

KIỂM TRA THÔNG TIN ĐỊA LÍ DO AI TẠO RA BẰNG NHIỀU NGUỒN: ĐÁNH GIÁ ĐỘ TIN CẬY VÀ LẬP LUẬN ĐỊA LÍ CỦA HỌC SINH LỚP 11

NGUYỄN VIỆT HẢI HIỆP

Trường Đại học Sư phạm, Đại học Đà Nẵng

Nhận bài ngày 20/8/2026. Sửa chữa xong 20/9/2026. Duyệt đăng 22/9/2026.

Abstract

This study presents a preliminary evaluation of the effectiveness of a procedure for checking AI-generated geographic information against multiple sources in supporting grade 11 students' ability to evaluate credibility and construct geographic arguments. A quasi-experimental pretest–posttest study with a control group was conducted over four lessons with 82 students at one upper secondary school, with 41 students in each group. After adjustment for pre-intervention scores using analysis of covariance (ANCOVA), the experimental class scored 0.452 points higher than the control class in geographic argumentation on a 0–3 scale (95% CI [0.292, 0.612], $p < 0.001$; Hedges' $g = 1.35$) and 0.455 points higher in credibility evaluation (95% CI [0.303, 0.607], $p < 0.001$; Hedges' $g = 1.13$). These preliminary findings suggest that the procedure may help students verify information and use geographic evidence. Further research across more classes and schools is needed to corroborate these findings.

Keywords: Credibility of geographic information, generative AI, geographic argumentation, geography teaching, multi-source information checking.

1. Đặt vấn đề

Trí tuệ nhân tạo (AI) tạo sinh đang làm thay đổi cách học sinh (HS) tìm kiếm và tiếp nhận thông tin nhưng đầu ra AI vẫn có thể chứa dữ kiện sai, thông tin lỗi thời, thiếu bối cảnh hoặc kết luận vượt quá bằng chứng. Các tổng quan và hướng dẫn gần đây đều nhấn mạnh yêu cầu người học phải xem xét độ phù hợp, độ tin cậy và cách sử dụng thông tin do AI tạo ra thay vì mặc nhiên coi đó là nguồn kiến thức đáng tin cậy (Fulsher et al., 2026; Miao & Holmes, 2023; Park, 2025).

Trong môn Địa lí, việc kiểm tra thông tin còn đòi hỏi HS xem xét thời gian, không gian, quy mô, đơn vị lãnh thổ và cách xác định chỉ tiêu. Yêu cầu phù hợp với đặc trưng của tư duy Địa lí (Bendl et al., 2025), Khung năng lực số cho người học (Bộ Giáo dục và Đào tạo, 2025a) và Chương trình giáo dục phổ thông môn Địa lí hiện hành (Bộ Giáo dục và Đào tạo, 2025b). Các nghiên cứu hiện có vẫn chưa làm rõ cách tổ chức cho HS kiểm tra thông tin Địa lí do AI tạo ra bằng nhiều nguồn trước khi sử dụng thông tin để xây dựng lập luận.

Nghiên cứu tổ chức cho HS kiểm tra thông tin Địa lí do AI tạo ra và hướng dẫn lớp thực nghiệm xem xét nguồn, đối chiếu nhiều nguồn, phân tích bằng chứng Địa lí và điều chỉnh phán đoán trước khi lập luận. Ba câu hỏi được đặt ra: sau khi điều chỉnh điểm trước can thiệp, hai nhóm khác nhau như thế nào về: 1) Lập luận Địa lí có căn cứ; 2) Khả năng đánh giá độ tin cậy của thông tin Địa lí; 3) Số nguồn độc lập được sử dụng, mức tích hợp bằng chứng và mức hoàn thành kiểm chứng?

2. Phương pháp nghiên cứu

Nghiên cứu sử dụng thiết kế bán thực nghiệm pretest-posttest có nhóm đối chứng trên 82 HS lớp 11 tại một trường trung học phổ thông, gồm 41 HS ở lớp thực nghiệm và 41 HS ở lớp đối chứng. Hai lớp học cùng nội dung và thời lượng trong bốn tiết. Lớp thực nghiệm được hướng dẫn kiểm tra thông tin Địa lí do AI tạo ra theo quy trình gồm: nêu phán đoán ban đầu, tách nội dung cần kiểm tra, xem

Email: nvhhiiep@ued.udn.vn

xét nguồn, đối chiếu nhiều nguồn, phân tích bằng chứng Địa lí, điều chỉnh nhận định và xây dựng lập luận cuối cùng. Lớp đối chứng học cùng nội dung và sử dụng các dạng dữ liệu tương tự nhưng không thực hiện quy trình trên. Do mỗi điều kiện chỉ gồm một lớp, kết quả được diễn giải trong phạm vi mẫu nghiên cứu và không dùng để khẳng định tác động nhân quả độc lập với đặc điểm lớp.

