

ĐẶC ĐIỂM THIẾT KẾ NGHIÊN CỨU VÀ KỸ THUẬT PHÂN TÍCH DỮ LIỆU TRONG CÁC NGHIÊN CỨU ĐỊNH LƯỢNG THUỘC LĨNH VỰC KHOA HỌC GIÁO DỤC TẠI VIỆT NAM (2020-2024)

CHARACTERISTICS OF RESEARCH DESIGN AND DATA ANALYSIS TECHNIQUES IN QUANTITATIVE STUDIES IN EDUCATIONAL SCIENCES IN VIETNAM (2020-2024)

Trịnh Thị Phương Thảo^{1,+},
Hứa Thị Toàn²,
Lê Minh Cường³

¹Trường Đại học Sư phạm - Đại học Thái Nguyên;

²Trường Đại học Nông Lâm - Đại học Thái Nguyên;

³Trường Đại học Đồng Tháp

+Tác giả liên hệ • Email: thaottp@tmue.edu.vn

Article history

Received: 28/02/2026

Accepted: 25/3/2026

Published: 20/5/2026

Keywords

Quantitative research,
research design, data analysis
techniques, educational
science

ABSTRACT

Quantitative research has been increasingly adopted in educational sciences due to its capacity to quantify data and systematically test hypotheses. This study aims to examine the characteristics of research design, sampling strategies, data collection methods, and analytical techniques in quantitative studies within the field of educational sciences in Vietnam. Secondary data analysis was conducted on articles published in nine domestic academic journals during the 2020-2024 period. The findings indicate a gradual growth in the volume of quantitative studies over time. Most studies employed non-probability sampling, questionnaire-based surveys, and traditional statistical methods, while a smaller number have begun to utilize digital data sources and apply more advanced analytical approaches. These results provide an overall picture of the characteristics of quantitative research methods in educational sciences in Vietnam and highlight the need to enhance transparency in methodological reporting, expand sample sizes, and promote the application of advanced analytical models in future research.

1. Mở đầu

Trong bối cảnh đổi mới căn bản và toàn diện giáo dục, nghiên cứu khoa học giáo dục (KHGD) ngày càng khẳng định vai trò then chốt trong việc cung cấp các bằng chứng tin cậy phục vụ hoạch định chính sách và cải tiến thực tiễn (Iftimescu và cộng sự, 2020; Tran và cộng sự, 2024). Để đáp ứng yêu cầu này, các nghiên cứu KHGD chú trọng vận dụng các lý thuyết thống kê và kỹ thuật phân tích dữ liệu trong thiết kế nghiên cứu, trong đó phương pháp nghiên cứu định lượng (NCĐL) được sử dụng rộng rãi nhằm nâng cao tính khách quan và độ tin cậy của kết quả nghiên cứu (Trần Trung, 2025). NCĐL là cách tiếp cận dựa trên các biến đo lường được và dữ liệu số, sử dụng các kỹ thuật thống kê để mô tả và phân tích mối quan hệ giữa các biến, kiểm định hiệu quả của các tác động hoặc giải pháp giáo dục, từ đó rút ra những kết luận có khả năng khái quát cao (Creswell, 2017).

Trên thế giới, NCĐL đã trở thành một nền tảng quan trọng trong KHGD, không chỉ phục vụ kiểm định giả thuyết mà còn góp phần nâng cao tính chặt chẽ phương pháp và độ tin cậy của bằng chứng thực nghiệm (Putra và Lafayette, 2023). Xu hướng nổi bật trong giai đoạn gần đây là sự đa dạng hóa các mô hình phân tích dữ liệu đi cùng với yêu cầu chuẩn hóa quy trình thiết kế nghiên cứu nhằm nâng cao khả năng đối chiếu kết quả giữa các nghiên cứu và tạo cơ sở cho việc phát triển tri thức khoa học (Claro và cộng sự, 2024; Tan và Gao, 2025). Tổng quan của Mahat và cộng sự (2024) trên 30 công trình giai đoạn 1999-2023 cho thấy, các nghiên cứu ngày càng chú trọng chuẩn hóa thiết kế, chiến lược chọn mẫu và xác định cỡ mẫu tối thiểu, qua đó nâng cao độ tin cậy và tính giá trị của kết quả. Đồng thời, nhiều phân tích phương pháp nghiên cứu gần đây cũng nhấn mạnh sự phù hợp giữa giả thuyết nghiên cứu, kỹ thuật phân tích thống kê và mức độ nghiêm ngặt của thiết kế nghiên cứu (McCrudden và cộng sự, 2025; Zhang, 2025). Nhìn chung, các xu hướng quốc tế cho thấy NCĐL trong KHGD đang chuyển dịch theo hướng chuẩn hóa quy trình, tăng cường tính minh bạch và nâng cao chất lượng bằng chứng thực nghiệm.

Tại Việt Nam, một số nghiên cứu tổng quan đã bước đầu phản ánh sự phát triển của NCĐL trong lĩnh vực KHGD, cho thấy loại hình nghiên cứu này ngày càng chiếm ưu thế trong các công bố khoa học những năm gần đây (Nguyễn Tiến Trung và cộng sự, 2024; Phan Thị Thanh Thảo và cộng sự, 2022). Bên cạnh đó, thực trạng và định hướng phát

triển của lĩnh vực KHGD cũng được phân tích thông qua việc mô tả các hướng nghiên cứu và loại hình phương pháp chủ yếu (Nguyễn Đức Minh và cộng sự, 2023). Tuy nhiên, các nghiên cứu hiện có phần lớn mới dừng lại ở mức khái quát về xu hướng công bố hoặc phân loại phương pháp chung, trong khi những đặc trưng kỹ thuật của NCDL như loại hình thiết kế nghiên cứu, chiến lược chọn mẫu, quy mô mẫu và các công cụ phân tích dữ liệu cụ thể vẫn chưa được xem xét một cách hệ thống. Việc thiếu vắng những phân tích chi tiết về phương pháp nghiên cứu làm cho bức tranh toàn diện về cách thức triển khai NCDL trong KHGD Việt Nam chưa thực sự rõ nét.

