

HIỆU QUẢ HỌC TẬP TỪ VIDEO GIÁO DỤC DO AI TẠO SINH: MỘT ĐÁNH GIÁ TƯỜNG THUẬT

LEARNING EFFECTIVENESS OF AI-GENERATED EDUCATIONAL VIDEOS: A NARRATIVE REVIEW

Trần Triệu Hải^{1,2},
Nguyễn Văn Hạnh^{1,+},
Đương Thị Thùy Mai¹

¹Đại học Bách khoa Hà Nội;
²Trường Đại học Mở Hà Nội
+Tác giả liên hệ • Email: hanh.nguyenvan@hust.edu.vn

Article history

Received: 19/6/2026

Accepted: 22/7/2026

Published: 20/9/2026

Keywords

AI-generated educational
video, generated instructor,
learning effectiveness,
generative artificial
intelligence, narrative review

ABSTRACT

The development of generative artificial intelligence (AI) has introduced new ways of producing educational videos. However, evidence on whether these videos improve learning remains inconclusive. This narrative review synthesized empirical studies that examined learning outcomes from AI-generated educational videos (AIGV). Fifteen studies published between 2023 and 2026 met the inclusion criteria and were included in the synthesis. Existing studies indicated that AIGVs may produce comparable immediate learning performance to human-made videos. Some conditional advantages were reported for retention, idea generation, lipreading, and accuracy in performing procedures. However, no consistent advantage was found for transfer, long-term retention, or overall learning outcomes. Videos featuring real instructors were often rated more favorably in terms of social presence and engagement. Studies that integrated AIGV into flipped classrooms, project-based learning, or blended learning reported positive results, but they did not isolate the specific effect of the video itself. A cautious conclusion is that AIGV may be a feasible alternative medium in some contexts, but current evidence is insufficient to claim that it is generally superior.

1. Mở đầu

Trong quá trình phát triển của công nghệ giáo dục, video đã chuyển từ vị trí của một phương tiện minh họa sang vai trò của một nguồn học liệu chủ đạo (Nguyễn Văn Hạnh và Nguyễn Hữu Hợp, 2016). Trước đây, phim và tài liệu nghe nhìn chủ yếu được sử dụng để làm rõ những hiện tượng khó quan sát trực tiếp, tái hiện các sự kiện hoặc trình diễn những quy trình mà lời giảng và văn bản khó thể hiện đầy đủ. Khi công nghệ số và Internet trở nên phổ biến, video không còn bị giới hạn bởi thời gian phát sóng, không gian lớp học hoặc thiết bị chuyên dụng (Long và Hanh, 2020; Means và cộng sự, 2013; Hung và cộng sự, 2020). Người học có thể truy cập nội dung theo nhu cầu, dừng lại, xem lại và điều chỉnh tốc độ tiếp nhận. Nhờ đó, video ngày càng được tích hợp thường xuyên vào thiết kế khóa học thay vì chỉ được sử dụng như tài liệu bổ trợ (Berk, 2009; Nguyễn Văn Hạnh và Nguyễn Hữu Hợp, 2016; Kay, 2012). Vai trò này đặc biệt rõ trong E-learning và dạy học kết hợp (blended learning). Trong các khóa học trực tuyến, video thường đảm nhiệm những chức năng trước đây gắn với hoạt động trình bày trực tiếp của giảng viên, như giải thích kiến thức, minh họa ví dụ, hướng dẫn nhiệm vụ và định hướng quá trình tự học. Trong dạy học kết hợp, video giúp chuyển một phần nội dung ra ngoài thời gian học trực tiếp, tạo điều kiện để giờ học trên lớp được dành nhiều hơn cho thảo luận, thực hành và giải quyết vấn đề. Sự phát triển của khóa học trực tuyến đại chúng, lớp học đảo ngược và hệ thống quản lý học tập vì thế làm cho chất lượng video trở thành một yếu tố có ảnh hưởng trực tiếp đến trải nghiệm học tập số (Guo và cộng sự, 2014; Kay, 2012; Long và Hanh, 2020; Means và cộng sự, 2013; Hung và cộng sự, 2020). Tuy nhiên, giá trị giáo dục của video không xuất phát đơn thuần từ việc nội dung được trình bày bằng hình ảnh chuyển động. Theo lý thuyết học tập đa phương tiện, người học có thể xử lý và tích hợp thông tin tốt hơn khi nội dung ngôn ngữ và hình ảnh được tổ chức phù hợp với khả năng nhận thức. Ngược lại, video chứa quá nhiều chi tiết, trình bày thiếu mạch lạc hoặc không tạo cơ hội cho người học xử lý thông tin có thể làm tăng tải nhận thức (Mayer, 2014). Vì vậy, video chỉ phát huy tác dụng khi hình thức thể hiện, trình tự nội dung, nhịp độ trình bày và hoạt động học tập đi kèm được thiết kế có chủ đích. Hiệu quả học tập phụ thuộc vào chất lượng thiết kế sư phạm nhiều hơn là bản thân định dạng video.