Kết quả học tập được đánh giá bằng nhiệm vụ thực hiện và bảng tiêu chí chấm điểm, sử dụng hai dạng đề tương ứng ở pretest và posttest. Hai biến chính đều được chấm trên thang 0-3: Y1 phản ánh khả năng đánh giá độ tin cậy của thông tin Địa lí, còn Y2 phản ánh khả năng lập luận Địa lí có căn cứ. Mỗi biến gồm bốn thành phần và điểm chung được tính bằng trung bình cộng. Y2 được xác định là kết quả chính, Y1 là kết quả phụ.

Ngoài điểm pretest và posttest, nghiên cứu còn ghi nhận số nguồn HS sử dụng, số nguồn độc lập, mức tích hợp bằng chứng và mức hoàn thành kiểm chứng. Các chỉ báo quá trình chỉ được dùng để mô tả cách HS thực hiện nhiệm vụ, chưa được xem là thang đo độc lập đã được thẩm định.

Phân tích chính sử dụng ANCOVA, trong đó điểm sau can thiệp là biến phụ thuộc, còn nhóm nghiên cứu và điểm trước can thiệp tương ứng là các biến dự báo. Cách phân tích giúp điều chỉnh chênh lệch trước can thiệp giữa hai lớp trong thiết kế đo trước-sau (Vickers & Altman, 2001). Do phương sai phần dư không đồng đều, sai số chuẩn HC3 được sử dụng để tính khoảng tin cậy và kiểm định thống kê (Long & Ervin, 2000). Phân tích bổ sung đưa thêm thành tích Địa lí trước can thiệp vào mô hình để kiểm tra độ ổn định của kết quả. Kích thước hiệu ứng được tính bằng Hedges' g cho thiết kế đo trước-sau có nhóm đối chứng theo Morris (2008). Kết quả ở 8 thành phần được xem xét theo hướng thăm dò còn các chỉ báo về quá trình thực hiện nhiệm vụ chỉ được trình bày ở mức mô tả.

3. Nội dung nghiên cứu

3.1. Cơ sở lý thuyết

Học sinh vẫn gặp khó khăn khi kiểm tra nguồn và bằng chứng trên môi trường trực tuyến (Breakstone et al., 2021). Các nghiên cứu cho thấy khả năng đánh giá độ tin cậy có thể được cải thiện khi HS được hướng dẫn xem xét thông tin về nguồn, đối chiếu nhiều tài liệu và kiểm tra chéo giữa các nguồn (Barzilai et al., 2020; Fendt et al., 2025; Hämäläinen et al., 2023; Marten & Stadtler, 2025; McGrew, 2024; Pimentel, 2025). Quá trình đánh giá cần xem xét chủ thể cung cấp thông tin, mục đích, thời điểm công bố và mức độ phù hợp với nội dung cần kiểm tra. Nhiệm vụ thực hiện có thể phản ánh trực tiếp khả năng đánh giá nguồn của HS thay vì chỉ dựa vào tự báo cáo (Sparks et al., 2021). Đối với thông tin Địa lí, HS còn phải xem xét sự phù hợp về thời gian, không gian, chỉ tiêu, đơn vị, phạm vi lãnh thổ và quy mô.