Xuất phát từ khoảng trống nêu trên, nghiên cứu này tiến hành phân tích có hệ thống các ấn phẩm thuộc lĩnh vực KHGD đăng trên 9 tạp chí khoa học tiêu biểu tại Việt Nam giai đoạn 2020-2024 nhằm nhận diện và mô tả các đặc điểm thiết kế nghiên cứu và kỹ thuật phân tích dữ liệu trong NCDL. Thông qua phương pháp phân tích nội dung, nghiên cứu hướng tới việc làm rõ xu hướng vận dụng các chiến lược thiết kế nghiên cứu và công cụ thống kê trong thực tiễn nghiên cứu KHGD, qua đó góp phần cung cấp cơ sở tham khảo cho các nhà nghiên cứu, cơ quan quản lý và các tạp chí học thuật trong việc định hướng và phát triển các phương pháp nghiên cứu phù hợp với bối cảnh hội nhập quốc tế.

2. Phương pháp nghiên cứu

2.1. Nguồn dữ liệu và phạm vi khảo sát

Nghiên cứu thực hiện phân tích dữ liệu thứ cấp, kết hợp thống kê mô tả nhằm khảo sát đặc điểm của NCDL trong các công bố thuộc lĩnh vực KHGD tại Việt Nam giai đoạn 2020-2024. Nguồn dữ liệu bao gồm toàn bộ các ấn phẩm về KHGD được công bố trên 09 tạp chí khoa học trong nước giai đoạn này bao gồm: Tạp chí Giáo dục thuộc Bộ GD-ĐT (VJE); Tạp chí Khoa học Giáo dục Việt Nam (VJES); Tạp chí Khoa học, Trường Đại học Sư phạm Hà Nội (HNUEJ); Chuyên san Nghiên cứu Giáo dục thuộc Tạp chí Khoa học của Đại học Quốc gia Hà Nội (VNUJ); Tạp chí Khoa học Trường Đại học Sư phạm Hà Nội 2: Chuyên san Khoa học Giáo dục (HPU2J); Tạp chí Khoa học và Công nghệ Đại học Thái Nguyên (TNUJ); Tạp chí Khoa học Đại học Vinh (VUJS); Tạp chí Khoa học Đại học Huế (HUEJ) và Tạp chí Khoa học Trường Đại học Sư phạm TP. Hồ Chí Minh (HCMUEJ). Việc lựa chọn 09 tạp chí này nhằm bảo đảm nguồn dữ liệu nghiên cứu có tính đại diện cho các cơ sở đào tạo và nghiên cứu giáo dục ở cả ba miền Bắc, Trung, Nam. Đây đều là những tạp chí có uy tín học thuật, định hướng công bố phù hợp với mục tiêu và phạm vi của nghiên cứu. Giai đoạn 2020-2024 được lựa chọn để thu thập dữ liệu vì đây là giai đoạn Chương trình Giáo dục phổ thông 2018 được triển khai trong thực tiễn, qua đó bước đầu bộc lộ những vấn đề cần tiếp tục xem xét và đánh giá (Thuyet, 2024). Đây cũng là giai đoạn giáo dục Việt Nam đẩy mạnh chuyển đổi số (Lương Đình Hải, 2023). Việc xác định phạm vi này giúp nghiên cứu tập trung vào các công bố xuất hiện trong bối cảnh thực tiễn giáo dục gần đây với nhiều thay đổi đáng chú ý.

Bộ dữ liệu ban đầu của nghiên cứu gồm 4.476 ấn phẩm, được các tạp chí cung cấp dưới dạng tệp Excel, bao gồm các thông tin cơ bản như: Tiêu đề, tác giả, đơn vị công tác, năm công bố, từ khóa và tóm tắt. Đây là cơ sở để nghiên cứu triển khai quy trình sàng lọc, chuẩn hóa và phân tích dữ liệu ở bước tiếp theo.

2.2. Quy trình sàng lọc, chuẩn hóa và phân tích dữ liệu

Quy trình sàng lọc và chuẩn hóa được thực hiện thông qua 3 bước như sau:

Bước 1: Loại bỏ các bài không thuộc lĩnh vực KHGD. Ở bước này, nhóm nghiên cứu sử dụng thông tin phân loại của ấn phẩm trên website của tạp chí để xác định phạm vi học thuật của từng bài. Kết quả, 02 ấn phẩm trùng lặp và 655 ấn phẩm không thuộc lĩnh vực KHGD bị loại bỏ.

