



Bài báo nghiên cứu
XÂY DỰNG VÀ CHUẨN HÓA CÔNG CỤ ĐÁNH GIÁ
NĂNG LỰC STEM CỦA HỌC SINH TRUNG HỌC PHỔ THÔNG
TẠI THÀNH PHỐ HỒ CHÍ MINH

Tạ Thanh Trung, Trần Thị Xuân Quỳnh, Nguyễn Phương Uyên, Nguyễn Thành Nga*

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

*Tác giả liên hệ: Nguyễn Thành Nga – Email: nganthanhanh@hcmue.edu.vn

Ngày nhận bài: 31-3-2022; ngày nhận bài sửa: 04-4-2022; ngày duyệt đăng: 05-4-2022

TÓM TẮT

Hiện nay, đánh giá năng lực đang là xu hướng đánh giá của nền giáo dục trên thế giới. Theo đó, việc xây dựng và chuẩn hóa công cụ đánh giá năng lực đang được các nhà giáo dục đặc biệt quan tâm. Đối với giáo dục STEM, công cụ đánh giá năng lực STEM chuẩn hóa, phù hợp thực tiễn đóng vai trò quan trọng để định hướng phát triển trong tương lai. Nghiên cứu này nhằm mục đích xây dựng và chuẩn hóa được một công cụ đánh giá năng lực STEM của học sinh trung học phổ thông tại Thành phố Hồ Chí Minh trong quá trình học sinh học tập theo định hướng giáo dục STEM. Để đảm bảo độ tin cậy và độ giá trị của công cụ, quy trình nghiên cứu được thực hiện nghiêm ngặt với ba giai đoạn, gồm: thiết kế câu hỏi, phát triển thang đo, đánh giá thang đo. Kết quả phân tích định lượng cho thấy công cụ có độ tin cậy, độ giá trị và phù hợp để sử dụng đánh giá năng lực STEM của học sinh tại địa phương. Qua đó, bộ công cụ này là phương tiện để học sinh tự đánh giá năng lực của bản thân hoặc trở thành khung năng lực hỗ trợ giáo viên phát triển các công cụ đánh giá khác trong quá trình học sinh học tập theo định hướng giáo dục STEM.

Từ khóa: CB-SEM; đánh giá năng lực; học sinh trung học; năng lực STEM

1. Giới thiệu

Trong những năm gần đây, cùng với sự chuyển đổi chương trình giáo dục theo định hướng phát triển năng lực (competency-based approach), việc đánh giá trong giáo dục có xu hướng chung chuyển từ đánh giá khả năng tái hiện tri thức sang đánh giá năng lực của người học (Shulman, 2008). Đánh giá năng lực không chỉ đưa ra những nhận xét, kết luận có tính khách quan về quá trình và kết quả học tập, rèn luyện của học sinh, mà còn chú trọng xác định được những điểm mạnh, điểm hạn chế trong cấu trúc năng lực của học sinh. Điều này có ý nghĩa phản hồi định tính lẫn định lượng mức độ hiệu quả quá trình học tập của học sinh và dạy học của giáo viên. Qua đó, học sinh tìm ra phương hướng điều chỉnh hợp lý để nâng cao chất lượng học tập của bản thân hoặc giáo viên tìm ra những biện pháp giúp học sinh không ngừng tiến bộ, phát triển các năng lực và phẩm chất cần thiết nhằm chuẩn bị cho việc

Cite this article as: Ta Thanh Trung, Tran Thi Xuan Quynh, Nguyen Phuong Uyen, & Nguyen Thanh Nga (2022). Develop and standardize a STEM competency assessment tool for high school students in Ho Chi Minh City. *Ho Chi Minh City University of Education Journal of Science*, 19(8), 1255-1270.

thích ứng với sự biến động của thế giới, góp phần đảm bảo các mục tiêu giáo dục đã đề ra. Nghị quyết Hội nghị Ban Chấp hành Trung ương Đảng lần thứ tám (khóa XI) chỉ rõ: “Việc thi, kiểm tra và đánh giá kết quả giáo dục đào tạo cần tiến hành từng bước theo các tiêu chí tiên tiến được xã hội và cộng đồng giáo dục thế giới tin cậy và công nhận. Phối hợp sử dụng kết quả đánh giá trong quá trình học với đánh giá cuối kì, cuối năm học; đánh giá của người dạy với tự đánh giá của người học; đánh giá của nhà trường với đánh giá của gia đình và của xã hội” (General Secretary, 2013). Như vậy, có thể nhận định rằng hoạt động đánh giá hay đánh giá năng lực không phải là hoạt động riêng lẻ của giáo viên mà còn cần có sự kết hợp với hoạt động tự đánh giá, tự theo dõi quá trình phát triển năng lực của bản thân người học.