Đánh giá được nguồn chưa đủ để hình thành một lập luận Địa lí. HS còn phải lựa chọn bằng chứng phù hợp, kết hợp thông tin từ nhiều nguồn và giải thích vì sao bằng chứng hỗ trợ cho kết luận. Khi các nguồn không cùng hướng, người học cần nhận ra giới hạn của dữ liệu và điều chỉnh kết luận cho phù hợp. Trong môn Địa lí, yêu cầu lập luận còn gắn với bối cảnh không gian, sự khác biệt giữa các lãnh thổ, quy mô phân tích, biến đổi theo thời gian và quan hệ giữa các hiện tượng. Các nghiên cứu cho thấy HS thường gặp khó khăn khi phải đồng thời khai thác nhiều tài liệu, lựa chọn bằng chứng và kết nối dữ liệu với phần giải thích của mình (Gebele et al., 2022; Hindmarsh & Budke, 2023; Schwerdtfeger & Budke, 2021). Từ hai yêu cầu trên, hoạt động kiểm tra thông tin do AI tạo ra được tổ chức theo hướng HS không mặc định AI đúng hay sai. Các phát biểu do AI tạo được xem là nội dung cần kiểm tra. HS đưa ra phán đoán ban đầu, xác định phần thông tin cần kiểm tra, xem xét nguồn, đối chiếu nhiều nguồn, phân tích bằng chứng Địa lí, điều chỉnh nhận định và xây dựng lập luận cuối cùng. Trọng tâm không nằm ở việc tìm một nguồn để xác nhận câu trả lời của AI mà ở việc sử dụng nhiều nguồn độc lập và giải thích mức độ phù hợp của bằng chứng trước khi kết luận. Cách tổ chức phù hợp với các nghiên cứu về đánh giá và đối chiếu nguồn (Fendt et al., 2025; Hämäläinen et al., 2023; Marten & Stadtler, 2025; Pimentel, 2025), đồng thời đáp ứng yêu cầu sử dụng tài liệu để xây dựng lập luận trong dạy học Địa lí (Engelen & Budke, 2022; Gebele et al., 2022; Hindmarsh & Budke, 2023; Schwerdtfeger & Budke, 2021).

3.2. Kết quả nghiên cứu

3.2.1. Đặc điểm mẫu và kết quả trước can thiệp

Phân tích được thực hiện trên 82 HS, gồm 41 HS ở lớp đối chứng và 41 HS ở lớp thực nghiệm; tất cả đều có đầy đủ điểm trước và sau can thiệp. Trước can thiệp, thành tích Địa lí của hai lớp lần lượt là $7,12 \pm 0,63$ và $7,30 \pm 0,83$ (Hedges' $g = 0,23$). Điểm Y1 gần tương đương, với $1,80 \pm 0,42$ ở lớp đối chứng và $1,78 \pm 0,39$ ở lớp thực nghiệm ($g = -0,05$). Đối với Y2, lớp thực nghiệm có điểm thấp hơn lớp đối chứng, lần lượt là $1,88 \pm 0,36$ và $2,03 \pm 0,37$ ($g = -0,42$). Vì Y2 đã có chênh lệch giữa hai lớp trước can thiệp, ANCOVA được sử dụng để điều chỉnh điểm trước can thiệp khi so sánh kết quả sau can thiệp.

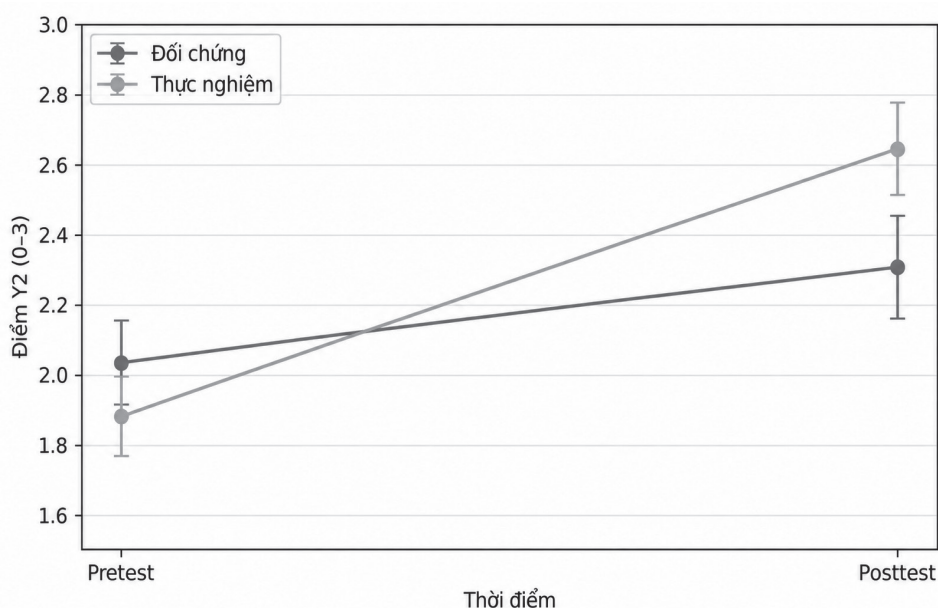
Sau can thiệp, điểm Y1 tăng từ $1,80 \pm 0,42$ lên $1,95 \pm 0,54$ ở lớp đối chứng và từ $1,78 \pm 0,39$ lên $2,39 \pm 0,44$ ở lớp thực nghiệm. Điểm Y2 tăng từ $2,03 \pm 0,37$ lên $2,30 \pm 0,47$ ở lớp đối chứng và từ $1,88 \pm 0,36$ lên $2,65 \pm 0,42$ ở lớp thực nghiệm. Mức tăng trung bình của lớp thực nghiệm cao hơn lớp đối chứng ở cả Y1 (0,61 so với 0,15) và Y2 (0,77 so với 0,27).

