

Vận dụng LPA phân nhóm học sinh dựa trên kết quả học tập và tự đánh giá năng lực STEM

Applying LPA to classify students based on learning outcomes and self-efficacy in STEM

Tạ Thanh Trung^{1*}, Lê Châu Đạt²

¹Trường Đại học Sư phạm Thành phố Hồ Chí Minh, Thành phố Hồ Chí Minh, Việt Nam

²Trường THPT Gia Định, Thành phố Hồ Chí Minh, Việt Nam

*Tác giả liên hệ, Email: trungtphysics@gmail.com

THÔNG TIN

DOI:10.46223/HCMCOUJS.
soci.vi.19.2.3070.2024

Ngày nhận: 10/11/2023

Ngày nhận lại: 13/05/2024

Duyệt đăng: 20/05/2024

Từ khóa:

hứng thú nghề nghiệp; lý thuyết nhận thức xã hội nghề nghiệp; năng lực STEM; phân tích nhóm ẩn; sự tự tin về năng lực STEM

Keywords:

career interest; Social Cognitive Career Theory (SCCT); STEM competencies; Latent Profile Analysis (LPA); STEM self-efficacy

TÓM TẮT

Trong bối cảnh các ngành nghề STEM ngày càng giữ vị trí quan trọng, việc định hướng nghề nghiệp học sinh, phân luồng đào tạo nguồn nhân lực cho những lĩnh vực này ngày càng nhận được quan tâm và triển khai sớm tại các trường phổ thông. Mục đích chính của bài báo này là cung cấp hướng dẫn về phân tích nhóm ẩn (LPA - Latent Profile Analysis), một phương pháp định lượng để xác định số lượng nhóm học sinh tiềm ẩn được phân chia dựa trên kết quả học tập STEM và mức độ tự tin về năng lực STEM trên đối tượng học sinh Trung Học Phổ Thông (THPT) trên địa bàn Thành phố Hồ Chí Minh. Kết quả phân tích trên mẫu nghiên cứu 1,074 học sinh trung học phổ thông cho thấy có thể phân thành ba nhóm đối tượng học sinh tiềm ẩn khi học tập STEM (hai nhóm đa số và một nhóm thiểu số) có sự phân hóa rõ rệt về kết quả học tập STEM và mức độ tự tin về năng lực STEM. Bên cạnh đó, kết hợp với các phép phân tích hồi quy logistic, kiểm định T-test cho hai mẫu độc lập, đánh giá chỉ số Cohen's d, nghiên cứu cũng chỉ ra sự ảnh hưởng của giới tính, tần suất tham gia trải nghiệm STEM lên sự chia nhóm, cũng như mức độ ảnh hưởng của nhóm đến định hướng nghề nghiệp của học sinh trung học phổ thông.

ABSTRACT

In the context of the increasingly vital STEM fields, guiding students' careers and implementing workforce training streams in these areas are gaining attention and being implemented early in secondary schools. The main purpose of this article is to provide guidance on Latent Profile Analysis (LPA), a quantitative method to determine the number of latent student profiles based on STEM learning outcomes and self-efficacy in STEM among high school students in Ho Chi Minh City. The analysis results on a sample of 1,074 high school students indicate the potential division into three latent student groups (two majority and one minority group) with significant differentiation in STEM learning outcomes and self-efficacy. Additionally, combined with logistic regression analysis, a T-test for independent samples, and Cohen's d index evaluation, the study also highlights the influence of gender and frequency of STEM experiences on different profiles, as well as the impact of these profiles on the career orientation of high school students.

1. Giới thiệu

Phân tích nhóm tiềm ẩn (LPA - Latent Profile Analysis) là phương pháp thống kê đang nhận được sự quan tâm của các nhà nghiên cứu trong những năm gần đây (Spurk, Hirschi, Wang, Valero, & Kauffeld, 2020). Phương pháp này được nghiên cứu và phát triển với mục đích phát hiện ra những hiện ra các *nhóm ẩn* (latent class) có chung nhiều đặc điểm trong một tập hợp đối tượng đa dạng về đặc điểm (Ferguson, Moore, & Hull, 2020). LPA được liệt vào nhóm phương pháp tiếp cận lấy con người làm trung tâm (person - centered approaches), bởi mô hình này hướng tới việc phân nhóm những đối tượng khảo sát dựa vào các đặc điểm cụ thể, qua đó giúp tìm ra những nhóm người có nhiều đặc điểm tương tự nhau (Ferguson & ctg., 2020). Hiện nay, nghiên cứu vận dụng LPA trải rộng trên nhiều lĩnh vực khác nhau, từ xã hội học (Spurk & ctg., 2020), giáo dục học (Ferguson & ctg., 2020; Ning & Downing, 2015; Rohatgi & Scherer, 2020), y học (Reyes & ctg., 2022), đến tâm lý học (Achterhof, Huntjens, Meewisse, & Kiers, 2019). Trong giáo dục, phương pháp này thường được vận dụng nghiên cứu phân loại đối tượng người học dựa trên các tiêu chí như: sự hứng thú khi học các môn học khoa học và điểm số ở các môn khoa học của học sinh trung học (Ferguson & ctg., 2020), động lực của sinh viên khi học toán và khoa học (Rangel, Vaval, & Bowers, 2020), thái độ của học sinh trung học khi học toán và khoa học (Berger, Mackenzie, & Holmes, 2020), mức độ kì vọng, mức độ tự tin về năng lực bản thân và kết quả khi học tập STEM của sinh viên (Rice, Lopez, & Richardson, 2013). Ở Việt Nam, LPA cũng đã được ứng dụng trong nghiên cứu xây dựng hệ thống bài trắc nghiệm thích nghi thông qua việc phân loại người học theo điểm số (Nguyen, 2018). Mặc dù phương pháp LPA cho thấy nhiều tiềm năng ứng dụng trong phân tích thống kê giáo dục, số lượng nghiên cứu sử dụng LPA trong giáo dục ở Việt Nam vẫn còn rất hạn chế.

Trong bối cảnh các ngành nghề Khoa học, Công nghệ, Kỹ thuật, Toán học (STEM) ngày càng giữ vị trí quan trọng trong sự phát triển kinh tế các nước, việc đào tạo nguồn nhân lực cho các ngành nghề STEM trở thành một nhiệm vụ cấp thiết của các nền giáo dục. Tại Việt Nam, chương trình Giáo dục Phổ thông 2018 ra đời đã quy định giai đoạn từ lớp 10 đến lớp 12 là giai đoạn giáo dục định hướng nghề nghiệp, theo đó kể từ lớp 10, học sinh sẽ lựa chọn học theo các nhóm môn tự chọn phù hợp với định hướng nghề sau này của mình (MOET, 2018). Nhằm góp phần thực hiện mục tiêu giáo dục hướng nghiệp, Bộ Giáo dục và Đào tạo cũng đã ban hành công văn 3089 năm 2020 về việc triển khai thực hiện giáo dục STEM ở trường trung học (MOET, 2020). Việc sử dụng mô hình nghiên cứu lấy người học làm trung tâm như LPA giúp các nhà giáo dục phân tích được một cách chi tiết những đặc điểm đa dạng của một số học sinh, từ đó có thể đưa ra những biện pháp giáo dục phù hợp cho các nhóm học sinh khác nhau (Berger & ctg., 2020). Việc áp dụng LPA trong giáo dục STEM ở Việt Nam giúp phân loại học sinh theo đặc điểm (năng lực, sở thích, mục tiêu, nhu cầu, ...) từ đó xây dựng biện pháp giáo dục tối ưu hóa hiệu suất học tập STEM và hỗ trợ định hướng nghề nghiệp STEM cho học sinh. Điều này sẽ góp phần đem lại hiệu quả trong giáo dục STEM và giáo dục hướng nghiệp có sự phân hóa học sinh theo định hướng của chương trình giáo dục phổ thông 2018.

Những cơ sở nghiên cứu nêu trên cho thấy việc phân loại học sinh dựa trên điểm số các môn STEM và sự tự đánh giá năng lực STEM, đánh giá sự khác biệt trong định hướng nghề nghiệp STEM của học sinh là một nhiệm vụ nghiên cứu cần thiết trong bối cảnh của Việt Nam hiện nay. Do đó, nghiên cứu này được thực hiện với các mục tiêu như sau: (1) *Vận dụng phương pháp LPA để xác định số lượng nhóm học sinh tiềm ẩn được phân chia dựa trên kết quả học tập STEM và mức độ tự tin về năng lực STEM và phân tích những đặc trưng của từng nhóm;* (2) *Phân tích tác động của các yếu tố giới tính, lớp, mức độ thường xuyên tham gia hoạt động STEM đến khả năng học sinh rơi vào một trong các nhóm tiềm ẩn;* (3) *Phân tích sự khác biệt về mức độ hứng thú nghề nghiệp STEM giữa các nhóm học sinh tiềm ẩn.* Kết quả của nghiên cứu

này sẽ đóng góp vào việc mở rộng phạm vi nghiên cứu về hướng nghiệp STEM cho học sinh tại Việt Nam. Đồng thời, nghiên cứu cung cấp một nền tảng cho việc áp dụng mô hình nghiên cứu LPA trong lĩnh vực giáo dục, giúp các nhà nghiên cứu có thể áp dụng mô hình này trong các nghiên cứu định lượng khác tại Việt Nam trong tương lai.