Đối với giáo dục STEM, vai trò của việc đánh giá năng lực người học càng được khăng định khi các công cụ đánh giá, đo lường năng lực STEM của học sinh được cho rằng rất cần thiết để nâng cao chất lượng của giáo dục STEM thông qua ý nghĩa phản hồi mức độ hiệu quả của mỗi giai đoạn học tập (Arikhan, Erktin, & Pesen, 2020). Tuy nhiên, để đánh giá năng lực một cách chuẩn xác của con người nói chung và của học sinh nói riêng là điều không dễ dàng, cần có một quy trình nghiên cứu nghiêm ngặt để đảm bảo độ tin cậy lẫn độ giá trị – các yếu tố thể hiện chất lượng của thang đo (Baartman et al., 2006). Do đó, các nhà nghiên cứu giáo dục trên thế giới đang không ngừng nỗ lực phối hợp giữa phương pháp nghiên cứu định tính và phương pháp nghiên cứu định lượng trong quá trình xây dựng công cụ đánh giá năng lực. Tại Việt Nam, phần lớn các công cụ đánh giá năng lực của học sinh khi học tập theo định hướng giáo dục STEM được xây dựng dựa trên khung lý thuyết kế thừa từ các công trình nghiên cứu trên thế giới, mà chưa có sự chuẩn hóa, điều chỉnh cho phù hợp với đặc điểm và mức độ triển khai giáo dục STEM trong thực tiễn bằng số liệu thực tế. Điều này dẫn đến khó khăn khi triển khai đánh giá đồng bộ hay lựa chọn công cụ đánh giá kết quả, quá trình học sinh học tập theo định hướng giáo dục STEM. Vậy nên, việc xây dựng và chuẩn hóa công cụ đánh giá năng lực STEM đảm bảo độ tin cậy, độ giá trị, phù hợp với mức độ triển khai giáo dục STEM đối với từng địa phương tại Việt Nam là điều cần thiết.

Chính vì lí do trên, chúng tôi thực hiện xây dựng và chuẩn hóa công cụ đánh giá năng lực STEM của học sinh phổ thông tại Thành phố Hồ Chí Minh (TPHCM). Qua quá trình khảo sát tài liệu, phỏng vấn chuyên sâu với đội ngũ chuyên gia, thu thập và phân tích thống kê dữ liệu từ 544 học sinh trung học phổ thông (THPT) đã trải qua quá trình học tập bài học STEM tại địa phương, kết quả nghiên cứu thu được công cụ đánh giá năng lực STEM của học sinh đảm bảo các tiêu chuẩn về độ tin cậy, độ giá trị và sự phù hợp với tổng thể.

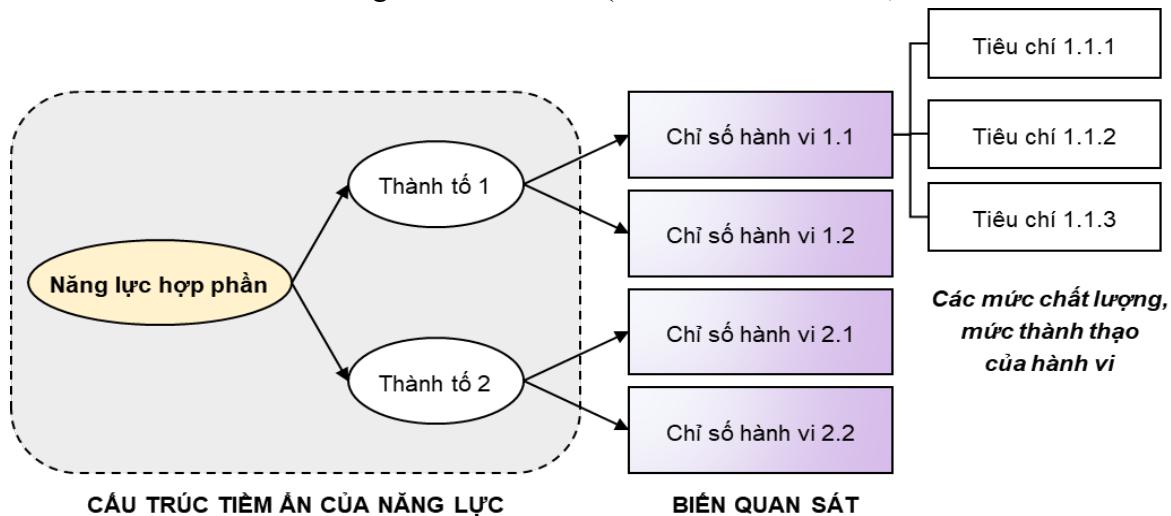
2. Đối tượng và phương pháp nghiên cứu

2.1. Đánh giá năng lực của học sinh với công cụ đánh giá chuẩn hóa

2.1.1. Đánh giá học sinh theo cách tiếp cận năng lực

Đánh giá năng lực của người học là quá trình quan trọng nhằm định hướng phát triển năng lực cho học sinh. Trong đó, năng lực được hiểu là thuộc tính cá nhân được hình thành, phát triển nhờ tố chất sẵn có và quá trình học tập, rèn luyện, cho phép con người huy động tổng hợp các kiến thức, kỹ năng và các thuộc tính cá nhân khác như hứng thú, niềm tin, ý chí,... thực hiện thành công một loại hoạt động nhất định, đạt kết quả mong muốn trong