3.2.2. Kết quả về lập luận Địa lí có căn cứ (Y2)

Sau khi điều chỉnh theo điểm Y2 trước can thiệp, lớp thực nghiệm có điểm Y2 sau can thiệp cao hơn lớp đối chứng 0,452 điểm trên thang 0-3 (SE HC3 = 0,082; KTC 95% [0,292; 0,612]; $p < 0,001$). Điểm sau ước tính lần lượt là 2,25 ở lớp đối chứng và 2,70 ở lớp thực nghiệm. Kích thước hiệu ứng trước-sau có nhóm đối chứng đạt $g = 1,35$.

Khi tính thêm thành tích Địa lí trước can thiệp, chênh lệch giữa hai lớp còn 0,428 điểm (KTC 95% [0,270; 0,586]; $p < 0,001$) và vẫn giữ cùng chiều. Hình 1 thể hiện sự thay đổi điểm Y2 trước và sau can thiệp của hai lớp; kết quả so sánh chính được xác định bằng ANCOVA.

Hình 1: Điểm lập luận Địa lí có căn cứ (Y2) trước và sau can thiệp ở hai nhóm



3.2.3. Đánh giá độ tin cậy của thông tin Địa lí (Y1)

Sau khi điều chỉnh theo điểm Y1 trước can thiệp, lớp thực nghiệm có điểm Y1 sau can thiệp cao hơn lớp đối chứng 0,455 điểm (SE HC3 = 0,078; KTC 95% [0,303; 0,607]; $p < 0,001$). Điểm sau ước tính lần lượt là 1,94 ở lớp đối chứng và 2,40 ở lớp thực nghiệm. Kích thước hiệu ứng trước-sau có nhóm đối chứng đạt $g = 1,13$.

Khi tính thêm thành tích Địa lí trước can thiệp, chênh lệch giữa hai lớp còn 0,430 điểm (KTC 95% [0,277; 0,583]; $p < 0,001$) và vẫn giữ cùng chiều. Kết quả ở cả Y1 và Y2 đều cho thấy lớp thực nghiệm có mức cải thiện cao hơn lớp đối chứng trong phạm vi mẫu nghiên cứu.

Bảng 1: Kết quả ANCOVA và phân tích bổ sung

Biến kết quả	Mô hình phân tích	Chênh lệch điều chỉnh	SE HC3	KTC 95%	p
Y2	Điều chỉnh theo điểm Y2 trước can thiệp	0,452	0,082	[0,292; 0,612]	< 0,001
Y2	Bổ sung thành tích Địa lí trước can thiệp	0,428	0,081	[0,270; 0,586]	< 0,001
Y1	Điều chỉnh theo điểm Y1 trước can thiệp	0,455	0,078	[0,303; 0,607]	< 0,001
Y1	Bổ sung thành tích Địa lí trước can thiệp	0,430	0,078	[0,277; 0,583]	< 0,001

Ghi chú: Chênh lệch điều chỉnh được tính theo hướng thực nghiệm - đối chứng. Hedges' g là kích thước hiệu ứng trước-sau có nhóm đối chứng theo Morris (2008); không tính cho phân tích bổ sung.

3.2.4. Kết quả theo từng thành phần và quá trình thực hiện nhiệm vụ

Cả 8 thành phần của Y1 và Y2 đều có chênh lệch theo hướng lớp thực nghiệm đạt điểm cao hơn lớp đối chứng sau khi điều chỉnh điểm trước can thiệp. Với Y1, mức chênh lệch dao động từ 0,351 đến 0,512 điểm; với Y2, từ 0,330 đến 0,463 điểm. Các kết quả chỉ được dùng để xem xét xu hướng thay đổi ở từng thành phần, không nhằm xác định riêng thành phần nào tạo ra hiệu quả.