2. Cơ sở lý thuyết

2.1. Giới thiệu phân tích LPA

2.1.1. Lý thuyết phân tích LPA

LPA được xây dựng dựa trên giả định rằng mỗi cá nhân trong tập hợp đều có thể được phân loại vào các nhóm khác nhau (nhóm tiềm ẩn) ứng với mức độ xác suất khác nhau và mỗi nhóm tiềm ẩn đều có những đặc điểm hoàn toàn riêng biệt (Spurk & ctg., 2020). Ban đầu, phương pháp này được nghiên cứu để giải quyết các biến đầu vào dưới dạng nhị thức (dichotomous variable). Sau đó LPA đã được mở rộng cho các loại biến rời rạc (discrete variable) cũng như cho biến liên tục (continuous variable) (Tein, Coxe, & Cham, 2013).

Về mặt kỹ thuật, LPA chú trọng đến sự không đồng nhất của đối tượng nghiên cứu. Sự không đồng nhất có thể phân biệt thành hai loại là sự không đồng nhất khả kiến và bất khả kiến. Sự không đồng nhất khả kiến là trường hợp các nhóm đối tượng được phân biệt qua một biến cụ thể, có thể thu thập trực tiếp (giới tính, độ tuổi, ...). Trong trường hợp này, mô hình phân tích theo kiểu tập trung vào biến (variable - centered analysis) như kiểm định T-test, phân tích phương sai (ANOVA), mô hình phân tích cấu trúc tuyến tính (SEM) có thể được sử dụng để so sánh giữa các nhóm (Spurk & ctg., 2020). Các mô hình này hướng tới phân tích mối liên hệ giữa các biến, chủ yếu khai thác đặc điểm chung của một nhóm đối tượng (Ferguson & ctg., 2020). Tuy nhiên, các mô hình phân tích trên sẽ gặp trở ngại khi có hiện tượng không đồng nhất bất khả kiến, khi có nhiều biến cùng một lúc gây ra sự không đồng nhất, tạo thành những nhóm ẩn, những đặc điểm của các nhóm ẩn này không thể biểu hiện ra khi thống kê trên cả tập hợp (Rohatgi & Scherer, 2020). LPA sẽ phát huy hiệu quả ở các trường hợp này. Kết quả từ phân tích LPA sẽ là các biến phân loại ẩn với giá trị là xác suất để một cá nhân thuộc vào một trong những nhóm đã phân chia (Spurk & ctg., 2020). Thông qua việc vận dụng phương pháp phân tích LPA, nhà nghiên cứu có thể khai thác các thông tin về phân tích nhóm ẩn như: số lượng nhóm ẩn tối ưu; số lượng thành viên của mỗi nhóm; đặc trưng của mỗi nhóm (thông qua các chỉ số như trung bình, độ lệch chuẩn, ...). Ngoài ra, khi kết hợp thêm phương pháp ANOVA (Ning & Downing, 2015; Pastor, Barron, Miller, & Davis, 2007) và hồi quy logistic (Ferguson & ctg., 2020; Morin & Marsh, 2015; Rohatgi & Scherer, 2020), kết quả thu được còn giúp chỉ ra sự tác động của các yếu tố (covariates) tới kết quả phân nhóm, dựa trên tính khả năng một cá nhân thuộc về một trong các nhóm đã phân chia.

2.1.2. Tiêu chí lựa chọn số lượng nhóm ẩn tối ưu trong LPA

Mục tiêu chính của nghiên cứu là phân loại các nhóm học sinh có đặc điểm giống nhau về các năng lực thành phần trong năng lực tư duy thiết kế. Trước tiên phép phân tích LPA sẽ được áp dụng cho các mô hình có số lượng nhóm khác nhau, sau đó tiến hành đánh giá các chỉ số thu được để chọn ra mô hình có số nhóm tối ưu nhất. Các chỉ số thường được dùng nhất để chọn số nhóm tối ưu là *Akaike Information Criterion (AIC)*, *Bayesian Information Criterion (BIC)*, và *Sample-size Adjusted BIC (SABIC)* (Rohatgi & Scherer, 2020). AIC được xác định dựa trên log - likelihood và số lượng tham số (parameters), không xét đến kích thước mẫu. BIC và SABIC được xác định dựa trên log - likelihood số lượng tham số có xét tới kích thước mẫu. Giá trị của các chỉ số này càng nhỏ chứng tỏ rằng phép phân nhóm càng khớp với số liệu (Marsh, Lüdtke, Trautwein, & Morin, 2009; Morin & Marsh, 2015). Tuy nhiên, các chỉ số AIC, BIC và

SABIC không phải luôn là tiêu chí hàng đầu trong việc đánh giá, điều này tùy thuộc vào độ phức tạp của mô hình nghiên cứu (Bechter, Dimmock, Howard, Whipp, & Jackson, 2018). Các chỉ số trên có thể giảm liên tục khi tăng số nhóm lên, dẫn tới kết quả là mô hình được chọn có số nhóm quá lớn, đặc biệt là nếu các biến đo lường không tuân theo phân bố chuẩn (Choi, Nylund-Gibson, Israel, & Mendez, 2019; Spurk & ctg., 2020). Một số nghiên cứu khác cho rằng mô hình có số nhóm tối ưu nhất không nhất thiết phải có giá trị AIC, BIC và SABIC cực tiểu, chỉ cần các chỉ số này giảm chậm lại rõ rệt khi tăng số nhóm lên (ví dụ: mô hình 04 nhóm có thể xem là tối ưu nếu chỉ số BIC giảm nhanh từ 01 nhóm đến 04 nhóm, khi tăng hơn 04 nhóm thì BIC giảm chậm lại rõ rệt) (Ferguson & ctg., 2020).

Một tiêu chí khác để đánh giá mô hình đó là so sánh sự khác biệt thống kê giữa 02 mô hình có số nhóm liền kề nhau (giữa mô hình có k nhóm và mô hình có $k - 1$ nhóm), thông qua thực hiện **phép kiểm định LMR (Lo-Mendell-Rubin adjusted test)** và **kiểm định BLRT (Bootstrapped Likelihood Ratio Test)**. Nếu kết quả cho giá trị p lớn hơn 0.05 sẽ cho thấy mô hình k nhóm không khác biệt rõ rệt so với mô hình $k - 1$ nhóm, như vậy cần loại bỏ và giữ lại mô hình có ít nhóm hơn. Phương pháp này chỉ dùng để lọc đi các mô hình có số nhóm quá lớn thay vì chọn mô hình tối ưu nhất. Trong một số trường hợp, các mô hình có số nhóm khá lớn vẫn có giá trị p nhỏ hơn 0.05 nên vẫn chưa loại được bằng phương pháp này (Celeux & Soromenho, 1996; Morin & Marsh, 2015; Spurk & ctg., 2020).

Một chỉ số quan trọng cũng thường được dùng trong đánh giá mô hình đó là chỉ số **entropy**, là chỉ số thể hiện mức độ tách biệt rõ rệt giữa các nhóm trong một mô hình (Celeux & Soromenho, 1996). Entropy có giá trị nằm trong khoảng từ 0 đến 1, chỉ số này càng lớn càng thể hiện mức độ tách biệt rõ rệt giữa các nhóm, càng có tính tối ưu (Celeux & Soromenho, 1996; Ning & Downing, 2015; Rohatgi & Scherer, 2020). Giá trị entropy lớn hơn 0.8 được xem là phù hợp (Spurk & ctg., 2020). Một số nhà nghiên cứu cho rằng chỉ số entropy đạt cực đại có thể xem là một tiêu chí để lựa chọn nhóm tối ưu (Rohatgi & Scherer, 2020), tuy nhiên, điều này cần sự cân nhắc kỹ bởi entropy có thể tăng như một hàm theo số nhóm (Cleveland, Collins, Lanza, Greenberg, & Feinberg, 2010) và là một chỉ số có độ tin cậy kém hơn các chỉ số thống kê đã kể trên (Spurk & ctg., 2020).