những điều kiện cụ thể (MOET, 2018). Về cấu trúc, có thể mô hình hóa năng lực như một hệ thống được cấu thành từ các thành phần gồm: năng lực hợp phần, năng lực thành tố, chỉ số hành vi được biểu hiện thông qua quá trình học tập (Bình, 2015). Hợp phần của năng lực liên quan đến lĩnh vực chuyên môn tạo nên năng lực. Từ đó, các năng lực thành tố được bóc tách từ năng lực hợp thành mang những tính chất đặc trưng từ các lĩnh vực chuyên môn. Cả hai thành phần này đều là các yếu tố tiềm ẩn, không thể được đánh giá trực tiếp mà cần được đúc kết thông qua các biểu hiện của tập hợp các chỉ số hành vi tương ứng, trong một khoảng thời gian nhất định. Chỉ số hành vi là dữ liệu cung cấp thông tin về năng lực phục vụ kiểm tra đánh giá xuyên suốt quá trình học tập. Các hành vi được thể hiện ra bên ngoài qua các hoạt động: viết được, phát biểu được, tạo ra được... là cơ sở để đo lường nhằm xác định mức độ đạt được của năng lực. Như vậy, bản chất của đánh giá năng lực của học sinh là quá trình kiểm tra và đo lường mức độ đạt được của năng lực khi học sinh thực hiện thành công nhiệm vụ cũng như vấn đề trong quá trình học tập thông qua thu thập và phân tích các biểu hiện hành vi được bộc lộ ra bên ngoài của học sinh (Baartman et al., 2006).



Hình 1. Mô hình cấu trúc của một năng lực hợp phần

Kết quả đánh giá giúp giáo viên xác định những điểm nổi bật và hạn chế trong năng lực từ đó có định hướng hỗ trợ học sinh trong việc học tập và bồi dưỡng năng lực. Ngoài ra, đánh giá năng lực có tác động đến tầng sâu nhận thức của học sinh, thúc đẩy động cơ học tập, hình thành thói quen tốt, nuôi dưỡng đam mê, nhận ra giá trị của riêng mình, từ đó có định hướng tự rèn luyện, tự hoàn thiện bản thân. Xây dựng và chuẩn hóa khung cấu trúc năng lực trở thành tiêu chí tất yếu để tiến hành đánh giá năng lực của học sinh. Giáo viên có thể sử dụng khung cấu trúc để thiết kế bảng kiểm, rubric, bài kiểm tra; để đánh giá năng lực; xem xét xây dựng kế hoạch bài dạy phù hợp bồi dưỡng và phát triển năng lực của học sinh. Ngoài ra, học sinh có thể dựa vào khung cấu trúc năng lực để có kế hoạch tự theo dõi và rèn luyện bản thân.

2.1.2. Các tiêu chí về độ tin cậy, độ giá trị của công cụ đánh giá năng lực được chuẩn hóa

Theo trường phái thực chứng, độ giá trị và độ tin cậy là hai công cụ nhận thức cơ bản và là tiêu chuẩn căn bản để chuẩn hóa một công cụ đánh giá, ví dụ như công cụ đánh giá

năng lực STEM (Winter, 2000). Trong đó, độ giá trị (validity) của công cụ đo lường thể hiện tính đo lường chính xác đối tượng cần đo của công cụ. Cụ thể, để một công cụ đánh giá có độ giá trị thì nó phải thỏa mãn các tiêu chí: tính giá trị về nội dung (content validity); tính giá trị đo lường (measurement validity); tính giá trị cấu trúc (construct validity). Giá trị về nội dung phản ánh mức độ chính xác về mặt diễn đạt nội dung của công cụ đo lường. Như vậy, một công cụ đo lường cần được chuẩn hóa về mặt nội dung thông qua việc diễn đạt rõ nghĩa, trong tâm, hàm súc và để hiệu chỉnh về mặt nội dung hiệu quả thì công cụ cần được sự tham vấn về cách trình bày bởi các chuyên gia trong lĩnh vực qua phương pháp nghiên cứu Delphi. Giá trị đo lường được xác định thông qua giá trị hội tụ và giá trị phân biệt. Giá trị hội tụ cho biết mức độ tương quan giữa các biến đo lường (chỉ số hành vi) trong cùng một thành tố. Giá trị phân biệt thể hiện sự phân biệt giữa các năng lực thành tố này so với thành tố khác trong cùng cấu trúc năng lực. Trong nghiên cứu định lượng về công cụ đo lường năng lực, giá trị hội tụ được xác định bởi mức độ tương quan của các chỉ số hành vi, còn giá trị phân biệt được kiểm tra bằng cách xem xét giữa các thành tố trong năng lực có thực sự phân biệt với nhau hay không. Giá trị cấu trúc liên quan đến mức độ mà công cụ đo lường khái niệm, hành vi, ý tưởng, chất lượng của cấu trúc lí thuyết. Trong đánh giá năng lực, thông qua đo lường kiểm định giá trị cấu trúc, nhà nghiên cứu sẽ có thể mô hình hóa cấu trúc lí thuyết, chấp nhận hay hiệu chỉnh hoặc đề xuất cấu trúc mới phù hợp với kết quả định lượng.