Y1-S4 và Y2-A1 có mức chênh lệch thấp hơn tương đối so với các thành phần còn lại. Kết quả gợi ý rằng việc điều chỉnh phán đoán và bước đầu kết hợp bằng chứng có thể cần thêm thời gian luyện tập. Nhận định chỉ mang tính gợi ý và cần được kiểm tra thêm bằng công cụ đo được thẩm định đầy đủ.

Bảng 2: Chênh lệch điều chỉnh theo từng thành phần của Y1 và Y2

Thành phần	Chênh lệch	KTC 95%	p
Y1-S1	0,512	[0,290; 0,735]	< 0,001
Y1-S2	0,512	[0,279; 0,745]	< 0,001
Y1-S3	0,416	[0,197; 0,635]	< 0,001
Y1-S4	0,351	[0,075; 0,627]	0,013
Y2-A1	0,330	[0,127; 0,533]	0,001
Y2-A2	0,434	[0,205; 0,663]	< 0,001
Y2-A3	0,403	[0,153; 0,653]	0,002
Y2-A4	0,463	[0,253; 0,673]	< 0,001

Dữ liệu về quá trình thực hiện nhiệm vụ cũng cho thấy sự khác biệt giữa hai lớp. HS lớp thực nghiệm sử dụng trung bình 3,15 nguồn, trong đó có 2,56 nguồn độc lập; ở lớp đối chứng, các giá trị tương ứng là 2,17 và 1,76. Điểm tích hợp bằng chứng của lớp thực nghiệm đạt $2,15 \pm 0,69$, cao hơn lớp đối chứng với $1,22 \pm 0,72$. Tỷ lệ hoàn thành kiểm chứng ở hai lớp đều cao, đạt 95,1% ở lớp thực nghiệm và 90,2% ở lớp đối chứng. Các chỉ số được dùng để mô tả cách HS thực hiện nhiệm vụ, chưa được xem là thang đo độc lập đã được thẩm định.

Bảng 3: Một số chỉ báo về quá trình thực hiện nhiệm vụ

Chỉ báo	Đối chứng	Thực nghiệm
Số nguồn đã sử dụng	$2,17 \pm 1,00$	$3,15 \pm 0,76$
Số nguồn độc lập	$1,76 \pm 0,97$	$2,56 \pm 0,84$
Điểm tích hợp bằng chứng (0-3)	$1,22 \pm 0,72$	$2,15 \pm 0,69$
Hoàn thành kiểm chứng, n (%)	37/41 (90,2%)	39/41 (95,1%)

Ghi chú: Các chỉ báo được sử dụng để mô tả quá trình thực hiện nhiệm vụ. Giá trị ở ba dòng đầu được trình bày dưới dạng $M \pm SD$.

3.3. Thảo luận

Sau khi điều chỉnh điểm Y2 trước can thiệp, lớp thực nghiệm có điểm lập luận Địa lí có căn cứ cao hơn lớp đối chứng khoảng 0,45 điểm trên thang 0-3; Y1 cũng có mức chênh lệch gần tương tự. Kết quả phù hợp với các nghiên cứu cho thấy HS đánh giá nguồn tốt hơn khi được hướng dẫn cụ thể cách kiểm tra, đối chiếu và mở rộng sang các nguồn độc lập thay vì chỉ được cung cấp kiến thức về thông tin sai lệch (Fendt et al., 2025; Hämäläinen et al., 2023; Marten & Stadler, 2025; McGrew, 2024; Pimentel, 2025). Hoạt động học không chỉ yêu cầu HS xác định mức độ đáng tin cậy của nguồn mà còn sử dụng

