307/30036540>