Mặt khác, để trình bày khái niệm độ tin cậy, Joppe (2000) đã diễn đạt như sau: “Mức độ mà tại đó kết quả không thay đổi theo thời gian và thể hiện tính đại diện chính xác cho tổng thể được nghiên cứu gọi là độ tin cậy. Nếu kết quả nghiên cứu có thể được nhân rộng theo phương pháp luận tương tự thì khi đó, công cụ đo lường được xem là đáng tin cậy”. Như vậy, có thể hiểu độ tin cậy của công cụ đánh giá năng lực là mức độ ổn định của công cụ đó. Độ tin cậy của thang đo phải thỏa mãn những các tiêu chí sau: tính ổn định (stability); tính tương đương (equivalence); tính nhất quán (internal consistency). Tính ổn định của công cụ đánh năng lực thể hiện ở chỗ công cụ đo lường phải đảm bảo sự ổn định theo thời gian qua các mẫu tương đương. Nghĩa là, nếu có một thang đo được kiểm tra trên các nhóm mẫu có cùng đặc tính trong khoảng thời gian phù hợp thì thang đo sẽ cho kết quả tương tự. Cụ thể, khi kiểm tra tính ổn định của công cụ đánh giá năng lực, cần tiến hành khảo sát sơ bộ với mẫu nhỏ và khảo sát đại trà nhằm loại bỏ các biến không phù hợp. Tính tương đương công cụ đo lường được kiểm tra với các nhóm chuyên gia. Mỗi thành viên trong nhóm chuyên gia có cùng quan điểm về công cụ đo thì khi đó, công cụ đo đảm bảo tính tương đương. Công cụ đánh giá năng lực được kiểm chứng tính tương bằng cách xin ý kiến chuyên gia, những biến được sự đồng thuận cao sẽ được giữ nguyên, các biến có sự khác biệt về quan điểm sẽ được xem xét hiệu chỉnh. Tính nhất quán các nhóm nội tại của mẫu khi đo lường bằng một công cụ phải cho kết quả nhất quán. Chia mẫu thành các nhóm nhỏ chứa đầy đủ dữ liệu của mẫu, thực hiện đo lường trên từng nhóm mẫu, kết quả thu được phải có sự trùng khớp, nếu chưa có sự không phù hợp phải xem xét hiệu chỉnh.

2.2. Năng lực STEM của học sinh

2.2.1. Khái niệm về năng lực STEM

Giáo dục STEM là mô hình giáo dục dựa trên cách tiếp cận liên môn, giúp học sinh áp dụng các kiến thức khoa học, công nghệ, kỹ thuật và toán học vào giải quyết một số vấn đề thực tiễn trong bối cảnh cụ thể (MOET, 2018). Để đánh giá những tác động của mô hình giáo dục này đối với học sinh, Boon Ng (2019) đã nghiên cứu năng lực STEM của học sinh trong phạm vi giáo dục dựa trên các môn học STEM và kể cả các môn học không liên quan đến STEM (ngôn ngữ, văn hóa, văn học...). Kết quả nghiên cứu cho thấy khi học tập các môn học STEM, học sinh bộc lộ các biểu hiện hành vi đặc trưng từ các lĩnh vực khoa học, công nghệ, kỹ thuật, toán học (Boon Ng, 2019). Trong quá trình học tập, học sinh tiếp cận các nhiệm vụ học tập theo cách thức các nhà khoa học, nhà toán học, kỹ sư giải quyết vấn đề và tiến hành nghiên cứu. Học sinh chủ động khám phá kiến thức thông qua các hoạt động tìm kiếm, phân tích, tổng hợp sử dụng kiến thức phục vụ mục đích học tập. Khi tham gia vào các bài học STEM, học sinh tìm hiểu và xác định vấn đề liên quan đến thực tiễn. Từ những kiến thức, kỹ năng đã tích lũy, học sinh đề xuất và lựa chọn phương án tối ưu, thực hiện thao tác kỹ thuật để chế tạo sản phẩm hoặc mô hình. Trong suốt quá trình học tập, học sinh trao đổi, thảo luận nhóm nhằm trình bày các ý tưởng của bản thân, đóng góp ý kiến xây dựng bài học, phản biện bảo vệ quan điểm của nhóm. Học sinh thực hiện chia sẻ quá trình, kết quả, sản phẩm của dự án đến với mọi người. Từ kết quả phân tích, trong phạm vi bài báo nghiên cứu năng lực STEM của học sinh được hiểu là *khả năng huy động tổng hợp kiến thức, kỹ năng về khoa học, công nghệ, kỹ thuật và toán học để giải quyết vấn đề thực tiễn trong bối cảnh cụ thể, mang lại giá trị cho cá nhân và cộng đồng*.