bằng chứng đã kiểm tra để điều chỉnh nhận định và xây dựng lập luận Địa lí. Dữ liệu quá trình cho thấy lớp thực nghiệm sử dụng nhiều nguồn hơn, có số nguồn độc lập cao hơn và đạt điểm tích hợp bằng chứng tốt hơn lớp đối chứng. Kết quả phản ánh sự chuyển dịch từ việc tìm một nguồn để xác nhận thông tin sang đối chiếu nhiều nguồn và cân nhắc bằng chứng trước khi đưa ra kết luận. Barzilai và cộng sự (2020) cho thấy các công cụ hỗ trợ số có thể giúp HS đánh giá nguồn và kết hợp thông tin từ nhiều tài liệu. Hämäläinen và cộng sự (2023) ghi nhận việc hướng dẫn khai thác thông tin về nguồn góp phần cải thiện khả năng đánh giá độ tin cậy, trong khi Marten và Stadtler (2025) cho thấy việc rèn luyện chiến lược đánh giá thông tin trực tuyến mang lại kết quả tích cực ở HS trung học cơ sở. Các chỉ báo về quá trình thực hiện nhiệm vụ chỉ được dùng để mô tả sự khác biệt giữa hai lớp, chưa đủ cơ sở để xác định mức độ đóng góp trực tiếp của từng chỉ báo vào sự thay đổi kết quả học tập.

Sự khác biệt giữa hai lớp không chỉ liên quan đến khả năng đánh giá thông tin nói chung mà còn gắn với yêu cầu của môn Địa lí. HS phải kiểm tra thời điểm, chỉ tiêu, đơn vị lãnh thổ và quy mô phân tích trước khi quyết định một nhận định được dữ liệu hỗ trợ đến mức nào. Bendl và cộng sự (2025) cho thấy tư duy Địa lí gắn với quan hệ không gian, quy mô và các khái niệm đặc trưng của môn học. Các nghiên cứu về lập luận trong dạy học Địa lí cũng chỉ ra khó khăn của HS khi phải sử dụng nhiều tài liệu và kết nối dữ liệu với lí lẽ (Gebele et al., 2022; Hindmarsh & Budke, 2023; Schwerdtfeger & Budke, 2021). Y1 và Y2 vì thế cần được xem là hai yêu cầu có liên hệ nhưng không đồng nhất. Y1 phản ánh khả năng xác định nguồn và bằng chứng có thể sử dụng, còn Y2 đòi hỏi HS giải thích mối liên hệ giữa bằng chứng và kết luận theo không gian, thời gian, quy mô và quan hệ địa lí.

Cả hai lớp đều làm việc với thông tin do AI tạo ra nên AI không phải yếu tố tạo nên sự khác biệt trực tiếp giữa hai lớp. Điểm khác biệt nằm ở cách giáo viên tổ chức nhiệm vụ kiểm tra thông tin. Thay vì tiếp nhận đầu ra AI như một đáp án có sẵn, HS được yêu cầu kiểm tra và đối chiếu thông tin trước khi sử dụng. Việc tổ chức nhiệm vụ theo hướng kiểm tra, đánh giá và sử dụng đầu ra AI phù hợp với yêu cầu về sử dụng GenAI trong giáo dục (Miao & Holmes, 2023; Park, 2025). Trong dạy học Địa lí, giáo viên có thể lựa chọn các thông tin do AI tạo ra với mức độ chính xác khác nhau, yêu cầu HS đối chiếu ít nhất hai nguồn độc lập, kết hợp bản đồ, bảng số liệu hoặc nguồn thống kê, đồng thời xem xét sự phù hợp về không gian, thời gian, chỉ tiêu và quy mô trước khi đưa ra kết luận. Hoạt động kiểm tra thông tin số cũng phù hợp với yêu cầu đánh giá dữ liệu và thông tin trong Khung năng lực số cho người học (Bộ Giáo dục và Đào tạo, 2025a), trong khi mục tiêu của hoạt động học vẫn cần bám sát yêu cầu cần đạt của môn Địa lí trong chương trình hiện hành (Bộ Giáo dục và Đào tạo, 2025b).

Phạm vi suy luận của kết quả còn hạn chế vì mẫu chỉ gồm hai lớp tại một trường, mỗi lớp tương ứng với một điều kiện. Ảnh hưởng của đặc điểm lớp chưa thể tách hoàn toàn khỏi ảnh hưởng của can thiệp. Y2 cũng có chênh lệch trước can thiệp với $g = -0,42$; ANCOVA và phân tích bổ sung theo thành tích Địa lí giúp điều chỉnh một phần khác biệt nhưng không thay thế được thiết kế ngẫu nhiên. Can thiệp kéo dài bốn tiết và chưa có đánh giá trì hoãn nên chưa xác định được mức duy trì hoặc khả năng chuyển sang nhiệm vụ khác. Công cụ đánh giá cũng chưa có đầy đủ bằng chứng về mức tương đương giữa hai dạng đề và độ tin cậy giữa người chấm; các chỉ báo quá trình mới được dùng ở mức mô tả. Các nghiên cứu tiếp theo cần mở rộng số lớp và số trường, kiểm tra độ tương đương của công cụ, sử dụng chấm độc lập và bổ sung đánh giá trì hoãn cùng nhiệm vụ chuyển giao.