Sự xuất hiện của trí tuệ nhân tạo (AI) tạo sinh làm thay đổi đáng kể cách video giáo dục được tạo ra. Trước đây, sản xuất một video thường đòi hỏi sự tham gia của giảng viên, người trình bày, người viết kịch bản, thiết bị ghi hình và các công đoạn biên tập. Hiện nay, một phần đáng kể của quy trình này có thể được thực hiện bằng các hệ thống tạo sinh. Điều này không chỉ rút ngắn quá trình sản xuất mà còn cho phép tạo nhiều phiên bản nội dung theo ngôn ngữ, mục đích hoặc nhóm người học khác nhau. Trong bài báo này, *video giáo dục do AI tạo sinh (AI-generated educational video - AIGV) được hiểu là một sản phẩm video giáo dục trong đó các đầu ra do AI tạo sinh được tích hợp vào sản phẩm cuối cùng như một thành phần nghe nhìn cốt lõi*. Các yếu tố này có thể là chuỗi hình ảnh, hoạt ảnh, đoạn phim ngắn, người dẫn/nhân vật ảo (avatar), hoặc lời thuyết minh/lồng tiếng tổng hợp nhằm truyền tải nội dung giảng dạy thực chất. Tùy thuộc vào mức độ can thiệp của con người trong khâu thiết kế sơ phạm và sản xuất, AIGV có thể được tạo ra hoàn toàn hoặc một phần bởi AI. Theo đó, việc chỉ sử dụng AI cho các công việc hỗ trợ ngoài lề như lấy ý tưởng, viết kịch bản, phân cảnh (storyboarding), tạo phụ đề, cắt ghép, nén video, đề xuất nội dung, phân tích dữ liệu thì không đủ tiêu chuẩn để xếp loại là AIGV. Sự quan tâm đối với AIGV phần lớn bắt nguồn từ khả năng giảm thời gian và nguồn lực cần thiết cho việc sản xuất học liệu. Trong các môi trường có số lượng người học lớn, nội dung thường xuyên thay đổi hoặc yêu cầu nhiều phiên bản ngôn ngữ, AIGV có thể cung cấp một giải pháp linh hoạt hơn so với việc ghi hình lại toàn bộ bài giảng. Công nghệ AI tạo sinh cũng mở ra khả năng minh họa những bối cảnh khó quay, xây dựng nhân vật ảo và tạo các tình huống mô phỏng phục vụ học tập. Tuy nhiên, khả năng tự động hóa sản xuất không đồng nghĩa với hiệu quả học tập, bởi AIGV vẫn cần tuân thủ các nguyên tắc thiết kế đa phương tiện và kiểm soát tải nhận thức.

Sự gia tăng nhanh chóng của các nghiên cứu và tri thức liên quan đến AIGV đã thúc đẩy nhiều nghiên cứu tổng quan về chủ đề này (Dou và cộng sự, 2025; Impito, 2026; Shu và cộng sự, 2025; Tran và cộng sự, 2025; Xie và cộng sự, 2025). Sự phát triển nhanh của các công cụ tạo video đã thúc đẩy nhiều nghiên cứu tổng quan về AIGV trong giáo dục. Một số công trình sử dụng phân tích thư mục hoặc trích lược khoa học để xác định xu hướng công nghệ và chủ đề nghiên cứu (Dou và cộng sự, 2025; Xie và cộng sự, 2025), trong khi các nghiên cứu khác tập trung vào ứng dụng, lợi ích và thách thức của AIGV trong giáo dục đại học, đào tạo y khoa và các bối cảnh học tập cụ thể (Impito, 2026; Shu và cộng sự, 2025; Tran và cộng sự, 2025). Các nghiên cứu này cung cấp một bức tranh khái quát có giá trị, nhưng chưa làm rõ đầy đủ sự không thống nhất của bằng chứng thực nghiệm về hiệu quả học tập. Một số nghiên cứu cho thấy kết quả từ AIGV tương đương video do con người sản xuất, trong khi một số khác chỉ ghi nhận lợi thế ở những nhiệm vụ cụ thể; đồng thời, video có người thật thường được đánh giá cao hơn về sự gắn kết và hiện diện xã hội. Sự không thống nhất này đặt ra nhu cầu cho một nghiên cứu tổng quan nhằm tổng hợp các nghiên cứu thực nghiệm đo lường trực tiếp kết quả học tập. Trên cơ sở đó, đánh giá tường thuật này trả lời 03 câu hỏi: (1) Các nghiên cứu thực nghiệm về kết quả học tập từ AIGV có những đặc điểm chính nào?; (2) Các nghiên cứu hiện có báo cáo những kết quả học tập nào khi so sánh AIGV với video do con người sản xuất và các học liệu thông thường khác?; (3) Những lợi thế, giới hạn và điều kiện sử dụng AIGV nào được hỗ trợ bởi bằng chứng hiện có?

Thông qua việc trả lời 03 câu hỏi, bài báo đóng góp vào việc làm rõ liệu AIGV có thể mang lại hiệu quả học tập so với video do con người sản xuất hoặc các học liệu đối chứng khác không. Điều này có thể có ý nghĩa thực tiễn đối với các cơ sở giáo dục đang cân nhắc sử dụng AIGV để thay thế một phần hoạt động trình bày của giảng viên.

2. Phương pháp nghiên cứu

Chúng tôi thực hiện theo phương pháp đánh giá tường thuật nhằm tổng hợp và diễn giải các bằng chứng thực nghiệm về kết quả học tập từ AIGV. Nguồn tài liệu khởi đầu là bộ dữ liệu “*Dataset for a Scoping Review on AI-Generated Video in Education (May 2026)*”, gồm 65 công trình liên quan đến AIGV trong giáo dục được thu thập từ Scopus, Web of Science và Google Scholar, được công khai trên Figshare (Hanh, 2026). Bộ dữ liệu này được sử dụng làm nguồn sàng lọc ban đầu vì có phạm vi trực tiếp và tương đối bao quát đối với chủ đề nghiên cứu. Để cập nhật tài liệu và giảm khả năng bỏ sót các nghiên cứu phù hợp, chúng tôi thực hiện tìm kiếm bổ sung trên Google Scholar vào tháng 6/2026 với chuỗi tìm kiếm đã được sử dụng trong (Hanh và Mai, 2026): (“AI-generated video*” OR “AI generated video*” OR “generative AI video*” OR “synthetic video*” OR “text-to-video” OR “text to video” OR “image to video” OR “image-to-video” OR “images-to-video” OR “AI video generation” OR “AI avatar video*” OR “AI-avatar video*” OR “AI-generated Avatar” OR “deepfake video*” OR “Synthesia” OR “HeyGen” OR “Sora” OR “Runway” OR “Veo” OR “D-ID” OR “Pika”) AND (“education” OR “teaching” OR “online learning” OR “e-learning” OR “blended learning” OR “instruction” OR “pedagogy” OR “classroom” OR “school” OR “K-12” OR “MOOC” OR “instructional design”). Giới hạn thời gian tìm kiếm được xác định từ 2023 trở đi do đây là giai đoạn ChatGPT cùng các công cụ AI tạo sinh bắt đầu phổ biến rộng rãi (Mai và cộng sự, 2024; Mai và Hanh, 2024).