2.2.2. Mô hình lý thuyết về năng lực STEM của học sinh

Cấu trúc năng lực STEM của học sinh được xây dựng dựa trên sự phân tích và tổng hợp các thành tố năng lực lặp đi lặp lại trong 13 đề tài nghiên cứu về khung cấu trúc năng lực STEM trên thế giới. Cấu trúc năng lực STEM gồm 5 thành tố:

Thành tố 1. Thu thập thông tin

Năng lực thành tố thu thập thông tin được hiểu là khả năng xác định thông tin cần tìm kiếm, tìm nguồn thông tin, ghi nhận nhận thông tin theo nhu cầu, mục đích ban đầu. Nghiên cứu của Bundy (2004) chỉ ra rằng năng lực thành tố thu thập thông tin giúp người học nhận ra nhu cầu tìm hiểu về thông tin, xác định mức độ thông tin cần thiết và truy cập thông tin hiệu quả. Căn cứ vào mô hình cấu trúc năng lực thông tin SCONUL của Bainton (2001), năng lực thành tố thu thập thông tin có các biểu hiện hành vi như: nhận ra nhu cầu thu thập thông tin, tìm kiếm nguồn thông tin, xây dựng chiến lược tìm kiếm thông tin, định vị và truy cập được thông tin. Năng lực thành tố thu thập thông tin ghi nhận được thông qua hành vi: lắng nghe và trả lời câu hỏi, quan sát, tìm kiếm nguồn thông tin, điều tra và khám phá, thu thập thông tin (lập kế hoạch và ghi nhận dữ liệu), đưa ra kết luận/quyết định (Boon Ng, 2019).

Thành tố 2. Xử lí và sử dụng thông tin

Nhóm biểu hiện hành vi xử lí và sử dụng thông tin của năng lực STEM biểu hiện tổng hợp các hành vi như tổng hợp dữ liệu, so sánh dữ liệu, phân loại dữ liệu, quản lý và lưu trữ dữ liệu, đánh giá, phân tích dữ liệu (Boon Ng, 2019). Mô hình SCONUL xác định năng lực xử lí và sử dụng thông tin được người học thực hiện thông qua việc so sánh, đánh giá thông

tin, tổ chức, tổng hợp thông tin (Bainton, 2001). Tác giả Bundy (2004) xác định các biểu hiện tương ứng với nhóm biểu hiện xử lý thông tin bao gồm: đánh giá nguồn thông tin, phân loại, lưu trữ thông tin, sử dụng thông tin để giải quyết vấn đề.

Thành tố 3. Thực hiện giải pháp

Năng lực thành tố thực hiện giải pháp là nhóm biểu hiện hành vi tiêu biểu của năng lực giải quyết vấn đề. Học sinh hiện thực hóa các giải pháp, thiết kế và tạo ra sản phẩm nhằm giải quyết các vấn đề (Watkins, Spencer, & Hammer, 2014). English (2016) xác định một số biểu hiện hành vi liên quan đến năng lực thực hiện giải pháp bao gồm: lập kế hoạch, chuẩn bị dụng cụ, thiết kế bản vẽ, đánh giá tính khả thi và chế tạo sản phẩm. Ngoài ra nghiên cứu của Sen và cộng sự (2018) chỉ ra năng lực thực hiện giải pháp được bộc lộ qua hành vi xác định nguyên nhân, thiếu sót của sản phẩm, thực hiện cải tiến và định hướng phát triển trong bối cảnh mới.

Thành tố 4. An toàn kĩ thuật

Năng lực thành tố an toàn kĩ thuật là năng lực thành tố bộc lộ kĩ năng thao tác kĩ thuật của học sinh đảm bảo các quy tắc an toàn trong quá trình thực hiện chế tạo sản phẩm. Năng lực thành tố an toàn kĩ thuật giúp cho học sinh hiểu về vai trò và cách thức sử dụng các công cụ/thiết bị kĩ thuật; lưu ý các quy tắc an toàn kĩ thuật nhằm bảo vệ bản thân và mọi người tránh những tai nạn đáng tiếc. Căn cứ vào nghiên cứu của Boon Ng (2019), biểu hiện của nhóm hành vi thực hành kĩ thuật bao gồm: sử dụng trang thiết bị (xác định được chức năng và sử dụng chính xác), bảo trì trang thiết bị (bảo quản và đảm bảo an toàn), hiệu chỉnh thiết bị và dụng cụ khi cần thiết.

Thành tố 5. Chia sẻ cộng đồng

Năng lực thành tố chia sẻ cộng đồng giúp học sinh đồng cảm với đối tượng đang gặp các vấn đề khó khăn, tạo động lực để các em chia sẻ quá trình, kết quả, sản phẩm học tập đến với cộng đồng. Trong khung chương trình giáo dục STEM của tổ chức The New York Academy of Sciences, yêu cầu cần tạo điều kiện cho học sinh có cơ hội thể hiện khả năng truyền đạt rõ ràng, chính xác, thuyết phục cùng việc phối hợp sử dụng nhiều phương thức để truyền đạt ý tưởng; sau đó, yêu cầu cần tạo môi trường để học sinh thảo luận, phản biện được các lập luận STEM (GSA, 2016). Tác giả Trilling (2009) đề xuất các hành vi chia sẻ cộng đồng: trình bày ý tưởng bằng các hình thức đã xác định (báo cáo, văn bản, poster, video...), lắng nghe góp ý và phát triển ý tưởng, xác định mục đích chia sẻ.

2.3. Phương pháp nghiên cứu

Nghiên cứu áp dụng quy trình của Boateng và cộng sự (2018) với ba giai đoạn: thiết kế câu hỏi, phát triển thang đo và đánh giá thang đo.