4. Kết luận

Trong mẫu 82 HS lớp 11, lớp thực nghiệm có mức cải thiện cao hơn lớp đối chứng ở cả khả năng đánh giá độ tin cậy của thông tin Địa lí và lập luận Địa lí có căn cứ. Sau khi điều chỉnh điểm trước can thiệp, chênh lệch giữa hai lớp đạt khoảng 0,45 điểm trên thang 0-3 ở cả hai biến. Khi bổ sung thành tích Địa lí trước can thiệp vào phân tích, hướng kết quả vẫn được giữ nguyên. Lớp thực nghiệm cũng sử dụng nhiều nguồn độc lập hơn và đạt điểm tích hợp bằng chứng cao hơn.

Thông tin Địa lí do AI tạo ra có thể được đưa vào hoạt động học như nội dung cần kiểm tra và đối chiếu, thay vì được tiếp nhận ngay như một câu trả lời đúng. HS cần xem xét nguồn, so sánh bằng chứng, nhận ra giới hạn của dữ liệu và điều chỉnh kết luận trước khi xây dựng lập luận Địa lí. Cách tổ

chức nhiệm vụ theo nhiều nguồn còn tạo điều kiện để HS kết hợp việc đánh giá thông tin số với các yêu cầu đặc thù của môn Địa lí như xem xét không gian, thời gian, quy mô, chỉ tiêu và phạm vi lãnh thổ.

Trong dạy học, giáo viên có thể sử dụng đầu ra AI để xây dựng các nhiệm vụ yêu cầu HS kiểm tra thông tin bằng bản đồ, bảng số liệu, nguồn thống kê và các tài liệu độc lập. Trọng tâm của hoạt động không nằm ở việc xác định AI đúng hay sai một cách đơn giản mà ở quá trình HS tìm bằng chứng, so sánh các nguồn và giải thích căn cứ cho kết luận. Cách tổ chức như vậy giúp giảm việc sử dụng AI với hoạt động học Địa lí, đồng thời duy trì vai trò chủ động của HS trong quá trình xử lí và đánh giá thông tin.

Do mỗi điều kiện chỉ có một lớp và công cụ đánh giá còn hạn chế về bằng chứng thẩm định, kết quả chỉ phản ánh hiệu quả bước đầu trong phạm vi mẫu. Thời gian can thiệp bốn tiết cũng chưa cho phép xác định mức độ duy trì của kết quả hoặc khả năng vận dụng sang những nhiệm vụ khác. Các nghiên cứu tiếp theo cần triển khai trên nhiều lớp và nhiều trường, hoàn thiện công cụ đánh giá, bổ sung đánh giá tri hoãn và nhiệm vụ chuyển giao để kiểm tra mức độ ổn định của kết quả.