Nghiên cứu được lựa chọn khi: (1) Có thiết kế thực nghiệm; (2) Sử dụng AIGV làm học liệu hoặc phương tiện hướng dẫn; (3) Đo ít nhất một kết quả về kiến thức, ghi nhớ, chuyển giao, kỹ năng, hành vi hoặc thực hành; (4) Có so sánh giữa các điều kiện, đo trước-sau hoặc đo lặp. Các nghiên cứu chỉ khảo sát thái độ, chất lượng kỹ thuật, hoặc ý định sử dụng; các bài khái niệm, tổng quan, mô tả công cụ và nghiên cứu không đo việc học từ video bị loại trừ. Ngoài ra, với tính chất của một đánh giá tường thuật, việc đánh giá chất lượng nghiên cứu và nguy cơ sai lệch không phải là yêu cầu bắt buộc hoặc không đòi hỏi quy trình chuẩn hóa nghiêm ngặt như trong đánh giá có hệ thống (systematic review). Vì vậy, bài báo này không áp dụng các công cụ đánh giá nguy cơ sai lệch chính thức. Kết quả là 15 nghiên cứu phù hợp được lựa chọn. Việc sàng lọc và trích xuất dữ liệu ban đầu do tác giả thứ nhất thực hiện; bảng dữ liệu tổng hợp sau đó được các đồng tác giả rà soát và trao đổi để thống nhất những trường hợp chưa rõ. Với mỗi nghiên cứu, các thông tin về năm công bố, bối cảnh, đặc điểm mẫu, thiết kế nghiên cứu, loại AIGV, điều kiện so sánh và kết quả học tập được trích xuất vào một bảng tổng hợp. Dữ liệu trước hết được mô tả theo đặc điểm của các nghiên cứu, sau đó được phân nhóm và tổng hợp tường thuật theo 03 nội dung: kết quả học tập so với điều kiện so sánh, những lợi ích được ghi nhận và các hạn chế và khoảng trống nghiên cứu của AIGV. Khi kết quả giữa các nghiên cứu không thống nhất, nhóm tác giả đối chiếu sự khác biệt về thiết kế, đối tượng, loại video, điều kiện so sánh và thước đo kết quả.

3. Kết quả nghiên cứu

3.1. Đặc điểm của các nghiên cứu được tổng hợp

Tổng số 15 nghiên cứu được công bố trong giai đoạn 2023-2026, gồm 01 nghiên cứu năm 2023, 02 nghiên cứu năm 2024, 08 nghiên cứu năm 2025 và 04 nghiên cứu năm 2026. Các nghiên cứu chủ yếu được thực hiện trong bối cảnh giáo dục đại học, đào tạo người trưởng thành và học tập trực tuyến, trong khi các nghiên cứu ở giáo dục phổ thông, giáo dục sức khỏe cộng đồng và phục hồi chức năng còn tương đối ít. Các công nghệ AIGV được sử dụng phổ biến bao gồm giảng viên hoặc nhân vật ảo (AI-generated avatars/virtual instructors), giọng nói tổng hợp (synthetic voice), video tạo từ văn bản (text-to-video generation) và các quy trình sản xuất kết hợp nhiều công cụ AI (multi-tool AI production workflows).

Về thiết kế nghiên cứu, 10 nghiên cứu sử dụng điều kiện đối chứng trực tiếp hoặc gián tiếp, 01 nghiên cứu so sánh hai phiên bản AIGV, trong khi 04 nghiên cứu còn lại đánh giá hiệu quả của các can thiệp trước và sau hoặc các mô hình dạy học tích hợp AIGV. Các chỉ số được đánh giá tập trung chủ yếu vào kết quả nhận thức và khả năng ghi nhớ, còn các bằng chứng về chuyển giao học tập, duy trì kết quả dài hạn, kỹ năng thực hành và thay đổi hành vi vẫn còn hạn chế. Đặc điểm và kết quả chính của toàn bộ các nghiên cứu được trình bày trong bảng 1.

Bảng 1. Đặc điểm và kết quả chính của 15 nghiên cứu được đưa vào tổng quan

Nguồn	Mẫu	Loại AIGV/tool	Mô tả thiết kế	Kết quả học tập được báo cáo
(Leiker và cộng sự, 2023)	Người lớn, n=83	AI-generated instructor/ Synthesia	So sánh giảng viên ảo và video do con người làm	Cả hai nhóm đều tăng điểm ($p<0,001$); không ghi nhận khác biệt về mức tăng hoặc cảm nhận của người học; kết quả chỉ cho thấy khả năng thay thế trong ngắn hạn.
(Arkün-Kocadere và Çağlar-Özhan, 2024)	Sinh viên, n=108	AI-generated instructor/ Vidext	Hai nhóm; cùng kịch bản/hình ảnh, khác người trình bày	Thành tích tăng tương đương; nhóm người thật gắn kết cao hơn ($p<0,05$), cho thấy lợi thế cảm xúc-xã hội.
(Kim, 2024)	HS THPT EFL, n=60	Text-to-video, text-to-image, text-to-voice và image-to-video/ Vrew và Gen3	Video AI so với video người thật; đo kiểm tra trước - sau giai đoạn giữa và kiểm tra sau độ trễ	Video do con người tạo ra được ưa chuộng hơn; video do AI tạo ra được đánh giá cao về hình ảnh hấp dẫn và hiệu quả; người tham gia đề xuất cải thiện tính chân thực của hình đại diện.
(Netland và cộng sự, 2025)	Người học trực tuyến, n=447	AI-generated instructor/ HeyGen	Thí nghiệm đa chủ đề; video AI so với video người thật	Không khác biệt tổng thể ($p = 0,685$); trải nghiệm hơi nghiêng về người thật và kết quả thay đổi theo chủ đề.
(Xu và cộng sự, 2025)	Sinh viên học từ vựng, n=76	AI-generated instructor/ AI Teacher GPT	Ngẫu nhiên 38/38; AIGV so với video ghi hình	AIGV cao hơn về ghi nhớ, nhưng không khác về chuyển giao/điểm tổng ($p>0,05$); video ghi hình có hiện diện xã hội cao hơn.
(Agarwal và cộng sự, 2025)	Người khiêm thính, n=120	AI-generated instructor/ LipMOOC	Video khẩu hình tổng hợp so với video người thật	Video tổng hợp không làm giảm kết quả; khóa có cấu trúc tốt hơn khóa chuẩn, gợi ý vai trò lớn của thiết kế khóa học.