Bước 2: Xác định các ấn phẩm sử dụng phương pháp NCDL. Từ 3.819 ấn phẩm thu được sau Bước 1, nhóm nghiên cứu tiến hành tải và đọc toàn văn. Việc phân loại ấn phẩm thuộc NCDL dựa trên nghiên cứu của Creswell (2017) gồm 3 tiêu chí: (1) có mô tả mẫu nghiên cứu với thông tin về quy mô, đặc điểm đối tượng hoặc tiêu chí lựa chọn mẫu; (2) sử dụng các công cụ thu thập dữ liệu tạo ra dữ liệu số như bảng hỏi, thang đo, bài kiểm tra hoặc các nền tảng khảo sát trực tuyến; (3) trình bày quy trình xử lý thống kê hoặc mô hình hóa dữ liệu trong phần phân tích kết quả. Trong nghiên cứu này, các ấn phẩm sử dụng phương pháp nghiên cứu hỗn hợp không được đưa vào phạm vi phân tích. Sau bước 2, có 620 bài đáp ứng đầy đủ tiêu chí được lựa chọn.

Bước 3: Mã hóa chi tiết. Bảng mã hóa được xây dựng dựa trên thông tin thu thập trực tiếp từ các ấn phẩm trong bộ dữ liệu, bao gồm:

- Phương pháp chọn mẫu: Chọn mẫu phi xác suất (chọn mẫu thuận tiện, chọn mẫu có chủ đích, chọn mẫu tình nguyện hoặc hạn ngạch); chọn mẫu xác suất (chọn mẫu ngẫu nhiên đơn giản, ngẫu nhiên phân tầng, chọn mẫu theo cụm hoặc nhiều giai đoạn); và trường hợp không nêu rõ phương pháp chọn mẫu đối với các nghiên cứu trắc lượng thư mục hoặc sử dụng dữ liệu thứ cấp.

- Quy mô mẫu: Số lượng đối tượng khảo sát/thực nghiệm hoặc số lượng tài liệu đối với các nghiên cứu trắc lượng thư mục.

- Công cụ và phần mềm xử lý dữ liệu: Phần mềm phân tích thống kê (SPSS, Excel, R, SmartPLS, Conquest,...) và các nền tảng thu thập dữ liệu trực tuyến (Google Forms, SurveyMonkey,...).

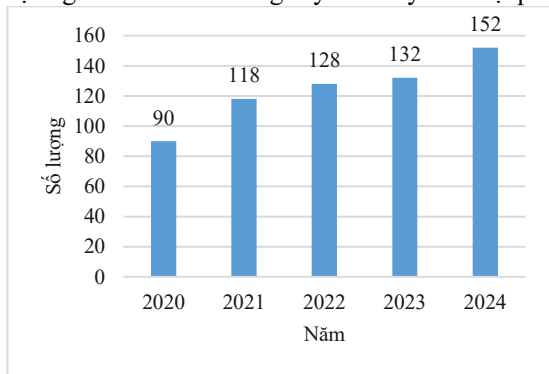
- Kỹ thuật phân tích dữ liệu: (1) thống kê mô tả (tần suất, tỉ lệ phần trăm, giá trị trung bình và độ lệch chuẩn); (2) phân tích độ tin cậy và giá trị thang đo (Cronbach's Alpha, EFA, CFA); (3) phân tích so sánh và kiểm định giả thuyết (T-test, ANOVA, Chi-square, Mann-Whitney, Kruskal-Wallis); (4) phân tích mối quan hệ và dự báo (trương quan Pearson/Spearman, hồi quy tuyến tính đơn và đa biến, hồi quy logistic/probit, phân tích trung gian và mô hình dữ liệu bảng); (5) mô hình hóa nâng cao (mô hình phương trình cấu trúc (SEM/PLS-SEM), lý thuyết hồi đáp câu hỏi (IRT/Rasch) và phân tích hiệu quả (DEA); (6) trắc lượng thư mục. Ngoài ra, một số nghiên cứu sử dụng kỹ thuật khai phá dữ liệu và xử lý văn bản ở mức độ khác nhau. Mỗi kỹ thuật được mã hóa theo tên gọi và mục đích sử dụng chính nhằm bảo đảm tính nhất quán trong quá trình tổng hợp và so sánh.

Dữ liệu sau khi sàng lọc và chuẩn hóa được lưu trữ dưới dạng tệp Excel (.xlsx). Tiếp đó, các kỹ thuật thống kê mô tả được sử dụng để tổng hợp số liệu, làm cơ sở cho các phân tích chi tiết tiếp theo.

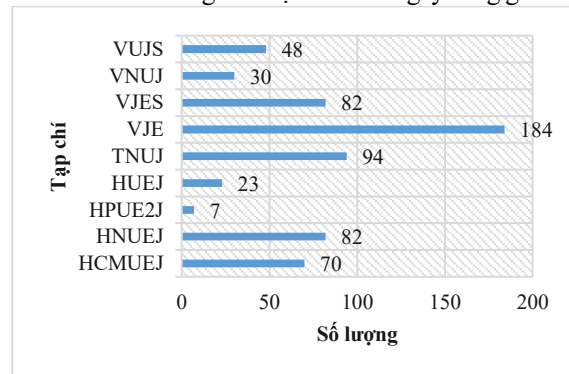
3. Kết quả nghiên cứu

3.1. Thông tin về xu hướng tăng trưởng của các ấn phẩm

Theo kết quả thống kê, giai đoạn 2020-2024 có tổng cộng 620 ấn phẩm sử dụng phương pháp NCDL được công bố. Hình 1 thể hiện sự tăng trưởng của các ấn phẩm theo năm. Từ 90 bài năm 2020, số ấn phẩm tăng lên đáng kể trong giai đoạn 2021-2022 và đạt 128 bài. Năm 2023 ghi nhận mức tăng chậm hơn với 132 bài, cho thấy sự ổn định tương đối sau giai đoạn tăng trưởng nhanh. Đến năm 2024, số lượng này tăng lên 152 bài, mức cao nhất trong giai đoạn nghiên cứu. Xu hướng này cho thấy mức độ quan tâm đối với NCDL trong lĩnh vực KHGD ngày càng gia tăng.