Một tiêu chí nữa, ít phổ biến hơn được sử dụng để loại đi các mô hình có số nhóm lớn là dựa vào **tỷ lệ của nhóm nhỏ nhất trong mô hình**. Theo đó, các nhóm chiếm ít hơn 5% kích thước của mẫu được xem là không có ý nghĩa đáng kể trong phân tích kết quả, cần loại đi các mô hình có nhóm nhỏ như vậy (Ning & Downing, 2015).

Các chỉ số và tiêu chí đánh giá được kể trên chỉ mang tính tương đối, không phải lúc nào cũng cho kết quả phù hợp. Nhiều nhà nghiên cứu đi đến quan điểm cho rằng việc đánh giá không nên chỉ dựa vào các chỉ số thống kê đã nêu mà còn phải căn cứ vào cơ sở lý thuyết (Ning & Downing, 2015; Spurk & ctg., 2020), các nghiên cứu trước đó, các đặc trưng của mỗi nhóm được phân tích, và hàm lượng thông tin thu được từ kết quả (Ferguson & ctg., 2020; Marsh & ctg., 2009). Dựa trên cơ sở các nghiên cứu đã nêu, chúng tôi đề xuất quy trình lựa chọn số nhóm tối ưu nhất như sau: trước tiên sẽ đánh giá kết quả kiểm định BLRT và LMR, loại đi các mô hình có giá trị p lớn hơn 0.05. Bước tiếp theo sẽ loại đi các mô hình có chỉ số entropy không đạt chuẩn (entropy < 0.8), những nhóm chiếm tỷ lệ quá nhỏ so với mẫu (chiếm chưa đến 5% mẫu). Sau đó sẽ đánh giá các chỉ số AIC, BIC và SABIC để lựa chọn mô hình có một trong các chỉ số này đạt cực tiểu, nếu không tìm được cực tiểu thì sẽ lựa chọn mô hình mà chỉ số BIC hoặc SABIC có sự giảm chậm lại rõ rệt khi tiếp tục tăng số nhóm. Nếu vẫn không tìm được mô hình phù hợp dựa trên tiêu chí này thì sẽ chọn mô hình có entropy lớn nhất trong số các mô hình còn lại. Bảng 1 tóm tắt các tiêu chí được sử dụng để lựa chọn số lượng nhóm tối ưu khi phân tích LPA.

Bảng 1

Các tiêu chí sử dụng trong lựa chọn số nhóm tối ưu

| Tên tiêu chí | Cách đánh giá |
|--------------------------|--|
| LMR | Nếu giá trị p không có ý nghĩa thống kê ($p > 0.05$) đối với mô hình có k nhóm, điều này cho thấy mô hình này không cải thiện đáng kể so với mô hình k - 1 nhóm, do đó nên giữ lại mô hình k - 1 nhóm. |
| BLRT | Nếu giá trị p không có ý nghĩa thống kê ($p > 0.05$) đối với mô hình có k nhóm, điều này cho thấy mô hình này không vượt trội hơn mô hình k - 1 nhóm, do đó nên giữ lại mô hình k - 1 nhóm. |
| Kích thước nhóm nhỏ nhất | Nhóm nhỏ nhất không được ít hơn 5% mẫu |
| AIC | Chọn mô hình có chỉ số AIC cực tiểu hoặc có sự giảm chậm rõ rệt khi tăng số nhóm |
| BIC | Chọn mô hình có chỉ số BIC cực tiểu hoặc có sự giảm chậm rõ rệt khi tăng số nhóm |
| SABIC | Chọn mô hình có chỉ số SABIC cực tiểu hoặc có sự giảm chậm rõ rệt khi tăng số nhóm |
| Entropy | Chọn mô hình có chỉ số entropy cực đại |

Nguồn: Kết quả xử lý từ dữ liệu điều tra

Theo khuyến nghị từ nhiều nghiên cứu (Ferguson & ctg., 2020; Marsh & ctg., 2009; Ning & Downing, 2015; Spurk & ctg., 2020), sau khi chọn được mô hình có số nhóm tối ưu, vẫn cần so sánh với cơ sở lý luận, các nghiên cứu trước đó cũng như kết quả rút ra được để đánh giá xem mô hình trên có phù hợp hay không, nếu vẫn chưa phù hợp thì cần tiến hành chọn lại mô hình một lần nữa.

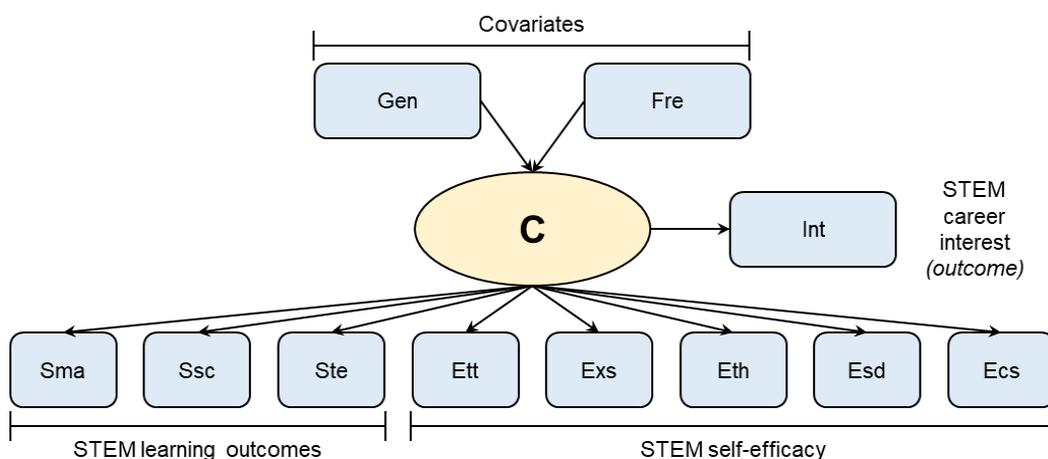
2.2. Mô hình lý thuyết về việc phân nhóm học sinh dựa trên thuyết nhận thức xã hội nghề nghiệp trong Giáo dục STEM

Các nghiên cứu trên thế giới cho thấy sự tự tin (self-efficacy) của học sinh trong một lĩnh vực có thể ảnh hưởng đến hứng thú nghề nghiệp (career interest) trong lĩnh vực đó (Fouad & Smith, 1996; Lent, Brown, & Hackett, 1994; Nugent & ctg., 2015). Sự hứng thú nghề nghiệp STEM được định nghĩa là sự hứng thú của một cá nhân trong việc lựa chọn các nghề nghiệp liên quan đến STEM trong tương lai (Luo, So, Wan, & Li, 2021). Sự tự tin về năng lực được hiểu là “sự đánh giá cá nhân về khả năng của bản thân trong việc tổ chức và thực hiện một chuỗi hoạt động cần thiết nhằm đạt được một mục tiêu mong muốn nào đó” (Bandura, 1986, tr. 391). Những niềm tin này được xem là yếu tố trung tâm và giữ vai trò quan trọng cấu thành sự tự chủ cá nhân (Lent & ctg., 1994). Sự tự tin về năng lực STEM (STEM self-efficacy) được hiểu là sự tự tin về khả năng nhận dạng, ứng dụng, tích hợp các kiến thức đã học được từ các môn Khoa học, Công nghệ, Kỹ thuật và Toán học để hiểu và giải quyết các vấn đề trong thực tiễn (Zollman, 2012). Kết quả nghiên cứu tổng quan của Nguyen, Tran, Nguyen, và Ta (2022) chỉ ra rằng năng lực STEM của học sinh trung học bao gồm 05 năng lực thành tố: (1) Thu thập thông tin; (2) Xử lý và sử dụng thông tin; (3) Thực hiện giải pháp; (4) An toàn kỹ thuật; (5) Chia sẻ cộng đồng. Đây là cơ sở quan trọng cho phép thực hiện phép phân tích LPA, cho phép đánh giá sự tác động của mỗi năng lực thành tố đến hứng thú nghề nghiệp STEM (STEM career interest) của học sinh. Sự tự tin về năng lực được hình thành thông qua các trải nghiệm thực tế của cá nhân bên cạnh những tác động từ

những chuẩn mực xã hội với các phản ứng về tâm sinh lý của họ (Ching & ctg., 2019; Falco & Summers, 2019). Người học khi có cơ hội trải nghiệm và tiếp xúc nhiều hơn với việc tự đánh giá trong quá trình học thì sẽ nâng cao khả năng tự đánh giá tốt hơn (Gehring, 2017). Vì vậy, các hoạt động trải nghiệm STEM được đánh giá là đóng vai trò tích cực trong việc cải thiện quan niệm về STEM và sự tự tin về năng lực STEM của học sinh (Ching & ctg., 2019).