2.3.1. Giai đoạn thiết kế câu hỏi

Trong giai đoạn thiết kế câu hỏi, dựa trên nghiên cứu lí thuyết, công cụ đánh giá năng lực STEM của học sinh bao gồm 5 thành tố: Thu thập thông tin (TT), Xử lí và sử dụng thông tin (XS), Thực hiện giải pháp (TH), An toàn kĩ thuật (SD), Chia sẻ thông tin (CS). Như vậy, bảng câu hỏi sơ bộ gồm 15 chỉ số hành vi dạng Likert bao gồm một câu mô tả hành vi và một phần phản hồi tùy chọn với 5 mức độ. Trong đó, các chỉ số hành vi được mã hóa theo

thứ tự từ A01 đến A15. Kỹ thuật Delphi được tiến hành với 21 chuyên gia về giáo dục STEM, gồm 4 giảng viên đại học sư phạm và 17 giáo viên STEM có kinh nghiệm dạy học STEM từ 2 năm trở lên tại các trường trung học phổ thông để hiệu chỉnh giá trị về mặt nội dung của từng câu mô tả hành vi của người học. Các mô tả bị đánh giá là không đảm bảo độ giá trị về mặt nội dung sẽ bị xóa hoặc thay thế và những mục chưa rõ ràng được sửa đổi. Ngôn từ được diễn đạt để đảm bảo sự phù hợp với học sinh phổ thông. Kết quả thang đo sau khi nghiên cứu định tính sẽ được sử dụng để thiết kế bảng câu hỏi khảo sát.

2.3.2. Giai đoạn phát triển thang đo

Tiếp theo, các nghiên cứu định lượng trên quy mô mẫu nhỏ sẽ được tiến hành nhằm đánh giá sơ bộ thang đo các khái niệm của nghiên cứu. Tiêu chuẩn đánh giá thang đo trong bước này bao gồm độ tin cậy và độ giá trị thang đo. Độ tin cậy được đánh giá qua tiêu chí hệ số Cronbach's Alpha đạt trên 0,7 và hệ số tương quan biến-tổng lớn hơn 0,3 (Nunnally & Bernstein, 1994). Giá trị hội tụ của thang đo được đánh giá qua hệ số tải (FL) trong phân tích nhân tố khám phá EFA với ngưỡng hệ số tải đạt giá trị tối thiểu 0,5; tỉ lệ phần trăm phương sai tích lũy đạt trên 50% thì thang đo đạt giá trị hội tụ (Gerbing & Anderson, 1988). Ngoài ra, đối với mỗi biến quan sát, chênh lệch giữa hệ số tải tối đa và các hệ số tải khác (xét trị tuyệt đối) phải đạt từ 0,3 trở lên để đảm bảo giá trị phân biệt (Bryant & Yarnold, 1995). Qua các phân tích sơ bộ ở giai đoạn này sẽ giúp làm giảm số lượng các chỉ số hành vi trong công cụ đánh giá năng lực cũng như trích được số lượng thành tố trong cấu trúc năng lực STEM của học sinh.

2.3.3. Giai đoạn đánh giá thang đo

Cuối cùng, nghiên cứu chính thức được thực hiện trên quy mô mẫu lớn. Kết quả thu được sẽ qua quá trình đánh giá các hiệu lực còn lại của độ tin cậy và độ giá trị, cũng như xem xét sự phù hợp của mô hình cấu trúc với tổng thể bằng phần mềm AMOS 24.

❖ Đánh giá độ tin cậy, độ hội tụ, độ phân biệt

Độ tin cậy nhất quán nội tại (internal consistency reliability) của thang đo được xác định dựa vào hai chỉ số là hệ số Cronbach's alpha và độ tin cậy tổng hợp CR (composite reliability) (Devellis & Thorpe, 2021). Hệ số Cronbach's alpha càng lớn thì độ tin cậy nhất quán nội tại càng cao. Độ tin cậy tổng hợp có giá trị từ 0 đến 1 và được cho là tốt khi có giá trị lớn hơn 0,7.

Độ giá trị hội tụ (convergent validity) cho biết mức độ hội tụ hay chia sẻ một tỉ lệ phương sai chung của các chỉ báo cho một khái niệm, nghĩa là để đạt được giá trị hội tụ thì tập hợp các biến đo lường một khái niệm nghiên cứu phải có mối tương liên hệ tương quan cao (Kline, 2016). Điều này được thể hiện qua các hệ số hồi quy chuẩn hóa đối với từng biến của khái niệm tiềm ẩn nếu nó là thang đo đơn hướng. Để một thang đo đạt giá trị hội tụ thì hệ số chuẩn hóa của các biến thành phần của thang đo phải lớn hơn 0,5 và có ý nghĩa thống kê (Gerbing & Anderson, 1988). Ngoài ra, giá trị hội tụ cũng được đánh giá qua phương sai trích trung bình AVE (average variance extracted). Hệ số AVE lớn hơn hoặc bằng 0,5 sẽ khẳng định được độ giá trị hội tụ (Hair et al., 2019).