Hình 1. Số lượng ấn phẩm theo năm



Hình 2. Số lượng ấn phẩm theo tạp chí

Hình 2 thể hiện sự phân bố số lượng ấn phẩm NCDL theo từng tạp chí. Kết quả cho thấy, VJE là tạp chí có số lượng công bố nhiều nhất với 184 bài, chiếm tỉ trọng lớn trong tổng số ấn phẩm. Tiếp theo là TNUJ với 94 bài; HNUEJ và VJES cùng đạt 82 bài. Các tạp chí còn lại có số lượng công bố giảm dần bao gồm: HCMUEJ (70 bài), VUJS (48 bài), VNUJ (30 bài), HUEJ (23 bài) và HPUE2J (7 bài). Sự khác biệt này một phần có thể được lý giải bởi đặc điểm và phạm vi xuất bản của từng tạp chí. Cụ thể, một số nguồn xuất bản tập trung vào lĩnh vực giáo dục, trong khi một số nguồn khác có phạm vi đa ngành. Vì vậy, số lượng ấn phẩm NCDL trong lĩnh vực KHGD được công bố trên các tạp chí là không đồng đều.

3.2. Đặc điểm thiết kế nghiên cứu của các nghiên cứu định lượng trong lĩnh vực khoa học giáo dục tại Việt Nam

3.2.1. Về phương pháp chọn mẫu

Kết quả thống kê cho thấy, các NCDL trong lĩnh vực KHGD tại Việt Nam chủ yếu sử dụng các phương pháp chọn mẫu phi xác suất. Trong nhóm này, chọn mẫu thuận tiện và chọn mẫu có chủ đích là hai hình thức được sử dụng phổ biến nhất, bên cạnh một số kỹ thuật phi xác suất khác như chọn mẫu tình nguyện, định ngạch hoặc quả cầu tuyết (Kiet, 2021; Tăng Đình Ngọc Thảo và cộng sự, 2021). Xu hướng này đặc biệt phổ biến trong các nghiên cứu khảo sát được triển khai tại các cơ sở giáo dục như trường phổ thông, trường đại học hoặc trong các chương trình đào tạo cụ thể. Ngược lại, các phương pháp chọn mẫu xác suất, bao gồm chọn mẫu ngẫu nhiên đơn giản, chọn mẫu phân tầng, chọn mẫu theo cụm hoặc chọn mẫu nhiều giai đoạn, xuất hiện với tần suất thấp hơn đáng kể. Những phương pháp này chủ yếu được áp dụng trong một số nghiên cứu có quy mô mẫu lớn hoặc phạm vi khảo sát diện

rộng, trong khi phần lớn các nghiên cứu khảo sát giáo dục vẫn ưu tiên các hình thức chọn mẫu phi xác suất do tính thuận tiện và khả năng triển khai thực tế.

Bên cạnh đó, một số nghiên cứu không áp dụng phương pháp chọn mẫu theo nghĩa thông thường mà tiến hành khảo sát toàn thể, đặc biệt trong các trường hợp phạm vi nghiên cứu được giới hạn trong một đơn vị cụ thể như một trường học, một khoa hoặc một nhóm đối tượng có quy mô nhỏ. Ngoài ra, đối với các nghiên cứu trắc lượng thư mục, dữ liệu thứ cấp được sử dụng thông qua việc khai thác các cơ sở dữ liệu học thuật lớn như Scopus, Web of Science hoặc Google Scholar. Đáng chú ý, có tới 45% số bài báo (279/620) không nêu rõ hoặc chỉ mô tả rất hạn chế phương pháp chọn mẫu. Việc thiếu minh bạch trong mô tả quy trình chọn mẫu gây khó khăn cho việc đánh giá tính đại diện của dữ liệu, mức độ khái quát hóa của kết quả nghiên cứu cũng như khả năng tái lập nghiên cứu.

3.2.2. Về hình thức thu thập dữ liệu

Theo thông tin từ bộ dữ liệu, khảo sát trực tuyến là hình thức thu thập dữ liệu được sử dụng phổ biến nhất trong các NCDL về KHGD tại Việt Nam. Trong nhóm này, Google Forms là nền tảng được sử dụng phổ biến nhờ tính miễn phí, dễ triển khai, khả năng phân phối linh hoạt và thuận tiện cho việc tổng hợp dữ liệu ban đầu. Bên cạnh Google Forms, một số nghiên cứu cũng sử dụng các nền tảng trực tuyến khác như Microsoft Teams, Zalo hoặc Facebook để gửi bảng hỏi và thu thập phản hồi. Đối với các nghiên cứu thực nghiệm sư phạm hoặc đánh giá quá trình học tập, dữ liệu còn được bổ sung từ các hệ thống quản lý học tập như: Moodle, Google Classroom hoặc các nền tảng học tập trực tuyến. Ngoài ra, một số nghiên cứu khai thác dữ liệu từ các công cụ học tập tương tác như Quizizz, Kahoot! và Mentimeter nhằm ghi nhận kết quả học tập, mức độ tham gia, cũng như phản hồi tức thời của người học trong quá trình triển khai hoặc hoạt động dạy học.