Bên cạnh đó, theo quan điểm đánh giá năng lực hiện nay, kết quả học tập (learning outcomes) không còn được sử dụng với mục tiêu chính là xếp hạng học sinh mà thay vào đó các nhà giáo dục kì vọng đây là một kênh thông tin phản hồi cho người học để giúp họ cải thiện và phát triển các khả năng của mình. Nghiên cứu của van Aalderen-Smeets, van der Molen, và Xenidou-Dervou (2019) trên đối tượng học sinh từ 14 đến 18 tuổi cho thấy điểm số học tập các môn STEM (STEM learning outcomes) (Toán, Vật lí, Hóa học, Sinh học, Khoa học tự nhiên, Công nghệ, Tin học) cao sẽ có tác động tích cực đến sự tự tin về năng lực STEM, đồng thời thúc đẩy học sinh lựa chọn học tập và nghề nghiệp STEM (Glynn, Taasobshirazi, & Brickman, 2009; van Aalderen-Smeets & ctg., 2019). Tuy nhiên, trong một nghiên cứu trên 65 sinh viên đại học, Kruger và Dunning (1999) cũng từng chỉ ra luôn có tồn tại sự chênh lệch giữa kết quả được đánh giá và tự đánh giá ở người học. Cụ thể, những học sinh có kết quả học tập ở mức thấp thường có xu hướng tự đánh giá năng lực mình cao hơn so với thực tế. Ngoài ra, phân tích tác động của giới tính đến sự tự tin về năng lực của học sinh cũng là vấn đề nghiên cứu được quan tâm trong việc triển khai giáo dục STEM (Falco & Summers, 2019; Stewart & ctg., 2020; van Tuijl & van der Molen, 2016). Các nghiên cứu cho thấy sự tự tin về năng lực STEM của nữ thường thấp hơn nam, kể cả khi điểm môn học STEM cao hơn (Stewart & ctg., 2020). Các quan điểm rập khuôn về việc nam giới nổi bật trong môn Toán và Khoa học có thể ảnh hưởng đến khả năng tự hiệu quả và khát vọng nghề nghiệp của nữ giới (van Tuijl & van der Molen, 2016).

Trong nghiên cứu này, chúng tôi sẽ sử dụng phương pháp LPA để phân chia học sinh thành các nhóm ẩn dựa trên 08 biến khảo sát định lượng, trong đó có 03 biến là điểm số trung bình của học sinh ở các môn hoặc nhóm môn (Sma: điểm Toán; Ssc: điểm Khoa học; Ste: điểm Tin học và Công nghệ), 05 biến do HS tự đánh giá 05 năng lực thành tố của năng lực STEM (Ett: thu thập thông tin; Exs: xử lý và sử dụng thông tin; Eth: thực hiện giải pháp; Esd: an toàn kĩ thuật; Ecs: chia sẻ cộng đồng) (Hình 1). Các biến bao gồm giới tính (Gen), mức độ thường xuyên tham gia hoạt động STEM (Fre) được sử dụng để đánh giá tác động của các yếu tố bên ngoài đến việc phân nhóm. Riêng biến hứng thú nghề nghiệp (Int) được sử dụng để đánh giá mức độ ảnh hưởng đầu ra của việc phân nhóm.



Hình 1. Mối liên hệ giữa các biến được sử dụng trong mô hình lý thuyết

3. Phương pháp nghiên cứu

3.1. Mẫu và cách thu mẫu

Nghiên cứu áp dụng phương pháp chọn mẫu ngẫu nhiên tích tụ để thu thập thông tin của các học sinh tại 09 trường THPT tại Thành phố Hồ Chí Minh. Việc thực hiện khảo sát đã được sự đồng ý của ban giám hiệu nhà trường, giáo viên bộ môn nhằm đảm bảo về mặt đạo đức nghiên cứu, và các học sinh thực hiện khảo sát một cách tự nguyện.

Tổng dữ liệu thu thập gồm 1,084 học sinh, sau khi lọc dữ liệu dựa trên tiêu chí các học sinh đánh chung một mức cho tất cả các câu hỏi, dữ liệu cuối cùng tiến hành phân tích gồm 1,074 học sinh. Các đặc điểm khác của mẫu được mô tả qua Bảng 2.

Bảng 2

Đặc điểm thống kê của mẫu

| Đặc điểm | | Tần số | Tỷ lệ (%) |
|---|---|--------|-----------|
| Giới tính | Nam | 558 | 52.0 |
| | Nữ | 516 | 48.0 |
| Loại hình trường | Trường công lập | 884 | 82.3 |
| | Trường ngoài công lập | 190 | 17.7 |
| Đặc điểm | | Tần số | Tỷ lệ (%) |
| Lớp | 10 | 600 | 55.9 |
| | 11 | 321 | 29.9 |
| | 12 | 153 | 14.2 |
| Mức độ thường xuyên tham gia hoạt động STEM | Chưa tham gia | 616 | 57.4 |
| | Thỉnh thoảng (01 - 03 lần/học kì) | 303 | 28.2 |
| | Thường xuyên (trên 03 lần/học kì) | 155 | 14.4 |
| Hình thức tiếp cận giáo dục STEM | Trong một tiết học STEM chính khóa | 296 | 27.6 |
| | Dự án học tập | 144 | 13.4 |
| | Trung tâm ngoài trường | 33 | 3.1 |
| | Khóa học trực tuyến | 52 | 4.8 |
| | Trải nghiệm không gian STEM tại trường | 253 | 23.6 |
| | Sân chơi, hội chợ khoa học | 169 | 15.7 |
| | Trải nghiệm thực tế | 31 | 2.9 |
| | Sinh hoạt câu lạc bộ | 80 | 7.5 |
| | Tham gia các cuộc thi khoa học kĩ thuật | 77 | 7.2 |

Ghi chú: * Mỗi học sinh có thể tiếp xúc nhiều hơn 01 hình thức
 Nguồn: Kết quả xử lý từ dữ liệu điều tra

3.2. Thu thập dữ liệu

Điểm trung bình môn Toán, Khoa học (lấy trung bình theo 03 môn: Lý, Hóa, Sinh), Kỹ thuật và Công nghệ (lấy trung bình theo 02 môn: Tin học và Công nghệ) do học sinh ghi nhận dựa vào điểm trung bình học kì vừa qua. Các biến về sự tự tin năng lực STEM được chia theo 05 năng lực thành phần, mỗi năng lực thành phần có 03 câu hỏi khảo sát, trong mỗi câu hỏi khảo sát học sinh sẽ tự đánh giá theo thang Likert 7, từ “Rất không đồng ý” đến “Rất đồng ý”. Tương tự, thang đo hứng thú nghề nghiệp cũng dựa trên 03 câu hỏi khảo sát, mỗi câu theo thang Likert 7.

3.3. Xử lý dữ liệu

Bước 1. Xử lý dữ liệu đầu vào: Kiểm tra và đảm bảo một số đặc tính thống kê như độ phù hợp với phân phối chuẩn thông qua các hệ số skewness và kurtosis, đánh giá độ tương quan giữa các biến. Điều kiện phân phối chuẩn được sử dụng là skewness từ -2 đến 2, kurtosis từ -7 đến 7 (Byrne, 2016). Hệ số tương quan (ρ) có độ lớn dưới 0.3 là yếu, từ 0.3 đến 0.6 là vừa phải, từ 0.6 đến 0.8 là khá cao và trên 0.8 là rất cao (Akoglu, 2018). Nếu có một số dữ liệu trên một biến quan sát (item) bị thiếu, có thể sử dụng phương pháp ước lượng hợp lý tối đa (FIML - Full Maximum Likelihood Estimation) để ước lượng dữ liệu còn thiếu (Ferguson & ctg., 2020; Spurk & ctg., 2020). Việc đảm bảo dữ liệu không bị mắc các lỗi như lỗi cú pháp, tham số vượt quá giới hạn, thông kê không đúng với lý thuyết cũng là điều rất quan trọng (Ferguson & ctg., 2020; Spurk & ctg., 2020). Để thực hiện phân tích LPA, giá trị mỗi biến sẽ được quy về thang đo có trung bình bằng 0 và độ lệch chuẩn bằng 1.

Bước 2. Chia nhóm và đánh giá mô hình: Dữ liệu sau đó được dựa vào chương trình phân tích LPA, bao gồm việc khai báo file dữ liệu vào, các biến cần sử dụng, phép kiểm định cần thực hiện (LMR và BLRT). Việc đánh giá các mô hình để chọn ra được số nhóm tối ưu nhất, dựa trên các phép kiểm định đã nêu. Ngoài ra, cần xem xét sự phù hợp của kết quả nghiên cứu với cơ sở lý thuyết.

Bước 3. Phân tích đặc trưng nhóm: Phân tích các tính chất đặc trưng của từng nhóm, dựa trên số lượng, tỷ lệ, giá trị trung bình, độ lệch chuẩn của các biến quan sát trên mỗi nhóm.