Độ giá trị phân biệt (discriminant validity) cho biết mức độ mà một khái niệm này thật sự phân biệt về mặt thống kê với một khái niệm khác cả về sự tương quan lẫn các biến đo lường

trong mỗi khái niệm (Hair et al., 2019). Cách tiếp cận truyền thống đánh giá tính phân biệt là tiêu chí Fornell-Larcker, được xác định bằng điều kiện hạch số căn bậc hai của AVE (square root of AVE – phải lớn hơn hạch số tương quan giữa các khái niệm (inter-construct correlations)). Ngoài ra, còn có tiêu chí đánh giá khác là tiêu chí tương quan giữa hai khái niệm phải nhỏ hơn 0,85 và phương sai riêng lớn nhất MSV (maximum shared variance) nhỏ hơn chỉ số AVE (Kline, 2016).

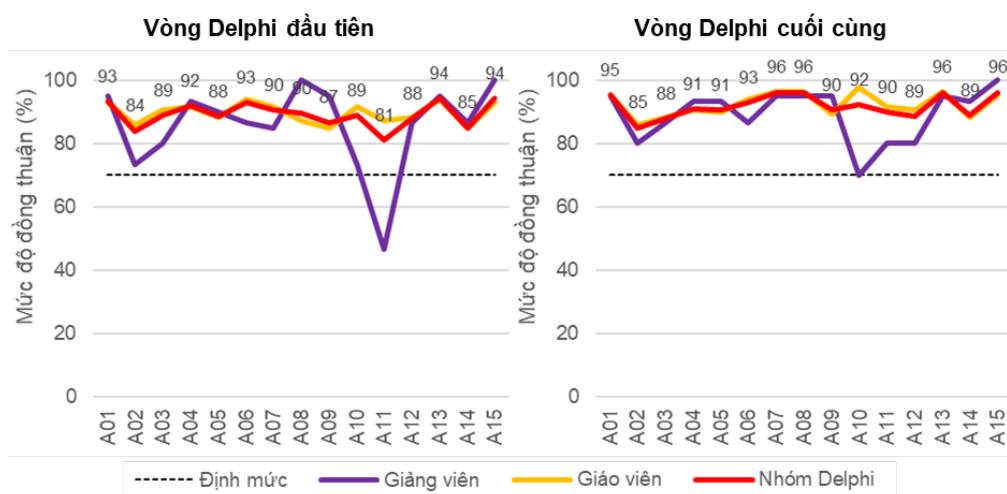
❖ *Kiểm định sự phù hợp của mô hình cấu trúc năng lực của học sinh*

Đầu tiên, chỉ bình phương/bậc tự do (χ^2 / df) là chỉ số để đánh giá mức độ phù hợp của mô hình đo lường và được đề xuất là nhỏ hơn 5 là phù hợp, nhỏ hơn 3 là tốt (Hair et al., 2019). Ngoài ra, các chỉ số phổ biến khác theo Kline (2016) là quan trọng bao gồm: chỉ số về độ phù hợp tuyệt đối – GFI (goodness-of-fit index); chỉ số điều chỉnh GFI theo bậc tự do trong mô hình – AGFI (adjusted goodness-of-fit statistics); chỉ số phù hợp tương đối – CFI (comparative fit index); chỉ số Tucker và Lewis – TLI (Tucker-Lewis index); chỉ số xác định mức độ phù hợp của mô hình so với tổng thể – RMSEA (root mean square error approximation). Các chỉ số GFI, AGFI, TLI và CFI có giá trị lớn hơn 0,9 được xem là mô hình thích hợp với dữ liệu thực tế (Hair et al., 2019). Hệ số RMSEA tối đa bằng 0,08 thì mô hình phù hợp tốt và RMSEA nhỏ hơn 0,03 là rất tốt.

3. Kết quả và thảo luận

3.1. Kết quả hiệu chỉnh độ giá trị về mặt nội dung

Kết quả từ phương pháp Delphi đã giúp nhóm nghiên cứu khai thác được hiệu quả các ý kiến về khung lý thuyết về năng lực STEM của học sinh phổ thông từ các chuyên gia. Trong vòng Delphi đầu tiên, ý kiến trong đội ngũ chuyên gia chưa có tính thống nhất cao tại nhiều chỉ số, cụ thể là mức độ đồng thuận có sự chênh lệch lớn (trên 10%) giữa các giảng viên đại học và giáo viên phổ thông tại các chỉ số A02, A03, A08, A09, A10, A11. Sau các vòng hiệu chỉnh Delphi, phần mô tả của các chỉ số hành vi trong khung năng lực STEM lý thuyết của học sinh đã được điều chỉnh về mặt nội dung theo sự thống nhất của đội ngũ chuyên gia Delphi (Hình 2).



Hình 2. Kết quả mức độ đồng thuận của đội ngũ chuyên gia Delphi trong vòng đầu (bên trái) và vòng cuối cùng (bên phải)

Với bảng hỏi cuối cùng, mức độ đồng thuận trung bình của các chuyên gia cho từng chỉ số hành vi của mỗi thành tố luôn từ 70% trở lên. Bên cạnh đó, nhóm giảng viên và giáo viên có sự tương đồng trong đánh giá mức độ đồng thuận ở đa số các chỉ số hành vi. Riêng các chỉ số hành vi về An toàn kỹ thuật (A10, A11, A12) thì vẫn còn sự chênh lệch. Trong khi nhóm giáo viên nhận định thành tố này là cần thiết được đánh giá và nội dung mô tả đã thể hiện rõ nội hàm thì nhóm giảng viên cho rằng các chỉ số này ít cần thiết hơn.