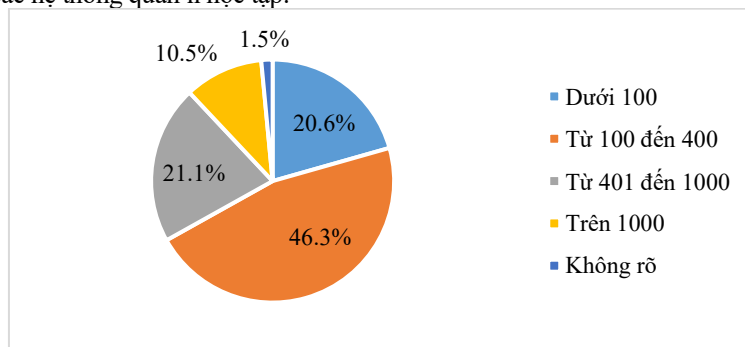
Bên cạnh dữ liệu sơ cấp, một số nghiên cứu sử dụng dữ liệu thứ cấp, chủ yếu được khai thác từ các cuộc điều tra quy mô quốc gia, báo cáo thống kê chính thức hoặc các cơ sở dữ liệu chuyên ngành trong và ngoài nước. Các nghiên cứu này không trực tiếp tiến hành thu thập dữ liệu mới mà tập trung vào phân tích lại các tập dữ liệu sẵn có nhằm nhận diện xu hướng, mối quan hệ hoặc sự thay đổi theo thời gian.

3.2.3. Về cỡ mẫu

Tỉ lệ cỡ mẫu theo các mức được trình bày trong hình 3. Kết quả cho thấy, cỡ mẫu của các NCDL phân bố không đồng đều, đồng thời phản ánh khá rõ đặc trưng thiết kế nghiên cứu và đối tượng khảo sát trong lĩnh vực KHGD.

Nhóm nghiên cứu có cỡ mẫu từ 100 đến 400 chiếm tỉ lệ cao nhất (46,3%). Đây là quy mô mẫu phổ biến trong các nghiên cứu khảo sát giáo dục, thường được triển khai tại một hoặc một số cơ sở giáo dục với các đối tượng nghiên cứu chủ yếu là người dạy và người học. Việc nhóm này chiếm ưu thế cho thấy, các NCDL trong lĩnh vực KHGD hiện nay chủ yếu tập trung vào việc mô tả thực trạng, khảo sát nhận thức, thái độ hoặc hành vi của các nhóm đối tượng trong phạm vi tương đối hẹp và thuận lợi cho tổ chức thu thập dữ liệu.

Nhóm nghiên cứu có cỡ mẫu từ 401 đến 1000 chiếm 21,1%, thường gắn với các khảo sát liên trường, liên cấp học hoặc khảo sát trên nhiều nhóm đối tượng khác nhau. So với nhóm trên, các nghiên cứu trong nhóm này có phạm vi triển khai rộng hơn, cho thấy xu hướng mở rộng đối tượng và bối cảnh khảo sát nhằm tăng khả năng khái quát hóa kết quả nghiên cứu. Trong nhóm này, hình thức thu thập dữ liệu chủ đạo là khảo sát trực tuyến, đôi khi kết hợp với dữ liệu trích xuất từ các hệ thống quản lý học tập.



Hình 3. Thông tin về cỡ mẫu trong bộ dữ liệu

Các nghiên cứu có cỡ mẫu dưới 100 chiếm 20,6%, chủ yếu xuất hiện trong các nghiên cứu trường học, nghiên cứu thử nghiệm công cụ đo, nghiên cứu thực nghiệm quy mô lớp học hoặc khảo sát ý kiến chuyên gia. Trong nhóm này, dữ liệu thường được thu thập thông qua bảng hỏi trực tiếp, phỏng vấn hoặc các hình thức khảo sát trực tuyến

với phạm vi hẹp. Trong khi đó, các nghiên cứu có cỡ mẫu trên 1000 chỉ chiếm 10,5%. Nhóm này chủ yếu bao gồm các nghiên cứu sử dụng dữ liệu thứ cấp, khảo sát diện rộng hoặc nghiên cứu trắc lượng thư mục. Tỷ lệ không cao của nhóm cỡ mẫu rất lớn cho thấy các NCDL trong lĩnh vực KHGD chưa thật sự phổ biến ở quy mô lớn, liên vùng hoặc cấp quốc gia, mà vẫn chủ yếu tập trung ở phạm vi cơ sở giáo dục hoặc nhóm đối tượng cụ thể.

Ngoài ra, có 9 nghiên cứu (1,5%) không báo cáo rõ cỡ mẫu hoặc không mô tả đầy đủ quy trình thu thập dữ liệu. Điều này phần nào cho thấy việc công bố thông tin về phương pháp trong một số nghiên cứu chưa đầy đủ, ảnh hưởng đến khả năng đánh giá độ tin cậy và mức độ khái quát của kết quả nghiên cứu.

3.3. Đặc điểm công cụ xử lý và kỹ thuật phân tích dữ liệu của các nghiên cứu định lượng trong lĩnh vực khoa học giáo dục tại Việt Nam

3.3.1. Về công cụ xử lý dữ liệu

Theo kết quả thống kê, SPSS là công cụ được sử dụng phổ biến nhất, đặc biệt trong các nghiên cứu khảo sát và thực nghiệm. Bên cạnh đó, một số phần mềm thống kê chuyên sâu như R/RStudio, Stata, JASP và ConQuest được lựa chọn trong các nghiên cứu có yêu cầu phân tích nâng cao, chẳng hạn như phân tích dữ liệu điều tra quy mô lớn, xử lý dữ liệu thứ cấp hoặc áp dụng các mô hình lý thuyết hồi đáp câu hỏi (IRT).