Bước 4. Phân tích các tác nhân: Thực hiện phép phân tích hồi quy logistic (logistic regression) để đánh giá tác động của các nhân tố khác đến sự phân nhóm. Kết quả phép phân tích logistic sẽ so sánh giữa hai nhóm A và B đã phân tích, từ đó cho biết các nhân tố trên sẽ có tác động làm tăng khả năng rơi vào nhóm A (hoặc nhóm B) hay không.

Bước 5. Phân tích ảnh hưởng đầu ra: Đánh giá mức độ ảnh hưởng của các nhóm đã phân chia đến định hướng nghề nghiệp STEM của học sinh được thực hiện thông qua phép kiểm định T-test cho hai mẫu độc lập và đánh giá chỉ số Cohen's d.

Nghiên cứu này sử dụng phần mềm SPSS ở Bước 1 và Bước 5, các bước còn lại được thực hiện trên phần mềm Mplus.

4. Kết quả và thảo luận

4.1. Thống kê mô tả và tương quan giữa các biến

Bảng 3 thể hiện thống kê mô tả trung bình, độ lệch chuẩn, hệ số tương quan, phân phối mẫu qua các giá trị độ lệch (skewness) và độ nhọn (kurtosis) của các biến trong mô hình nghiên cứu. Dựa trên điều kiện skewness dao động từ -2 đến 2 và kurtosis dao động từ -7 đến 7, dữ liệu ở các biến trên đều thỏa tiêu chí của phân phối chuẩn cho cỡ mẫu lớn.

Bảng 3

Thống kê mô tả các biến đánh giá (N = 1,074)

| Thang đo | Minimum | Maximum | Mean | Std. Deviation | Skewness | Kurtosis |
|----------|---------|---------|--------|----------------|----------|----------|
| Sma | 1 | 10 | 7.790 | 1.519 | -1.027 | 1.260 |
| Ssc | 2 | 10 | 8.079 | 1.283 | -1.359 | 2.890 |
| Ste | 2 | 10 | 8.561 | 1.279 | -1.604 | 3.927 |
| Ett | 1 | 7 | 4.971 | 1.045 | -0.776 | 1.652 |
| Exs | 1 | 7 | 4.976 | 1.045 | -0.676 | 1.206 |
| Eth | 1 | 7 | 4.441 | 1.163 | -0.320 | 0.540 |
| Eds | 1 | 7 | 4.735 | 1.153 | -0.381 | 0.593 |
| Esd | 1 | 7 | 4.544 | 1.173 | -0.322 | 0.640 |
| Int | 1 | 7 | 4.462 | 1.256 | -0.341 | 0.386 |
| Gen | 0 | 1 | 0.480 | 0.500 | - | - |
| Gra | 10 | 12 | 10.580 | 0.727 | - | - |
| Fre | 1 | 3 | 1.570 | 0.731 | - | - |

Nguồn: Kết quả xử lý từ dữ liệu điều tra

Bảng 4

Thống kê hệ số tương quan các biến

| | Sma | Ssc | Sts | Ett | Exs | Eth | Esd | Ecs |
|-----|---------|---------|---------|---------|---------|---------|---------|---------|
| Ssc | 0.819** | | | | | | | |
| Sts | 0.666** | 0.778** | | | | | | |
| Ett | 0.288** | 0.320** | 0.318** | | | | | |
| Exs | 0.244** | 0.289** | 0.308** | 0.773** | | | | |
| Eth | 0.143** | 0.172** | 0.198** | 0.620** | 0.603** | | | |
| Esd | 0.132** | 0.180** | 0.203** | 0.593** | 0.597** | 0.626** | | |
| Ecs | 0.167** | 0.197** | 0.204** | 0.635** | 0.656** | 0.673** | .599** | |
| Int | 0.117** | 0.135** | 0.124** | 0.371** | 0.410** | 0.462** | 0.447** | 0.436** |

** : giá trị $p < 0.001$

Nguồn: Kết quả xử lý từ dữ liệu điều tra

Bảng 4 thống kê hệ số tương quan cho thấy các biến dữ liệu phân thành hai nhóm rõ rệt, bao gồm nhóm các biến về điểm môn học (Sma, Ssc, Sts) và nhóm các biến về sự tự tin năng lực STEM (Ett, Exs, Eth, Esd, Ecs). Giữa các biến trong nhóm đều có tương quan thuận với nhau ở mức từ khá cao đến rất cao (ρ từ 0.593 đến 0.819), trong khi giữa các biến khác nhóm có tương quan thuận đều ở mức thấp (ρ từ 0.132 đến 0.320). Riêng biến Int có hệ số tương quan thấp với các biến điểm số (ρ từ 0.117 đến 0.135) và hệ số tương quan vừa với các biến về tự tin năng lực STEM (ρ từ 0.371 đến 0.462).

4.2. Lựa chọn số nhóm tối ưu

Kết quả phân tích các chỉ số đặc trưng ứng với các mô hình gồm 02 đến 08 nhóm thể hiện trong Bảng 5.

Bảng 5

So sánh giữa các mô hình

| Số nhóm | AIC | BIC | SABIC | Entropy | Giá trị PLMR | Giá trị PBLRT | Số nhóm ít hơn 5% mẫu |
|---------|----------|----------|----------|---------|--------------|---------------|-----------------------|
| 2 | 22099.36 | 22223.84 | 22144.44 | 0.806 | 0.0005 | < 0.0001 | 0 |
| 3 | 21064.23 | 21233.52 | 21125.53 | 0.864 | 0.0016 | < 0.0001 | 0 |
| 4 | 20459.43 | 20673.53 | 20536.95 | 0.859 | 0.3316 | < 0.0001 | 1 |
| 5 | 19919.86 | 20178.78 | 20013.62 | 0.869 | 0.0069 | < 0.0001 | 2 |
| 6 | 19546.40 | 19850.13 | 19656.38 | 0.872 | 0.0498 | < 0.0001 | 2 |
| 7 | 19236.45 | 19584.99 | 19362.65 | 0.860 | 0.2536 | < 0.0001 | 2 |
| 8 | 19011.76 | 19405.11 | 19154.19 | 0.865 | 0.4745 | < 0.0001 | 3 |

Nguồn: Kết quả xử lý từ dữ liệu điều tra

Bảng 5 cho thấy giá trị p của kiểm định LMR đạt giá trị nhỏ hơn 0.05 đối với mô hình có 02, 03 và 05 nhóm, riêng mô hình 04 nhóm có giá trị p vượt qua ngưỡng. Giá trị p của kiểm định BLRT thì đều có giá trị xấp xỉ 0, đảm bảo giá trị p nhỏ hơn 0.05. Các mô hình có từ 04 nhóm trở lên sẽ bị loại do kích thước nhóm nhỏ nhất chưa chiếm đến 5% mẫu. Chỉ số AIC, BIC và SABIC đều giảm khi số nhóm tăng từ 02 đến 05. Về chỉ số entropy, các mô hình 3, 4, 5 nhóm đều có chỉ số entropy lớn hơn 0.9. Như vậy mô hình có ba nhóm là tối ưu nhất (chỉ số AIC, BIC, SABIC đều thấp nhất và chỉ số entropy cao nhất trong số các nhóm không bị loại). Ngoài ra, mô hình gồm ba nhóm cũng phù hợp để thực hiện các phân tích, đánh giá sâu theo mô hình SCCT, vì với số lượng nhóm cao hơn sẽ làm xuất hiện các nhóm rất thiểu số với nhiều đặc điểm cá biệt, gây khó khăn cho việc phân tích đánh giá.

4.3. Đặc trưng của từng nhóm

Sau khi chọn ra mô hình có số nhóm tối ưu, bước tiếp theo là phân tích các tính chất đặc trưng của từng nhóm. Các nhóm lần lượt được mã hóa thành Class 1, Class 2 và Class 3 theo thứ tự tăng dần giá trị trung bình điểm số và điểm tự đánh giá năng lực khi học STEM.