(Struger và cộng sự, 2025)	n=55	AI-generated instructor/ HeyGen và ElevenLabs	Hỗn hợp; avatar AI so với giảng viên thật	Tỉ lệ trả lời đúng ngay sau video cao hơn nhẹ ở điều kiện avatar AI; phản ứng cảm xúc nhìn chung tích cực, nhưng còn lo ngại về tính chân thực.
(Yaojun và Tekwa, 2025)	Sinh viên năm nhất	GenAI text-to-video	AIGV trong động não so với đối chứng	Tạo nhiều ý tưởng hơn trong thời gian ngắn hơn, nhưng tác giả thận trọng và chưa chứng minh chất lượng viết dài hạn.
(Lind, 2025)	Nhân viên chuyên nghiệp, n>250	AI-generated instructor/ HeyGen và ElevenLabs	Ngẫu nhiên 2x2; phát ngôn viên tổng hợp và người thật	Không khác về quiz; khi người xem nhận ra tính tổng hợp, đánh giá cảm nhận giảm nhẹ.
(Pellas, 2025)	Giáo sinh khoa học, n=55	GenAI text-to-video/ Sudowrite, Visla và Jasper	Thiết kế trong nhóm, kết hợp với kiểm tra trước/sau; hai dạng AIGV	Cả hai dạng AIGV (xem trước và không xem trước các khái niệm chính) cải thiện ($p<0,05$); thiếu đối chứng không-AIGV nên chưa tách được hiệu ứng thời gian/thực hành.
(Aimara và cộng sự, 2025)	Thực hành sinh học động vật, n = 47 sinh viên	AI-generated instructor/ MidJourney và Synthesia	Học kết hợp: AIGV, video phòng thí nghiệm, trò chơi, quiz	Gắn kết, hiểu quy trình và hài lòng tăng so với phương pháp truyền thống.
(Pagrut và cộng sự, 2026)	HS 10-15 tuổi, n=250	Animation hoặc Live-action video/ STEVE AI	Một nhóm trước-sau; video địa phương 3 phút/ngày, 7 ngày	Kiến thức, hành vi và chỉ số răng miệng cải thiện ($p < 0,001$); không có nhóm chứng nên quy kết nhân quả yếu.
(Castillo-Cuesta và cộng sự, 2026)	Người học EFL B2, n=156	Live-action video / Fliki, Kapwing, SteveAI, InVideo; AI-generated instructor/ Voki	Bán thực nghiệm; AIGV + lớp học đảo ngược	Hiệu/ghi nhớ và tự chủ tăng; không tách được AIGV khỏi toàn bộ mô hình.
(Liu, 2026)	Sinh viên năm nhất, n=55	Animation hoặc Live-action video	Học theo dự án (trong đào tạo kế toán) + AIGV; trước-sau	Kiến thức bền vững và kết quả nhiệm vụ kế toán tăng; can thiệp đa thành phần, không có đối chứng.
(Ülgen và cộng sự, 2026)	Người bệnh, 159 ngẫu nhiên, 103 hoàn tất	Animation hoặc Live-action video/ Veo 3	RCT; video tập luyện AI so với tờ hướng dẫn	Tuân thủ tương đương; video chính xác hơn ở bài tập phức tạp (so với in giấy truyền thống), nhưng theo dõi ngắn và tỉ lệ bỏ cuộc đáng lưu ý.

Ghi chú: Animation hoặc Live-action video được hiểu là video hoạt hình hoặc cảnh thực; AI-generated instructor được hiểu là người hướng dẫn ảo AI.

3.2. AIGV thường đạt kết quả tức thời tương đương, nhưng chưa thể hiện lợi thế nhất quán

Mẫu hình phổ biến nhất trong các nghiên cứu được đánh giá là AIGV hỗ trợ người học tiếp thu nội dung, nhưng không thể hiện lợi thế nhất quán so với video có người trình bày thật. Leiker và cộng sự (2023), Arkün-Kocadere và Çağlar Özhan (2024), Kim (2024), Netland và cộng sự (2025) và Lind (2025) nhìn chung báo cáo kết quả học tập tương tự hoặc không ghi nhận khác biệt có ý nghĩa thống kê giữa hai điều kiện của thực nghiệm. Khi 02 loại video sử dụng cùng kịch bản và tài liệu học tập, việc thay người thật bằng nhân vật do AI tạo ra nhìn chung không làm thay đổi rõ rệt điểm số. Tuy nhiên, người học thường cảm thấy video có người thật hấp dẫn, tự nhiên hoặc dễ tạo kết nối hơn. Dù vậy, kết quả “không có khác biệt có ý nghĩa thống kê” chưa đủ để khẳng định hai loại video hoàn toàn tương đương. Phần lớn các nghiên cứu chưa được thiết kế riêng để kiểm định tính tương đương. Vì thế, chỉ có thể kết luận rằng AIGV tạo ra kết quả học tập ngắn hạn tương tự video có người thật trong một số nhiệm vụ cụ thể; chưa có đủ bằng chứng để khẳng định AIGV có thể thay thế hoàn toàn video do con người trình bày.