Đối với các nghiên cứu kiểm định mô hình lý thuyết, các phần mềm phục vụ mô hình cấu trúc tuyến tính (SEM) như SmartPLS và AMOS được sử dụng khá phổ biến. Đáng chú ý, một số nghiên cứu gần đây đã bắt đầu ứng dụng các thuật toán học máy như: Random Forest, Decision Tree hoặc Convolutional Neural Network (CNN) trong các bài toán dự đoán và phân tích dữ liệu phi cấu trúc (Nguyễn Bá Tiến và cộng sự, 2023; Nguyễn Việt Hưng và Trịnh Huy Hoàng, 2022).

3.3.2. Về kỹ thuật phân tích dữ liệu

Thống kê mô tả được sử dụng nhiều nhất và trong hầu hết các nghiên cứu. Các chỉ số như: Tần suất, tỉ lệ phần trăm, giá trị trung bình và độ lệch chuẩn được sử dụng phổ biến nhằm mô tả thực trạng, mức độ biểu hiện hoặc nhận thức của các nhóm đối tượng nghiên cứu. Nhóm kiểm định sự khác biệt cũng được sử dụng rộng rãi, bao gồm các phép kiểm định như T-test, ANOVA, Chi-square và các kiểm định phi tham số (Mann Whitney Test, Kruskal Wallis Test). Phân tích tương quan và hồi quy dùng trong các nghiên cứu nhằm xác định mối quan hệ giữa các biến, dự báo mức độ ảnh hưởng của các yếu tố độc lập lên biến phụ thuộc. Tuy nhiên, phần lớn các nghiên cứu dừng ở hồi quy tuyến tính cơ bản, với số lượng biến kiểm soát hạn chế.

Đối với các nghiên cứu sử dụng thang đo, phân tích độ tin cậy và phân tích nhân tố khám phá (EFA) được áp dụng phổ biến. Trong khi đó, phân tích nhân tố khẳng định (CFA) và mô hình cấu trúc tuyến tính (SEM) xuất hiện với tần suất thấp hơn, chủ yếu trong các nghiên cứu gần đây hoặc các công trình có yêu cầu kiểm định mô hình lý thuyết rõ ràng. Một số nghiên cứu chuyên biệt áp dụng lý thuyết hồi đáp câu hỏi (IRT), mô hình Rasch hoặc các kỹ thuật phân tích trắc lượng thư mục và phân tích mạng lưới, cho thấy sự đa dạng nhất định trong phương pháp phân tích dữ liệu.

4. Kết luận và bình luận

Các kết quả nghiên cứu cho thấy, số lượng các NCDL trong lĩnh vực KHGD tại Việt Nam giai đoạn 2020-2024 có xu hướng gia tăng rõ rệt. Điều này phù hợp với sự phát triển chung trên thế giới (Ketsman và cộng sự, 2025). Tuy nhiên, các đặc điểm về phương pháp nghiên cứu lại cho thấy một số khác biệt trong bối cảnh nghiên cứu tại Việt Nam.

Thứ nhất, các nghiên cứu trong nước chủ yếu sử dụng phương thức chọn mẫu phi xác suất, cho thấy quá trình lựa chọn mẫu thường phụ thuộc vào điều kiện thực tiễn của nhà nghiên cứu. Bên cạnh đó, tỉ lệ lớn các bài báo chưa mô tả rõ quy trình chọn mẫu. Điều này có thể ảnh hưởng đến việc đánh giá độ tin cậy cũng như tính khái quát của kết quả. Như vậy có thể thấy, việc đảm bảo tính đại diện thống kê của mẫu vẫn chưa phải là ưu tiên phổ biến trong các NCDL tại Việt Nam. Đây có thể được xem là một hạn chế so với các tiêu chuẩn phương pháp và yêu cầu minh bạch thường thấy trong nghiên cứu quốc tế, trong đó nhấn mạnh việc chuẩn hóa thiết kế nghiên cứu, chiến lược chọn mẫu và xác định cỡ mẫu tối thiểu nhằm nâng cao độ tin cậy và giá trị suy luận của NCDL (Mahat và cộng sự, 2024; McCrudden và cộng sự, 2025).

Thứ hai, các hình thức thu thập dữ liệu phản ánh xu hướng chuyển đổi số ngày càng rõ nét, khi dữ liệu từ các nền tảng trực tuyến và các công cụ khảo sát số được khai thác nhiều hơn. Trong đó, khảo sát bằng bảng hỏi trực tuyến nổi lên như một phương thức thu thập dữ liệu phổ biến. Theo Pregoner (2024), phân tích định lượng dựa trên dữ liệu bảng hỏi có ưu thế về khả năng khái quát hóa, nhưng vẫn tồn tại những hạn chế nhất định về chiều sâu bối cảnh. Sự phụ thuộc chủ yếu vào khảo sát bằng bảng hỏi cũng đặt ra yêu cầu kết hợp các phương pháp định tính nhằm phân tích đầy đủ hơn những quá trình học tập phức tạp, vượt ra ngoài các kết quả mô tả thuần túy của NCDL.