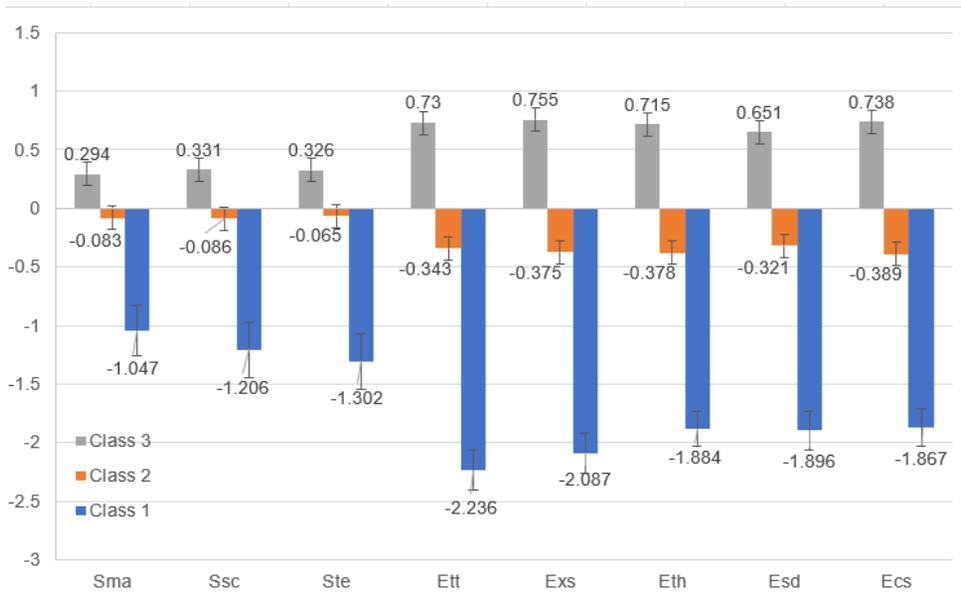
Bảng 6

Thống kê trung bình và sai số chuẩn của các biến trên mỗi nhóm

| Biến | Class 1 (N = 58) (5.4%) | | Class 2 (N = 562) (52.4%) | | Class 3 (N = 454) (42.2%) | |
|------|----------------------------|-------|------------------------------|-------|------------------------------|-------|
| | Means | S.E | Means | S.E | Means | S.E |
| Sma | -1.047 | 0.214 | -0.083 | 0.052 | 0.294 | 0.054 |
| Ssc | -1.206 | 0.235 | -0.086 | 0.052 | 0.331 | 0.049 |
| Sts | -1.302 | 0.237 | -0.065 | 0.051 | 0.326 | 0.043 |
| Ett | -2.236 | 0.172 | -0.343 | 0.052 | 0.730 | 0.054 |
| Exs | -2.087 | 0.173 | -0.375 | 0.055 | 0.755 | 0.051 |
| Eth | -1.884 | 0.149 | -0.378 | 0.044 | 0.715 | 0.072 |
| Esd | -1.896 | 0.163 | -0.321 | 0.046 | 0.651 | 0.063 |
| Ecs | -1.867 | 0.161 | -0.389 | 0.047 | 0.738 | 0.066 |

Nguồn: Kết quả xử lý từ dữ liệu điều tra

Bảng 6 cho thấy Class 1 có số lượng thành viên ít nhất (chiếm 5.4% mẫu), Class 2 đông nhất với 52.2% và Class 3 có 42.2%. Trong đó sai số chuẩn của Class 1 cao hơn hẳn so với Class 2 và Class 3.



Hình 2. Giá trị trung bình các biến đo lường ứng với ba nhóm tiềm ẩn

Đồ thị Hình 2 thể hiện những đặc trưng của mỗi nhóm. Trong đó, Class 1 là nhóm có điểm các môn thấp, mức độ tự tin khi học STEM cũng thấp. Class 2 là nhóm có điểm các môn trung bình, mức độ tự tin ở mức trung bình. Class 3 là nhóm có điểm các môn khá nhất, mức độ tự tin ở mức cao nhất. Sự khác biệt về mặt tự tin của ba nhóm là rất rõ rệt, về điểm số các môn thì Class 2 và Class 3 chênh lệch ít, nhóm 1 có điểm số thấp cách biệt khá xa.

4.4. Phân tích ảnh hưởng của giới tính và mức độ tham gia trải nghiệm STEM lên sự chia nhóm

Bảng 7

Kết quả phân tích hồi quy logistic

| | B | S.E | Odds Ratio | Giá trị p | Diễn giải kết quả |
|--------------------------|--------|-------|------------|-----------|---|
| Class 1 - Class 3 | | | | | |
| Gen | 0.632 | 0.316 | 1.882 | 0.045 | Học sinh nam có tỷ lệ rơi vào Class 1 so với Class 3 cao hơn học sinh nữ. Học sinh ít tham gia STEM có xu hướng rơi vào Class 1 nhiều hơn. |
| Fre | -0.947 | 0.230 | 0.388 | < 0.001 | |
| Class 2 - Class 3 | | | | | |
| Gen | -0.310 | 0.144 | 0.733 | 0.032 | Học sinh nam có tỷ lệ rơi vào Class 2 so với Class 3 ít hơn học sinh nữ. Học sinh ít tham gia STEM có xu hướng rơi vào Class 2 nhiều hơn. |
| Fre | -0.537 | 0.091 | 0.585 | < 0.001 | |
| Class 1 - Class 2 | | | | | |
| Gen | 0.942 | 0.313 | 0.390 | 0.003 | Học sinh nam có xu hướng rơi vào Class 1 so với Class 2 nhiều hơn học sinh nữ. |
| Fre | -0.410 | 0.229 | 0.664 | 0.073 | |

Nguồn: Kết quả xử lý dữ liệu điều tra

Nghiên cứu cho thấy ở hai nhóm đa số (Class 1 và Class 2), nam giới vẫn thể hiện xu hướng có điểm số và tự tin về năng lực STEM cao hơn so với nữ giới. Tuy nhiên, độ chênh lệch về điểm số học tập giữa Class 2 và Class 3 không rõ rệt bằng sự chênh lệch về mức độ tự tin. Tuy nhiên, riêng đối với nhóm thiểu số (Class 1) chỉ chiếm 5.4% mẫu, xu hướng này ngược lại, cho thấy khả năng nam giới thuộc Class 1 cao hơn so với nữ giới. Điều này làm nổi bật giá trị của phương pháp LPA trong việc phát hiện những nhóm ẩn với đặc điểm ngược lại so với xu hướng chung. Cần có thêm những nghiên cứu sâu hơn để làm rõ thêm các đặc điểm của những đối tượng thiểu số này. Kết quả phân tích cũng đã cho thấy học sinh tham gia trải nghiệm STEM nhiều hơn có xu hướng thuộc Class 3 (điểm số và tự tin về năng lực STEM cao) nhiều hơn. Điều này một lần nữa khẳng định vai trò quan trọng của các hoạt động STEM trong việc tăng cường sự tự tin về năng lực STEM, từ đó thúc đẩy học sinh có động lực học tập các môn STEM và đạt được kết quả tốt hơn.

4.5. Kết quả phân tích mức độ ảnh hưởng của nhóm đến định hướng nghề nghiệp

Bảng 8

Mức độ ảnh hưởng của nhóm đến định hướng nghề nghiệp

| | Cohen's d | 95% CI | | Diễn giải kết quả |
|--------------------------|-----------|-----------|----------|--|
| | | Thấp nhất | Cao nhất | |
| <i>Class 1 - Class 3</i> | -1.495 | -1.785 | -1.208 | Học sinh Class 1 có hứng thú nghề nghiệp STEM thấp hơn Class 3 |
| <i>Class 2 - Class 3</i> | -0.932 | -1.062 | -0.802 | Học sinh Class 2 có hứng thú nghề nghiệp STEM thấp hơn Class 3 |
| <i>Class 1 - Class 2</i> | -0.750 | -1.025 | -0.478 | Học sinh Class 1 có hứng thú nghề nghiệp STEM thấp hơn Class 2 |

Nguồn: Kết quả xử lý từ dữ liệu điều tra

Phân tích ở Bảng 8 trên cho thấy có sự khác biệt rõ rệt về hứng thú nghề nghiệp giữa các nhóm, theo thứ tự Class 1 thấp nhất, đến Class 2 trung bình và Class 3 là cao nhất, tương tự với thứ tự của điểm số và sự tự tin về năng lực STEM.