3.3. Lợi ích của AIGV chủ yếu xuất hiện trong những nhiệm vụ học tập cụ thể

Lợi thế của AIGV chỉ xuất hiện ở một số nhiệm vụ cụ thể. Xu và cộng sự (2025) ghi nhận điểm ghi nhớ từ vựng cao hơn, nhưng không khác về chuyển giao hoặc điểm tổng và video ghi hình có hiện diện xã hội cao hơn. Ülgen và cộng sự (2026) không thấy khác biệt về tuân thủ tập luyện, song video giúp thực hiện chính xác hơn động tác nhiều bước. Yaojun và Tekwa (2025) cho thấy, AIGV hỗ trợ tạo nhiều ý tưởng hơn trong thời gian ngắn hơn, nhưng chưa

chứng minh chất lượng viết dài hạn. Agarwal và cộng sự (2025) cho thấy, video khẩu hình tổng hợp không làm giảm việc học, còn khóa có cấu trúc tốt hơn tạo kết quả cao hơn khóa học tiêu chuẩn. Struger và cộng sự (2025) ghi nhận tỉ lệ trả lời đúng ngay sau video cao hơn nhẹ ở điều kiện avatar AI và phản ứng cảm xúc nhìn chung tích cực, nhưng người tham gia vẫn bày tỏ lo ngại về tính chân thực. Pellas (2025) đã ghi nhận cải thiện ở hai dạng AIGV nhưng thiếu đối chứng không - AIGV. Nhìn chung, lợi ích rõ hơn khi công nghệ đáp ứng trực tiếp yêu cầu trực quan hóa, mô phỏng chuyển động hoặc tạo nội dung chuyên biệt. Như vậy, bằng chứng hiện có không cho thấy một tác động chung của AIGV đối với mọi kết quả học tập; những lợi ích được ghi nhận chủ yếu phụ thuộc vào loại nhiệm vụ và chức năng cụ thể của video.

3.4. Các mô hình tích hợp AIGV cho kết quả tích cực nhưng khó quy kết tác động riêng

04 nghiên cứu sử dụng thiết kế trước-sau hoặc triển khai AIGV như một phần của gói can thiệp đều ghi nhận những thay đổi tích cực. Aimara và cộng sự (2025) kết hợp AIGV với video hướng dẫn thí nghiệm, trò chơi nghiêm túc và câu hỏi trắc nghiệm do AI hỗ trợ trong môi trường học kết hợp. Castillo-Cuesta và cộng sự (2026) sử dụng AIGV cùng lớp học đảo ngược và các hoạt động học trực tuyến đồng bộ. Liu (2026) tích hợp AIGV vào học theo dự án và nhiệm vụ tìm kiếm, xử lý thông tin. Pagrut và cộng sự (2026) cho người tham gia xem các video ngắn bằng ngôn ngữ địa phương lặp lại trong bảy ngày. Nhìn chung, các nghiên cứu này ghi nhận sự cải thiện về hiểu biết, mức độ tham gia, kết quả thực hiện nhiệm vụ, hành vi hoặc một số chỉ số lâm sàng. Tuy nhiên, không thể khẳng định những cải thiện này chỉ do AIGV tạo ra. Trong mỗi nghiên cứu, AIGV được sử dụng cùng nhiều yếu tố khác, như lớp học đảo ngược, làm việc nhóm, phản hồi, trò chơi, câu hỏi trắc nghiệm, ngôn ngữ địa phương hoặc việc lặp lại nội dung hằng ngày. Chính những yếu tố này cũng có thể góp phần tạo ra kết quả tích cực. Vì vậy, các nghiên cứu trên chủ yếu cho thấy mô hình dạy học có tích hợp AIGV có tính khả thi và có thể mang lại lợi ích, nhưng chưa cung cấp bằng chứng riêng biệt rằng AIGV hiệu quả hơn các hình thức học liệu khác.

3.5. Mức độ chặt chẽ của thiết kế và khoảng trống phương pháp

06 nghiên cứu có thiết kế đối chứng tương đối chặt chẽ và sử dụng các kết quả đo khách quan (Arkün-Kocadere và Çağlar-Özhan, 2024; Leiker và cộng sự, 2023; Lind, 2025; Netland và cộng sự, 2025; Ülgen và cộng sự, 2026; Xu và cộng sự, 2025). Một số nghiên cứu khác cũng có nhóm đối chứng, nhưng cỡ mẫu còn nhỏ hoặc chưa mô tả đầy đủ quy trình thực nghiệm (Agarwal và cộng sự, 2025; Kim, 2024; Struger và cộng sự, 2025; Yaojun và Tekwa, 2025). Các nghiên cứu chỉ so sánh trước và sau hoặc sử dụng AIGV trong một gói can thiệp gồm nhiều hoạt động có khả năng xác định quan hệ nhân quả thấp hơn. Bên cạnh đó, các nghiên cứu đo lường những kết quả rất khác nhau, như điểm quiz, chất lượng bài viết, ghi nhớ từ vựng, khả năng tạo ý tưởng, đọc khẩu hình, kiến thức sức khỏe và độ chính xác vận động. Sự khác biệt này khiến việc gộp các kết quả thành một ước lượng chung không phù hợp. Phần lớn nghiên cứu chỉ đánh giá kết quả ngay sau khi người học xem video; rất ít nghiên cứu kiểm tra khả năng ghi nhớ sau một khoảng thời gian hoặc khả năng vận dụng kiến thức vào nhiệm vụ mới. Nhiều nghiên cứu cũng chỉ được thực hiện tại một môn học hoặc với một/một số video duy nhất, nên khó khái quát kết quả sang các bối cảnh khác. Ngoài ra, thông tin về công cụ AI, phiên bản mô hình, kịch bản, quy trình biên tập, mức độ tham gia của con người, việc kiểm tra độ chính xác của nội dung và cách thông báo cho người học rằng video được tạo bởi AI chưa được báo cáo thống nhất. Do chất lượng của AIGV phụ thuộc không chỉ vào mô hình AI mà còn vào các quyết định thiết kế và kiểm duyệt của con người. Việc thiếu thông tin về quá trình sản xuất làm giảm khả năng lặp lại nghiên cứu và gây khó khăn khi giải thích vì sao các nghiên cứu đưa ra kết quả khác nhau.