Thứ ba, phần lớn các nghiên cứu sử dụng cỡ mẫu ở mức trung bình, phản ánh những hạn chế về nguồn lực và khả năng tiếp cận đối tượng nghiên cứu so với bối cảnh quốc tế. Tuy vậy, đây cũng có thể xem là cơ hội để thúc đẩy các chiến lược mở rộng hợp tác nghiên cứu và khai thác dữ liệu thứ cấp nhằm nâng cao quy mô mẫu trong tương lai.

Cuối cùng, kết quả về công cụ và kỹ thuật phân tích dữ liệu cho thấy sự chiếm ưu thế của các phương pháp thống kê truyền thống như thống kê mô tả, kiểm định sự khác biệt và hồi quy tuyến tính cơ bản. Bên cạnh đó, một số nghiên cứu gần đây đã bắt đầu sử dụng học máy và các kỹ thuật khai phá văn bản, xử lý ngôn ngữ tự nhiên. Điều này cho thấy, bên cạnh các phương pháp thống kê truyền thống, trong các nghiên cứu được tổng hợp đã xuất hiện bước đầu một số phương pháp phân tích dữ liệu hiện đại.

Như vậy, bài báo đã cung cấp một bức tranh tổng quan về đặc điểm thiết kế nghiên cứu và kỹ thuật phân tích dữ liệu trong các NCDL thuộc lĩnh vực KHGD tại Việt Nam giai đoạn 2020-2024. So sánh với xu hướng quốc tế, NCDL tại Việt Nam đang từng bước tiệm cận với các nghiên cứu toàn cầu, đặc biệt trong việc ứng dụng công cụ phân tích dữ liệu và khai thác nền tảng số. Tuy nhiên, những hạn chế về tính đại diện của mẫu, mức độ minh bạch và sự phụ thuộc vào các kỹ thuật phân tích cơ bản cũng cho thấy sự cần thiết phải nâng cao hơn nữa về chất lượng nghiên cứu trên các tạp chí trong nước. Với các kết quả này, bài báo không chỉ cung cấp dữ liệu cập nhật về đặc điểm NCDL lĩnh vực KHGD tại Việt Nam mà còn mở rộng các nghiên cứu trước đây thông qua phân tích sâu cấu trúc phương pháp nghiên cứu thay vì chỉ tập trung vào sản lượng công bố. Các phát hiện từ bài báo có thể hỗ trợ các nhà nghiên cứu, biên tập viên tạp chí và cơ sở đào tạo trong việc định hướng phát triển phương pháp nghiên cứu, nâng cao chất lượng thiết kế nghiên cứu và chuẩn hóa quy trình báo cáo khoa học. Trong tương lai, hướng nghiên cứu có thể mở rộng phạm vi phân tích sang các nghiên cứu hỗn hợp và nghiên cứu định tính, đồng thời khai thác dữ liệu liên cơ sở hoặc dữ liệu quy mô quốc gia để tăng tính khái quát của kết quả. Ngoài ra, việc thúc đẩy sử dụng các mô hình phân tích nâng cao và tăng cường hợp tác nghiên cứu quốc tế có thể góp phần nâng cao vị thế của nghiên cứu KHGD Việt Nam trong bối cảnh toàn cầu hóa học thuật.

Tuyên bố về vai trò của các tác giả: *Trịnh Thị Phương Thảo: Lên ý tưởng, xác định phương pháp và công cụ nghiên cứu, giám sát, chỉ đạo quá trình nghiên cứu; Hứa Thị Toàn: Phân tích dữ liệu và viết bản thảo; Lê Minh Cường: Trục quan hoá dữ liệu, sửa chữa bản thảo.*

Tuyên bố về GenAI và Quyền tác giả: *Trong quá trình chuẩn bị bản thảo này, tác giả đã sử dụng ChatGPT (phiên bản GPT-5.2, OpenAI) nhằm hỗ trợ rà soát diễn đạt, tránh lặp từ và cải thiện tính mạch lạc của câu văn. Nội dung học thuật, lập luận và các kết luận khoa học trong bài viết do tác giả chịu trách nhiệm hoàn toàn.*

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

Lời cảm ơn: *Nhóm tác giả cảm ơn sự tài trợ của Bộ Giáo dục và đào tạo qua đề tài với mã số: B2025-TNA-19.*

Tài liệu tham khảo

- Claro, M., Castro-Grau, C., Ochoa, J. M., Hinojosa, J. E., & Cabello, P. (2024). Systematic review of quantitative research on digital competences of in-service school teachers. *Computers & Education, 215*, 1-15. <https://doi.org/10.1016/j.compedu.2024.105030>
- Creswell, J. W. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage Publications.
- Iftimescu, S., Ion, G., Proteasa, C., Iucu, R., Marin, E., & Stingu, M. (2020). Closing the circle: research and policymaking in education. *European Higher Education Area: Challenges for a New Decade*, 323-340. <https://doi.org/10.1007/978-3-030-56316-5>
- Ketsman, O., Droog, A., & Qazi, S. (2025). Mapping the prevalence of mixed methods research in educational technology journals. *Computers & Education, 226*, 1-13. <https://doi.org/10.1016/j.compedu.2024.105207>
- Kiet, N. T. (2021). Factors influencing lecturers' attitude towards working at universities after retirement age. *HCMUE Journal of Science, 18*(11), 1940-1952. [https://doi.org/10.54607/hcmue.js.18.11.3106\(2021\)](https://doi.org/10.54607/hcmue.js.18.11.3106(2021))
- Lương Đình Hải (2023). Chuyển đổi số trong giáo dục giai đoạn 2020-2022 : Nghiên cứu trắc lượng khoa học. *Tạp chí Khoa học Giáo dục Việt Nam, 19*(4), 1-7.
- Mahat, D., Neupane, D., & Shrestha, S. (2024). Quantitative research design and sample trends : A systematic examination of emerging paradigms and best practices. *Cognizance Journal of Multidisciplinary Studies, 4*(2), 20-27. <https://doi.org/10.47760/cognizance.2024.v04i02.002>