4.6. Thảo luận

Kết quả phân tích mức độ ảnh hưởng của việc chia nhóm đến hứng thú nghề nghiệp đã giúp khẳng định hai yếu tố tác động đến hứng thú nghề nghiệp STEM, đó là sự tự tin về năng lực STEM và điểm số học tập các môn học STEM. Qua đó kết quả nghiên cứu, quy luật về mối liên hệ giữa hứng thú nghề nghiệp STEM và sự tự tin về năng lực STEM theo mô hình SCCT (Blotnicky, Franz-Ondendaal, French, & Joy, 2018; Lent & ctg., 1994; Nugent, Barker, Grandgenett, & Welch, 2016) cũng được chứng minh phù hợp với đối tượng học sinh phổ thông ở Việt Nam. Vì vậy, nhằm thúc đẩy học sinh lựa chọn theo học ngành nghề STEM nhiều hơn cần có các tác động sự phạm có thể ảnh hưởng đến sự tự tin vào năng lực STEM của học sinh. Kết quả phân nhóm bằng LPA cho thấy điểm số môn học STEM cao thì sự tự tin về năng lực STEM cũng cao, không có nhóm nào có sự đối lập (mức độ tự tin về năng lực cao nhưng điểm thấp hoặc ngược lại), điều này cho thấy mối liên hệ chặt chẽ giữa hai yếu tố trên. Sự tương quan giữa thái độ học tập các môn học STEM với kết quả học tập như trên phù hợp với kết quả của nhiều nghiên cứu trên thế giới (Berger & ctg., 2020; Glynn & ctg., 2009; van Aalderen-Smeets & ctg., 2019). Sự tương quan cao về kết quả học tập và sự tự tin về năng lực giữa các môn học, đặc biệt là giữa Toán với các môn Khoa học và Công nghệ cũng là một kết quả đáng lưu ý, bởi vì kiến

thức về toán học là không thể thiếu trong mọi lĩnh vực Khoa học, Công nghệ, sự tự tin về toán học của học sinh là một yếu tố quan trọng thúc đẩy sự tự tin về các môn học STEM khác (Berger & ctg., 2020). Dựa trên sự liên kết chặt chẽ giữa kết quả học tập STEM, sự tự tin về năng lực STEM với hứng thú nghề nghiệp STEM đã nêu, các nhà nghiên cứu trên thế giới đã đề ra nhiều biện pháp nhằm thúc đẩy học sinh lựa chọn học tập ngành nghề STEM nhiều hơn, bao gồm thay đổi phương pháp dạy học và đánh giá các môn học STEM, tăng cường phát triển năng lực giải quyết vấn đề, năng lực hợp tác trong dạy học (Berger & ctg., 2020), tổ chức các hoạt động ngoại khóa về STEM nhằm giúp học sinh có thêm được trải nghiệm về STEM (van Tuijl & van der Molen, 2016; Wang & ctg., 2023). Trong nghiên cứu này, kết quả phân tích hồi quy logistic từ ba nhóm được phân chia cũng cho thấy các học sinh được tham gia trải nghiệm STEM cũng có sự tự tin về năng lực STEM và điểm số môn học cao hơn so với nhóm học sinh không tham gia. Điều này đã chứng tỏ tính hiệu quả và cần thiết của việc tăng cường các hoạt động trải nghiệm STEM cho học sinh phổ thông ở Việt Nam hiện nay. Kết quả phân tích hồi quy logistic cũng cho thấy đối với hai nhóm chính, nam giới có xu hướng có điểm số và sự tự tin năng lực STEM cao hơn nữ giới. Đây là một vấn đề khiến nhiều nhà nghiên cứu trên thế giới quan tâm, vì là một trong những nhân tố chính gây ra sự thiếu hụt nhân lực nữ giới trong các ngành nghề STEM (Falco & Summers, 2019; Stewart & ctg., 2020; van Tuijl & van der Molen, 2016; Wang & ctg., 2023). Nam giới thường có xu hướng lựa chọn các ngành nghề STEM - khoa học tự nhiên, trong khi nữ giới có xu hướng ưu tiên các ngành nghề xã hội hơn (Wang & ctg., 2023). Định kiến xã hội ở nhiều quốc gia hiện nay vẫn cho rằng lĩnh vực STEM là khô khan, cực nhọc và không phù hợp với nữ giới (Falco & Summers, 2019; Luo & ctg., 2021; Wang & ctg., 2023). Nhiều phụ huynh vẫn không ủng hộ việc con gái của họ tham gia các ngành nghề STEM (van Tuijl & van der Molen, 2016; Wang & ctg., 2023). Những rào cản trên ảnh hưởng rất nhiều đến sự tự tin khi học STEM và lựa chọn nghề nghiệp STEM của các học sinh nữ. Để khắc phục tình trạng này, bên cạnh việc tăng cường các hoạt động trải nghiệm STEM cho học sinh (van Tuijl & van der Molen, 2016), các nhà nghiên cứu còn đề xuất việc tổ chức các hoạt động tham vấn hướng nghiệp riêng cho các học sinh nữ nhằm thúc đẩy lựa chọn nhóm ngành STEM (Falco & Summers, 2019). Đây là những giải pháp cần được xem xét và nghiên cứu khả năng áp dụng đối với học sinh phổ thông ở Việt Nam.

Kết quả nghiên cứu trên cũng cho thấy tiềm năng của phương pháp LPA trong nghiên cứu định lượng về giáo dục. Việc phân chia người học thành các nhóm ẩn giúp nhà nghiên cứu có thể đi sâu vào nghiên cứu những đặc điểm khác biệt trong một mẫu đối tượng đa dạng và nhiều khác biệt. Khả năng kết hợp với các phương pháp nghiên cứu khác như hồi quy logistic và kiểm định T-test càng khẳng định tính linh động và hiệu quả của phương pháp LPA. Bên cạnh những mặt tích cực, phương pháp LPA cũng có một số hạn chế nhất định. Kết quả phân nhóm phụ thuộc vào các biến đầu vào, tuy nhiên chưa có tiêu chí cụ thể để đánh giá số lượng biến đầu vào phù hợp cho phép phân tích, vốn ảnh hưởng rất lớn đến kết quả chia nhóm. Do vậy, nhà nghiên cứu cần cẩn thận trong việc nghiên cứu cơ sở lý thuyết để lựa chọn các biến đầu vào phù hợp. So với phương pháp mô hình cấu trúc tuyến tính (SEM), kết quả phân tích theo LPA cũng chưa cho thấy được sự tác động đa chiều giữa các yếu tố được khảo sát và mức độ tác động. Do vậy việc kết hợp LPA với SEM cũng là một hướng đi cần được xem xét.

5. Kết luận

Tóm lại, nghiên cứu này đã sử dụng phương pháp LPA phân nhóm học sinh dựa trên điểm số học tập và sự tự tin về năng lực STEM, kết hợp phân tích hồi quy logistic để đánh giá các nhân tố tác động đầu vào, gồm giới tính, lớp, trải nghiệm STEM và phân tích mức độ ảnh hưởng của phân nhóm lên hứng thú nghề nghiệp. Điều này cho phép các nhà giáo dục tiếp tục

nghiên cứu các biện pháp tác động sự phạm phù hợp cho từng nhóm đối tượng được phân chia, từ đó nâng cao kết quả học tập, sự tự đánh giá năng lực STEM và hứng thú nghề nghiệp STEM. Kết quả nghiên cứu cũng là nền tảng để mở rộng phương pháp LPA trong nghiên cứu định lượng về giáo dục cũng như các nghiên cứu dựa theo mô hình SCCT ở Việt Nam.

LỜI CẢM ƠN

Tạ Thanh Trung và Lê Châu Đạt được tài trợ bởi Chương trình học bổng đào tạo thạc sĩ, tiến sĩ trong nước của Quỹ Đổi mới sáng tạo Vingroup (VINIF), mã số lần lượt là VINIF.2022.ThS.097 và VINIF.2023.ThS.032.

Tài liệu tham khảo

- Achterhof, R., Huntjens, R. J. C., Meewisse, M. L., & Kiers, H. A. L. (2019). Assessing the application of latent class and latent profile analysis for evaluating the construct validity of complex posttraumatic stress disorder: Cautions and limitations. *European Journal of Psychotraumatology*, 10(1), 301-307. doi:10.1080/20008198.2019.1698223
- Akoglu, H. (2018). User's guide to correlation coefficients. *Turkish Journal of Emergency Medicine*, 18(3), 91-93. doi:10.1016/j.tjem.2018.08.001
- Bandura, A. (1986). Fearful expectations and avoidant actions as coefficients of perceived self-inefficacy. *American Psychologist*, 41(12), 1389-1391. doi:10.1037/0003-066X.41.12.1389
- Bechter, B. E., Dimmock, J. A., Howard, J. L., Whipp, P. R., & Jackson, B. (2018). Student motivation in high school physical education: A latent profile analysis approach. *Journal of Sport and Exercise Psychology*, 40(4), 206-216. doi:10.1123/jsep.2018-0028
- Berger, N., Mackenzie, E., & Holmes, K. (2020). Positive attitudes towards mathematics and science are mutually beneficial for student achievement: A latent profile analysis of TIMSS 2015. *The Australian Educational Researcher*, 47(3), 409-444. doi:10.1007/s13384-020-00379-8
- Blotnicky, K. A., Franz-Odenaal, T., French, F., & Joy, P. (2018). A study of the correlation between STEM career knowledge, mathematics self-efficacy, career interests, and career activities on the likelihood of pursuing a STEM career among middle school students. *International Journal of STEM Education*, 5(1), Article 22. doi:10.1186/s40594-018-0118-3
- Byrne, B. M. (2016). *Structural equation modeling with AMOS: Basic concepts, applications, and programming*. New York, NY: Routledge.
- Celeux, G., & Soromenho, G. (1996). An entropy criterion for assessing the number of clusters in a mixture model. *Journal of Classification*, 13(2), 195-212. doi:10.1007/BF01246098
- Ching, Y. H., Yang, D., Wang, S., Baek, Y., Swanson, S., & Chittoori, B. (2019). Elementary school student development of STEM attitudes and perceived learning in a STEM integrated robotics curriculum. *TechTrends*, 63(5), 590-601. doi:10.1007/s11528-019-00388-0
- Choi, A. Y., Nylund-Gibson, K., Israel, T., & Mendez, S. E. A. (2019). Correction to: A latent profile analysis of bisexual identity: Evidence of within-group diversity. *Archives of Sexual Behavior*, 48(7), 113-130. doi:10.1007/s10508-019-1423-8