4. Kết luận và bình luận

4.1. AIGV có cải thiện việc học không?

Câu trả lời được hỗ trợ tốt nhất là: AIGV có thể hỗ trợ việc học và trong nhiều trường hợp đạt kết quả ngắn hạn tương tự học liệu đối chứng, nhưng chưa có bằng chứng nhất quán rằng nó cải thiện việc học tốt hơn video do con người sản xuất. Kết luận này bác bỏ giả định rằng người trình bày tổng hợp tất yếu làm giảm việc học, nhưng cũng không ủng hộ tuyên bố rằng tự động hóa video tự thân tạo ra thành tích cao hơn. Kết quả tương đương phù hợp với quan điểm của học tập đa phương tiện: điều quyết định không chỉ là nguồn gốc con người hay AI, mà là mức độ nội dung được tổ chức, tin hiệu hóa, phân đoạn và phối hợp hình và lời phù hợp với xử lý nhận thức (Mayer, 2014). Khi hai điều kiện dùng cùng kịch bản và hình ảnh, như trong Arkün-Kocadere và Çağlar Özhan (2024), khác biệt về người trình bày không đủ để tạo khác biệt thành tích. Ngược lại, khi AIGV thực hiện tốt những nhiệm vụ mà học liệu đối chứng khó đáp ứng, như tái hiện khẩu hình nhất quán, trực quan hóa các động tác nhiều bước hoặc tạo nhanh các tình huống minh họa, lợi ích cục bộ có thể xuất hiện (Agarwal và cộng sự, 2025). Tuy nhiên, hiệu quả học tập không chỉ được phản ánh qua kết quả của các bài kiểm tra kiến thức. Các nghiên cứu của Arkün-Kocadere và Çağlar Özhan (2024), Kim (2024),

Netland và cộng sự (2025) và Xu và cộng sự (2025) cho thấy người thật thường được đánh giá tốt hơn về gắn kết, nhịp độ, tính chân thực hoặc hiện diện xã hội. Điều này không nhất thiết làm giảm điểm ngắn hạn, nhưng có thể ảnh hưởng đến động lực, niềm tin, kiên trì và chất lượng tương tác trong khóa học dài. Vì vậy, thay thế người thật bằng người dẫn/nhân vật ảo chỉ dựa trên sự tương đương của một bài kiểm tra ngay sau xem video là chưa đủ.

4.2. Hàm ý thực tiễn

Bài báo đề xuất 03 hàm ý thực tiễn chung cho giáo dục, cụ thể: *Thứ nhất*, AIGV nên được xem là một phương án sản xuất học liệu, không phải một phương pháp sư phạm độc lập. Giá trị của nó phụ thuộc vào mục tiêu, kịch bản, cấu trúc nội dung và hoạt động đi kèm. Các can thiệp lớp học đảo ngược cho thấy thiết kế khóa học có thể quan trọng hơn việc khuôn mặt trên video là tổng hợp hay thật (Agarwal và cộng sự, 2025; Castillo-Cuesta và cộng sự, 2026); *Thứ hai*, nên ưu tiên AIGV trong các nhiệm vụ có lợi thế chức năng rõ: tạo phiên bản đa ngôn ngữ, minh họa thao tác, mô phỏng tình huống khó quay, tạo nhiều biến thể luyện tập hoặc hỗ trợ tiếp cận. Với nội dung đòi hỏi quan hệ, cảm xúc, thảo luận đạo đức nhạy cảm hoặc xây dựng niềm tin, giảng viên thật hoặc mô hình kết hợp người và AI có thể phù hợp hơn. Avatar có thể đảm nhận phân hướng dẫn lặp lại, còn GV giữ vai trò giải thích, phản hồi và tương tác; *Thứ ba*, việc sử dụng AIGV còn đặt ra những yêu cầu về đạo đức và kiểm soát chất lượng. Trong y tế, khoa học, kỹ thuật và các lĩnh vực rủi ro cao, mọi AIGV phải được chuyên gia kiểm định trước sử dụng. Hình ảnh trôi chảy và giọng nói tự nhiên không bảo đảm tính đúng đắn. Cần lưu giữ kịch bản, nguồn nội dung, lời nhắc, công cụ, thời điểm tạo, bước chỉnh sửa và người phê duyệt. Việc công bố rõ video có yếu tố AI cũng cần được xem như một phần của đạo đức và thiết kế nghiên cứu, bởi nhận biết tính tổng hợp có thể làm thay đổi cách người học đánh giá nội dung (Lind, 2025).

Trong bối cảnh chuyển đổi số giáo dục tại Việt Nam, AIGV có thể được ưu tiên cho việc xây dựng học liệu trực tuyến, tạo phiên bản nội dung bằng tiếng Việt và trực quan hóa những hiện tượng hoặc quy trình khó tổ chức trực tiếp, như thí nghiệm khoa học, thao tác kỹ thuật và tình huống nghề nghiệp. Tuy nhiên, các cơ sở giáo dục nên triển khai thử nghiệm ở quy mô nhỏ trước khi sử dụng rộng rãi, đồng thời yêu cầu giảng viên kiểm tra độ chính xác của nội dung, chất lượng phát âm tiếng Việt và sự phù hợp với văn hóa học tập của người học. Việc lựa chọn công cụ cũng cần cân nhắc chi phí, bảo vệ dữ liệu và năng lực số của GV.

Nghiên cứu tiếp theo nên chuyển từ câu hỏi “AI hay con người tốt hơn” sang xác định loại AIGV nào phù hợp với từng mục tiêu, nhóm người học và bối cảnh sử dụng. Các nghiên cứu so sánh cần kiểm soát kịch bản, thời lượng, hình ảnh, giọng nói, nhịp độ và hoạt động học tập đi kèm; đồng thời đo cả kết quả tức thời, ghi nhớ dài hạn, chuyên giao, tải nhận thức, hiện diện xã hội và niềm tin của người học. Cũng cần có thêm nghiên cứu tại Việt Nam để xem xét mức độ chấp nhận GV ảo, chất lượng nội dung tiếng Việt và khả năng triển khai AIGV trong các điều kiện cơ sở vật chất khác nhau. Quy trình tạo video cũng cần được báo cáo rõ, bao gồm công cụ, nguồn kịch bản, lời nhắc, bước chỉnh sửa, mức độ tham gia của con người và quy trình kiểm định chuyên gia. Điều này giúp tăng khả năng lặp lại và giải thích sự khác biệt giữa các kết quả nghiên cứu.