- McCrudden, M. T., Bowman, M., & Oaxaca, G. S. C. (2025). Reporting of methodological rigor in empirical mixed methods research in educational psychology. *Educational Psychology Review*, 37(4), 1-29. <https://doi.org/10.1007/s10648-025-10090-8>
- Nguyễn Bá Tiến, Phạm Hương Thảo, Lê Thị Kim Huyền, Hoàng Đăng Trị, Nguyễn Tiên Thảo, Đồng Thị Tuyết Hạnh, Nguyễn Quốc Chính (2023). Nghiên cứu chuyên đổi điểm bài thi đánh giá năng lực của Đại học Quốc gia Hà Nội và Đại học Quốc gia Thành phố Hồ Chí Minh. *Tạp chí Khoa học Giáo dục Việt Nam*, 19(1), 9-13. <https://doi.org/10.15625/2615-8957/12310102>
- Nguyễn Đức Minh, Nguyễn Lê Văn Dung, Nguyễn Thị Thanh Thủy, Phạm Thị Bích Đào (2023). Khoa học giáo dục Việt Nam-Thực tiễn đặt ra và xu hướng nghiên cứu giai đoạn 2023-2030. *Tạp chí Khoa học Giáo dục Việt Nam*, 3(19). <https://doi.org/10.15625/2615-8957/12310301>
- Nguyễn Tiến Trung, Đinh Đức Tài, Phạm Hùng Hiệp, Lương Đình Hải, Phạm Thị Oanh (2024). Nghiên cứu khoa học giáo dục Việt Nam trên cơ sở dữ liệu Scopus trong 10 năm thực hiện Nghị quyết 29-NQ/TW: Xu hướng và hợp tác quốc tế. *Tạp chí Giáo dục*, 24(1), 1-6. <https://tcgd.tapchigiaoduc.edu.vn/index.php/tapchi/article/view/1176>
- Nguyễn Viết Hưng, Trịnh Huy Hoàng (2022). Áp dụng mô hình học sâu nhận dạng mức độ hài lòng của người học. *Tạp chí Khoa học, Trường Đại học Sư phạm Thành phố Hồ Chí Minh*, 19(12), 2053-2063. [https://doi.org/10.54607/hcmue.js.19.12.3630\(2022\)](https://doi.org/10.54607/hcmue.js.19.12.3630(2022))
- Phạm Thị Thanh Thảo, Phạm Hùng Hiệp, Lương Đình Hải (2022). 55 năm Khoa học Giáo dục Việt Nam, phân tích từ dữ liệu Scopus. NXB Đại học Quốc gia Hà Nội.
- Pregoner, J. D. (2024). Research approaches in education: A comparison of quantitative, qualitative and mixed methods. *IMCC Journal of Science*, 4(2), 12-17. <https://doi.org/10.2139/ssrn.5111007>
- Putra, A. S., & Lafayette, W. (2023). A Review of Quantitative Research Methods in Journal of Engineering Education from 2012 to 2022 By The Numbers : A Review of Quantitative Research Methods in Journal of Engineering Education From 2012 to 2022. *2023 Annual Conference & Exposition*, 1-15. <https://doi.org/10.18260/1-2--43097>
- Tan, C. Y., & Gao, L. (2025). Evaluating methodological quality of meta-analyses: A case study of meta-analyses on associations between parental involvement and students' learning outcomes. *Educational Research Review*, 47(4), 1-22. <https://doi.org/10.1016/j.edurev.2025.100678>
- Tăng Đình Ngọc Thảo, Nguyễn Văn Tròn, Võ Tấn Phát, Nguyễn Dương Thanh (2021). Các yếu tố ảnh hưởng đến sự hài lòng của sinh viên Khoa Khoa học Xã hội và Nhân văn đối với chất lượng dịch vụ đào tạo Trường Đại học Cần Thơ. *Tạp chí Khoa học, Trường Đại học Sư phạm Hà Nội*, 66(1), 120-135. <https://doi.org/10.18173/2354-1075.2021-0012>
- Thuyet N. M. (2024). General Education Curriculum 2018 and issues requiring further research attention. *Vinh University Journal of Science*, 53(Special), 5-16. <https://doi.org/10.56824/vujs.2024.htkhgd0m1>
- Trần Trung (2025). Năng lực phân tích dữ liệu trong nghiên cứu. *Tạp chí Giáo dục*, 25(3), 1-6. <https://tcgd.tapchigiaoduc.edu.vn/index.php/tapchi/article/view/2933>
- Tran, H.-H., Le Van Nguyen, D., Pham, D. T. B., & Tran, T. N. (2024). Exploring educators' perception regarding the field of educational research between 2025 and 2030. *Multidisciplinary Reviews*, 7(7), 1-5. <https://doi.org/10.31893/multirev.2024146>
- Zhang, L. (2025). Estimating school and teacher effects on students' academic performance using multilevel models with three or more levels: a literature review. *Discover Education*, 4(1), 1-20. <https://doi.org/10.1007/s44217-024-00392-4>