- Cleveland, M. J., Collins, L. M., Lanza, S. T., Greenberg, M. T., & Feinberg, M. E. (2010). Does individual risk moderate the effect of contextual-level protective factors? A latent class analysis of substance use. *Journal of Prevention & Intervention in the Community*, 38(3), 213-228. doi:10.1080/10852352.2010.486299
- Falco, L. D., & Summers, J. J. (2019). Improving career decision self-efficacy and STEM self-efficacy in high school girls: Evaluation of an intervention. *Journal of Career Development*, 46(1), 62-76. doi:10.1177/0894845317721651
- Ferguson, S. L., Moore, E. W., & Hull, D. M. (2020). Finding latent groups in observed data: A primer on latent profile analysis in Mplus for applied researchers. *International Journal of Behavioral Development*, 44(5), 458-468. doi:10.1177/0165025419881721
- Fouad, N. A., & Smith, P. L. (1996). A test of a social cognitive model for middle school students: Math and science. *Journal of Counseling Psychology*, 43(3), 338-346. doi:10.1037/0022-0167.43.3.338
- Gehring, E. (2017). Self-assessment to improve learning and evaluation. In *2017 ASEE Annual conference & exposition proceedings*. Columbus, Ohio: ASEE Conferences.
- Glynn, S. M., Taasobshirazi, G., & Brickman, P. (2009). Science motivation questionnaire: Construct validation with nonscience majors. *Journal of Research in Science Teaching*, 46(2), 127-146. doi:10.1002/tea.20267
- Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77(6), 1121-1134. doi:10.1037/0022-3514.77.6.1121
- Lent, R. W., Brown, S. D., & Hackett, G. (1994). Toward a unifying social cognitive theory of career and academic interest, choice, and performance. *Journal of Vocational Behavior*, 45(1), 79-122. doi:10.1006/jvbe.1994.1027
- Luo, T., So, W. W. M., Wan, Z. H., & Li, W. C. (2021). STEM stereotypes predict students' STEM career interest via self-efficacy and outcome expectations. *International Journal of STEM Education*, 8(1), Article 36. doi:10.1186/s40594-021-00295-y
- Marsh, H. W., Lüdtke, O., Trautwein, U., & Morin, A. J. S. (2009). Classical latent profile analysis of academic self-concept dimensions: Synergy of person - and variable-centered approaches to theoretical models of self-concept. *Structural Equation Modeling: A Multidisciplinary Journal*, 16(2), 191-225. doi:10.1080/10705510902751010
- Ministry of Education and Training (MOET). (2018). *Chương trình Giáo dục phổ thông tổng thể [General Education Program]*. Truy cập ngày 10/10/2023 tại <https://moet.gov.vn>
- Ministry of Education and Training (MOET). (2020). *Công văn 3089/BGDĐT-GDTrH về việc Triển khai giáo dục STEM trong giáo dục trung học [3089th MOET Document on the Implementation of STEM education in secondary education]*. Truy cập ngày 10/10/2023 tại <https://moet.gov.vn>
- Morin, A. J. S., & Marsh, H. W. (2015). Disentangling shape from level effects in person-centered analyses: An illustration based on university teachers' multidimensional profiles of effectiveness. *Structural Equation Modeling: A Multidisciplinary Journal*, 22(1), 39-59. doi:10.1080/10705511.2014.919825

- Nguyen, N. T., Tran, Q. T. X., Nguyen, U. P., & Ta, T. T. (2022). Một số nghiên cứu về năng lực STEM trên thế giới và đề xuất khung năng lực STEM cho học sinh phổ thông tại Việt Nam [An overview study on STEM competencies in the world and propose a STEM competency framework for high school students in Vietnam]. *The Vietnam Journal of Education*, 22(10), 48-53.
- Nguyen, T. T. V. (2018). *Ứng dụng phân tích lớp tiềm ẩn xây dựng hệ thống trắc nghiệm thích nghi [Application of latent class analysis in adaptive quiz system development] (Master's thesis)*. Đà Nẵng, Việt Nam: Danang University of Science and Technology.
- Ning, H. K., & Downing, K. (2015). A latent profile analysis of university students' self-regulated learning strategies. *Studies in Higher Education*, 40(7), 1328-1346. doi:10.1080/03075079.2014.880832
- Nugent, G., Barker, B., Grandgenett, N., & Welch, G. (2016). Robotics camps, clubs, and competitions: Results from a US robotics project. *Robotics and Autonomous Systems*, 75(B), 686-691. doi:10.1016/j.robot.2015.07.011
- Nugent, G., Barker, B., Welch, G., Grandgenett, N., Wu, C., & Nelson, C. (2015). A model of factors contributing to STEM learning and career orientation. *International Journal of Science Education*, 37(7), 1067-1088. doi:10.1080/09500693.2015.1017863
- Pastor, D. A., Barron, K. E., Miller, B. J., & Davis, S. L. (2007). A latent profile analysis of college students' achievement goal orientation. *Contemporary Educational Psychology*, 32(1), 8-47. doi:10.1016/j.cedpsych.2006.10.003
- Rangel, V. S., Vaval, L., & Bowers, A. (2020). Investigating underrepresented and first-generation college students' science and math motivational beliefs: A nationally representative study using latent profile analysis. *Science Education*, 104(6), 1041-1070. doi:10.1002/sce.21593
- Reyes, A., Hermann, B. P., Busch, R. M., Drane, D. L., Barr, W. B., Hamberger, M. J., ... McDonald, C. R. (2022). Moving towards a taxonomy of cognitive impairments in epilepsy: Application of Latent profile analysis to 1178 patients with temporal lobe epilepsy. *Brain Communications*, 4(6), Article fcac289. doi:10.1093/braincomms/fcac289
- Rice, K. G., Lopez, F. G., & Richardson, C. M. E. (2013). Perfectionism and performance among STEM students. *Journal of Vocational Behavior*, 82(2), 124-134. doi:10.1016/j.jvb.2012.12.002
- Rohatgi, A., & Scherer, R. (2020). Identifying profiles of students' school climate perceptions using PISA 2015 data. *Large-Scale Assessments in Education*, 8(1), Article 4. doi:10.1186/s40536-020-00083-0
- Spurk, D., Hirschi, A., Wang, M., Valero, D., & Kauffeld, S. (2020). Latent profile analysis: A review and "How to" guide of its application within vocational behavior research. *Journal of Vocational Behavior*, 120, Article 103445. doi:10.1016/j.jvb.2020.103445
- Stewart, J., Henderson, R., Michaluk, L., Deshler, J., Fuller, E., & Rambo-Hernandez, K. (2020). Using the social cognitive theory framework to chart gender differences in the developmental trajectory of STEM self-efficacy in science and engineering students. *Journal of Science Education and Technology*, 29(6), 758-773. doi:10.1007/s10956-020-09853-5

- Tein, J. Y., Coxe, S., & Cham, H. (2013). Statistical power to detect the correct number of classes in latent profile analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, 20(4), 640-657. doi:10.1080/10705511.2013.824781
- van Aalderen-Smeets, S. I., van der Molen, J. H. W., & Xenidou-Dervou, I. (2019). Implicit STEM ability beliefs predict secondary school students' STEM self-efficacy beliefs and their intention to Opt for a STEM field career. *Journal of Research in Science Teaching*, 56(4), 465-485. doi:10.1002/tea.21506
- van Tuijl, C., & van der Molen, J. H. W. (2016). Study choice and career development in STEM fields: An overview and integration of the research. *International Journal of Technology and Design Education*, 26(2), 159-183. doi:10.1007/s10798-015-9308-1
- Wang, N., Tan, A. L., Zhou, X., Liu, K., Zeng, F., & Xiang, J. (2023). Gender differences in high school students' interest in STEM careers: A multi-group comparison based on structural equation model. *International Journal of STEM Education*, 10(1), Article 59. doi:10.1186/s40594-023-00443-6
- Zollman, A. (2012). Learning for STEM literacy: STEM literacy for learning. *School Science and Mathematics*, 112(1), 12-19. doi:10.1111/j.1949-8594.2012.00101.x