Đánh giá này có một số hạn chế, cụ thể: *Thứ nhất*, tài liệu được xác định từ một bộ dữ liệu công khai và tìm kiếm bổ sung trên Google Scholar, thay vì thực hiện tìm kiếm mới trên nhiều cơ sở dữ liệu; do đó, một số nghiên cứu phù hợp có thể bị bỏ sót; *Thứ hai*, chỉ các công trình tiếng Anh đã qua phản biện được xem xét, làm hạn chế khả năng phản ánh nghiên cứu được công bố bằng các ngôn ngữ khác; *Thứ ba*, do đây là đánh giá tường thuật, nghiên cứu không thực hiện đánh giá nguy cơ sai lệch bằng công cụ chuẩn hóa hoặc tính lại kích thước hiệu ứng. Vì vậy, kết luận chủ yếu phản ánh các mô hình kết quả được báo cáo, không phải một ước lượng định lượng về hiệu quả chung của AIGV; *Thứ tư*, sự khác biệt lớn về đối tượng, loại video, thiết kế và thước đo khiến việc so sánh giữa các nghiên cứu cần được diễn giải thận trọng.

Tuyên bố về vai trò của các tác giả: Nguyễn Văn Hạnh: Xây dựng ý tưởng nghiên cứu, Thiết kế phương pháp nghiên cứu, Thực hiện nghiên cứu và thu thập dữ liệu, Phân tích dữ liệu, Viết bản thảo ban đầu, Chỉnh sửa và phản biện bản thảo, Giám sát, chỉ đạo quá trình nghiên cứu. Trần Triệu Hải và Dương Thị Thuỳ Mai: Thiết kế phương pháp nghiên cứu, Thực hiện nghiên cứu và thu thập dữ liệu, Phân tích dữ liệu, Viết bản thảo ban đầu.

Tuyên bố về GenAI và Quyền tác giả: Trong quá trình chuẩn bị bản thảo này, các tác giả đã sử dụng ChatGPT (GPT-5.5 Thinking) của OpenAI để hỗ trợ kiểm tra chính tả, rà soát diễn đạt và cải thiện văn phong.

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 được tài trợ bởi Quỹ Phát triển Khoa học và Công nghệ Quốc gia (NAFOSTED) theo mã số 503.01-2025.21.

Lời cảm ơn: Nhóm tác giả trân trọng cảm ơn Quỹ Phát triển Khoa học và Công nghệ Quốc gia (NAFOSTED) đã tài trợ cho nghiên cứu này thông qua đề tài “Bài giảng video do AI tạo ra có thay thế được video do con người làm trong môi trường blended learning/ e-learning không? Nghiên cứu về sự hài lòng, động lực và gắn kết học tập của sinh viên đại học”, mã số 503.01-2025.21.

Tài liệu tham khảo

- Agarwal, S., Namboodiri, V., & Jawahar, C. V. (2025). LipMOOC: A Modular Synthetic MOOC for Lipreading. *International Conference on Artificial Intelligence in Education*, 179-186. https://doi.org/10.1007/978-3-031-98462-4_23
- Aimara, S., Radid, M., Chems, G., Nfissi, I., Tabite, R., & Khiati, M. (2025). Hybridization of Animal Biology Practical Work: Innovative Pedagogical Design Through Generative Artificial Intelligence. *International Symposium on Generative AI and Education*, 314-327. https://doi.org/10.1007/978-3-031-98476-1_25
- Arkün-Kocadere, S., & Çağlar-Özhan, Ş. (2024). Video lectures with AI-generated instructors: Low video engagement, same performance as human instructors. *International Review of Research in Open and Distributed Learning*, 25(3), 350-369. <https://doi.org/10.19173/irrodl.v25i3.7815>
- Berk, R. A. (2009). Multimedia teaching with video clips: TV, movies, YouTube, and mtvU in the college classroom. *International Journal of Technology in Teaching & Learning*, 5(1), 1-21.
- Castillo-Cuesta, L., Cabrera-Solano, P., & Ochoa-Cueva, C. (2026). Integrating AI-generated videos and the flipped classroom approach to enhance active English language learning. *Studies in English Language and Education*, 13(1), 45-60. <https://doi.org/10.24815/siele.v13i1.129>
- Dou, A., Xu, W., & Liu, R. (2025). Starting From Sora: A Scientometric Analysis of Generative Videos for Educational Purposes. *International Journal of Distance Education Technologies (IJDET)*, 23(1), 1-19. <https://doi.org/10.4018/IJDET.373234>
- Guo, P. J., Kim, J., & Rubin, R. (2014, March). How video production affects student engagement: An empirical study of MOOC videos. In *Proceedings of the first ACM conference on Learning scale conference* (pp. 41-50).
- Hanh, N. V., & Mai, T. T. D. (2026). AI-Generated Video in Education: A Scoping Review of Technologies, Pedagogical Applications, Benefits, and Challenges. *IEEE Access*.
- Hanh, N. V. (2026). *Dataset for a scoping review on AI-generated video in education (May 2026)*. <https://doi.org/10.6084/m9.figshare.32386242>
- Hung, T. T., Tuyet, N. T., Long, N. T., & Van Hanh, N. V. (2020). The effect of positive learning culture in students' blended learning process. *Journal of E-Learning and Knowledge Society*, 16(3), 68-75. <https://doi.org/10.20368/1971-8829/1135218>
- Impito, P. F. (2026). Artificial Intelligence Generated Videos as Supportive Tools in Medical Education: A Scoping Review. *Journal of Medical Education and Curricular Development*, 13. <https://doi.org/10.1177/23821205261437352>
- Kay, R. H. (2012). Exploring the use of video podcasts in education: A comprehensive review of the literature. *Computers in Human Behavior*, 28(3), 820-831. <https://doi.org/10.1016/j.chb.2012.01.011>
- Kim, R. (2024). Comparative Analysis of AI-Generated and Human-Made Instructional Videos: Effects on Learning English Caused-Motion Constructions by Korean High School EFL Learners. *Multimedia-Assisted Language Learning*, 27(4), 179-199. <https://doi.org/10.15702/mall.2024.27.4.179>
- Leiker, D., Gyllen, A. R., Eldesouky, I., & Cukurova, M. (2023). Generative AI for Learning: Investigating the Potential of Learning Videos with Synthetic Virtual Instructors. In N. Wang, G. Rebolledo-Mendez, V. Dimitrova, N. Matsuda, & O. C. Santos (Eds.), *Artificial Intelligence in Education. Posters and Late Breaking Results, Workshops and Tutorials, Industry and Innovation Tracks, Practitioners, Doctoral Consortium and Blue Sky* (pp 523-529). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-36336-8_81
- Lind, S. J. (2025). Can AI-powered avatars replace human trainers? An empirical test of synthetic humanlike spokesperson applications. *Journal of Workplace Learning*, 37(1), 19-40. <https://doi.org/10.1108/JWL-04-2024-0075>
- Liu, X. (2026). Integrating eco-friendly consumption knowledge into an accounting curriculum using a project-based collaborative learning approach. *International Journal of Sustainability in Higher Education*, 1-28. <https://doi.org/10.1108/IJSHE-02-2025-0101>
- Long, N. T., & Hanh, N. V. (2020). A Structural Equation Model of Blended Learning Culture in the Classroom. *International Journal of Higher Education*, 9(4), 99-115. <https://doi.org/10.5430/ijhe.v9n4p99>