Tài liệu tham khảo

- Barzilai, S., Mor-Hagani, S., Zohar, A. R., Shlomi-Elouz, T., & Ben-Yishai, R. (2020). Making sources visible: Promoting multiple document literacy with digital epistemic scaffolds. *Computers & Education*, 157, 103980. <https://doi.org/10.1016/j.compedu.2020.103980>.
- Bendl, T., Krajiňáková, L., Marada, M., & Řezníčková, D. (2025). Geographical thinking in geography education: A systematic review. *International Research in Geographical and Environmental Education*, 34(4), 326-352. <https://doi.org/10.1080/10382046.2024.2354097>.
- Bộ Giáo dục và Đào tạo (2025a). *Thông tư số 02/2025/TT-BGDĐT ngày 24/01/2025 quy định Khung năng lực số cho người học*.
- Bộ Giáo dục và Đào tạo (2025b). *Thông tư số 17/2025/TT-BGDĐT ngày 12/9/2025 sửa đổi, bổ sung một số nội dung trong Chương trình giáo dục phổ thông ban hành kèm theo Thông tư số 32/2018/TT-BGDĐT ngày 26/12/2018 của Bộ trưởng Bộ Giáo dục và Đào tạo*.
- Breakstone, J., Smith, M. D., Wineburg, S., Rapaport, A., Carle, J., Garland, M., & Saavedra, A. (2021). Students' civic online reasoning: A national portrait. *Educational Researcher*, 50(8), 505-515. <https://doi.org/10.3102/0013189X211017495>.
- Engelen, E., & Budke, A. (2022). Promoting geographic internet searches and subsequent argumentation using an Open Educational Resource. *Computers and Education Open*, 3, 100090. <https://doi.org/10.1016/j.caeo.2022.100090>.
- Fendt, M., Muth, X., & Edelsbrunner, P. A. (2025). Judging a text by its author-A meta-analysis of interventions to foster source credibility assessment. *Learning and Individual Differences*, 124, 102782. <https://doi.org/10.1016/j.lindif.2025.102782>.
- Fulsher, A., Pagkratidou, M., & Kendeou, P. (2026). GenAI and misinformation in education: A systematic scoping review of opportunities and challenges. *AI & Society*, 41(2), 1373-1385. <https://doi.org/10.1007/s00146-025-02536-y>.
- Gebele, D., Zepter, A. L., Königs, P., & Budke, A. (2022). Metacognition in argumentative writing based on multiple sources in geography education. *European Journal of Investigation in Health, Psychology and Education*, 12(8), 948-974. <https://doi.org/10.3390/ejihpe12080069>.
- Hämäläinen, E. K., Kiili, C., Räikkönen, E., Lakkala, M., Ilomäki, L., Toom, A., & Marttunen, M. (2023). Teaching sourcing during online inquiry-Adolescents with the weakest skills benefited the most. *Instructional Science*, 51(1), 135-163. <https://doi.org/10.1007/s11251-022-09597-2>.
- Hindmarsh, K., & Budke, A. (2023). What are the competences & difficulties of school students in writing material-based & multi-perspective argumentations in geography classrooms? *European Journal of Geography*, 14(3), 55-67. <https://doi.org/10.48088/ejg.k.hin.14.3.055.067>.
- Long, J. S., & Ervin, L. H. (2000). Using heteroscedasticity consistent standard errors in the linear regression model. *The American Statistician*, 54(3), 217-224. <https://doi.org/10.1080/00031305.2000.10474549>.
- Marten, P. L., & Stadler, M. (2025). Building resilience against online misinformation: A teacher-led training promoting evaluation strategies among lower secondary students. *Computers in Human Behavior*, 165, 108548. <https://doi.org/10.1016/j.chb.2024.108548>.
- McGrew, S. (2024). Teaching lateral reading: Interventions to help people read like fact checkers. *Current Opinion in Psychology*, 55, 101737. <https://doi.org/10.1016/j.copsy.2023.101737>.
- Miao, F., & Holmes, W. (2023). *Guidance for generative AI in education and research*. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000386693>.
- Morris, S. B. (2008). Estimating effect sizes from pretest-posttest-control group designs. *Organizational Research Methods*, 11(2), 364-386. <https://doi.org/10.1177/1094428106291059>.
- Park, J. (2025). A systematic literature review of generative artificial intelligence (GenAI) literacy in schools. *Computers and Education: Artificial Intelligence*, 9, 100487. <https://doi.org/10.1016/j.caeai.2025.100487>.
- Pimentel, D. R. (2025). Learning to evaluate sources of science (mis)information on the internet: Assessing students' scientific online reasoning. *Journal of Research in Science Teaching*, 62(3), 684-720. <https://doi.org/10.1002/tea.21974>.
- Schwerdtfeger, S., & Budke, A. (2021). Reference to materials in written argumentations of students in geography lessons. *Journal of Curriculum and Teaching*, 10(3), 20-35. <https://doi.org/10.5430/jct.v10n3p20>.
- Sparks, J. R., van Rijn, P. W., & Deane, P. (2021). Assessing source evaluation skills of middle school students using learning progressions. *Educational Assessment*, 26(4), 213-240. <https://doi.org/10.1080/10627197.2021.1966299>.
- Vickers, A. J., & Altman, D. G. (2001). Analysing controlled trials with baseline and follow up measurements. *BMJ*, 323(7321), 1123-1124. <https://doi.org/10.1136/bmj.323.7321.1123>.