- Mai, T. T. D., & Hanh, N. V. (2024). Whether English Proficiency and English Self-Efficacy Influence the Credibility of ChatGPT-Generated English Content of EMI Students. *International Journal of Computer-Assisted Language Learning and Teaching (IJCALLT)*, 14(1), 1-21. <https://doi.org/10.4018/IJCALLT.349972>
- Mai, T. T. D., Da, C. V., & Hanh, N. V. (2024). The use of ChatGPT in teaching and learning: a systematic review through SWOT analysis approach. *Frontiers in Education*, 9, 1328769. <https://doi.org/10.3389/feduc.2024.1328769>
- Mayer, R. E. (2014). Introduction to Multimedia Learning. In R. E. Mayer (Ed.), *The Cambridge Handbook of Multimedia Learning* (2nd ed., pp. 1-24). Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369.002>
- Means, B., Toyama, Y., Murphy, R., & Baki, M. (2013). The effectiveness of online and blended learning: A meta-analysis of the empirical literature. *Teachers College Record*, 115(3), 1-47.
- Netland, T., von Dzengelevski, O., Tesch, K., & Kwasnitschka, D. (2025). Comparing human-made and AI-generated teaching videos: An experimental study on learning effects. *Computers & Education*, 224, 105164. <https://doi.org/10.1016/j.compedu.2024.105164>
- Nguyễn Văn Hanh, Nguyễn Hữu Hợp (2016). *Giáo trình Công nghệ giáo dục*. NXB Đại học Quốc gia Hà Nội.
- Pagrut, S., Reche, A. M., Madhu, P. P., Vishnani, R., Deotare, P., Toshniwal, K., & Wadekar, A. (2026). Evaluating the effect of artificial intelligence-based oral health education model among 10 to 15-year-old children in Adivasi Ashram School of Central India-An interventional study. *Journal of Education and Health Promotion*, 15(1), 73. https://doi.org/10.4103/jehp.jehp_561_25
- Pellas, N. (2025). The impact of AI-generated instructional videos on problem-based learning in science teacher education. *Education Sciences*, 15(1), 102. <https://doi.org/10.3390/educsci15010102>
- Shu, J., Kou, H., Zhang, J., Zhang, T., Zhang, H., Xu, T., & Zhou, Y. (2025). AI-Generated Instructional Videos: A Systematic Review of Learning Impacts, Applications, and Research Perspectives. *2025 7th International Conference on Computer Science and Technologies in Education (CSTE)*, 877-881. <https://doi.org/10.1109/CSTE64638.2025.11091997>
- Struger, P., Brünner, B., & Ebner, M. (2025). Synthetic Educators: Analyzing AI-Driven Avatars in Digital Learning Environments. *International Conference on Human-Computer Interaction*, 187-199. https://doi.org/10.1007/978-3-031-93567-1_13
- Tran, T. H., Duong, T. T. M., & Nguyen, V. H. (2025). A rapid review of using AI-generated instructional videos in higher education. *Frontiers in Computer Science*, 7, 1721093. <https://doi.org/10.3389/fcomp.2025.1721093>
- Ülgen, N. K., Nazlıgöl, A. S., Yiğit, N., Arslan, A. K., & Akkurt, M. O. (2026). AI-Generated Exercise Videos vs Traditional Leaflets: A Novel Approach to Home Exercise Education in Achilles Tendinitis and Plantar Fasciitis: A Randomized Controlled Trial. *Foot & Ankle Orthopaedics*, 11(2), 24730114261436388. <https://doi.org/10.1177/24730114261436388>
- Xie, W., Hu, A., Xie, Q., Chen, J., Wan, R., & Liu, Y. (2025). Bibliometric analysis and review of AI-based video generation: research dynamics and application trends (2020-2025). *Discover Computing*, 28(1), 130. <https://doi.org/10.1007/s10791-025-09628-9>
- Xu, T., Liu, Y., Jin, Y., Qu, Y., Bai, J., Zhang, W., & Zhou, Y. (2025). From recorded to AI-generated instructional videos: A comparison of learning performance and experience. *British Journal of Educational Technology*, 56(4), 1463-1487. <https://doi.org/10.1111/bjet.13530>
- Yaojun, R., & Tekwa, K. (2025). From text to moving image: Evaluating generative artificial intelligence text-to-video models for pre-writing idea generation in language instruction. *Education and Information Technologies*, 30(13), 18749-18778. <https://doi.org/10.1007/s10639-025-13516-6>