3.2. Mẫu nghiên cứu

Về quy mô mẫu, theo Hair và cộng sự (2019), cỡ mẫu cho các nghiên cứu có phân tích SEM và mô hình có từ 7 khái niệm nghiên cứu trở xuống thì cỡ mẫu tối thiểu là 300. Đề tài nghiên cứu áp dụng phương pháp chọn mẫu ngẫu nhiên tích tụ để thu thập thông tin của các học sinh tại 9 trường THPT, gồm: THPT Lương Thế Vinh (Quận 1), THPT Chuyên Lê Hồng Phong (Quận 5), Phổ thông Năng khiếu – Đại học Quốc gia (Quận 5), THPT Phạm Phú Thứ (Quận 6), THCS – THPT Nam Sài Gòn (Quận 7), THPT Nguyễn Du (Quận 10), THPT Nguyễn Hữu Cánh (Quận Bình Tân), THPT Gia Định (Quận Bình Thạnh) và THCS – THPT Hoa Sen (Thành phố Thủ Đức). Việc lấy mẫu nghiên cứu diễn ra trong khoảng thời gian từ tháng 12 năm 2021 đến đầu tháng 4 năm 2022. Sau khi gửi các phiếu khảo sát bằng cả hình thức trực tuyến và trực tiếp đến các nhà trường THPT trên địa bàn thành phố, kết quả có hơn 887 học sinh tham gia khảo sát. Sau quá trình kiểm tra và loại bỏ những mẫu không đáng tin cậy và những học sinh chưa từng học tập theo định hướng giáo dục STEM, mẫu gồm 544 học sinh đã được lựa chọn để tiến hành phân tích.

Bảng 2. Đặc điểm thống kê của mẫu

Đặc điểm	Tiêu chí	Tần số	Tỉ lệ
Giới tính	Nam	252	46,3 %
	Nữ	292	53,7 %
Khối lớp	10	299	55,0 %
	11	172	31,6 %
	12	73	13,4 %
Loại hình trường	Trường công lập	402	73,9 %
	Trường chuyên	65	11,9 %
	Trường dân lập	77	14,2 %
Hình thức tiếp cận giáo dục STEM*	- Trong một tiết học STEM chính khóa	163	30 %
	- Trong một dự án học tập hoặc dự án cộng đồng	99	18,2 %
	- Trung tâm ngoài nhà trường	14	2,5 %
	- Khóa học trực tuyến	31	5,7 %
	- Trải nghiệm không gian STEM tại trường (phòng trưng bày sản phẩm, mô hình, ứng dụng STEM...)	58	10,7 %
	- Sân chơi, hội chợ khoa học	122	22,4 %
	- Trải nghiệm thực tế tại các cơ sở nghiên cứu, cơ sở kinh doanh, doanh nghiệp...	20	3,7 %
	- Sinh hoạt câu lạc bộ STEM	51	9,4 %
- Tham gia các cuộc về khoa học kỹ thuật (chế tạo tên lửa nước, máy rửa tay...)		30	5,5 %
Hình thức khác		178	32,7 %

Ghi chú: * Mỗi học sinh có thể tiếp xúc nhiều hơn 1 hình thức.

Về phân phối của mẫu, bộ mẫu này cũng đảm bảo tiêu chuẩn về phân phối chuẩn khi trị trung bình và trung vị gần bằng nhau, giá trị tuyệt đối của độ xiên skewness luôn nhỏ hơn 2 (lớn nhất là $|skewness_{(A05)}| = 0,706$) và độ nghiên kurtosis nhỏ hơn 2 (lớn nhất là $|kurtosis_{(A05)}| = 1,076$). Như vậy, dữ liệu này đủ các điều kiện về quy mô và phân phối của mẫu để phân tích CB-SEM.

3.3. Đánh giá sơ bộ độ tin cậy và giá trị của thang đo

Trước khi tiến hành kiểm định mô hình đo lường năng lực STEM của học sinh bằng CB-SEM, phân tích nhân tố khám phá được tiến hành nhằm thu gọn hoặc loại các chỉ số hành vi không phù hợp trong thang đo năng lực và xác định số lượng thành tố năng lực có trong cấu trúc năng lực STEM của học sinh. Kết quả phân tích sơ bộ cho thấy biến A01 của thang đo Thu thập thông tin và A06 của thang đo Xử lí và sử dụng thông tin có chênh lệch hệ số tải nhân tố giữa các nhóm nhân tố dưới 0,3, cho thấy các biến này không đảm bảo độ phân biệt nên bị loại ra. Bên cạnh đó, phân tích EFA sơ bộ thu được 3 thành tố của năng lực thay vì 5 thành tố như ban đầu. Nhận thấy các chỉ số hành vi của học sinh vẫn tuân theo tiến trình học tập theo định hướng giáo dục STEM, chúng tôi đặt lại tên các thành tố là: Thu thập và xử lí thông tin (IN), Thực hiện giải pháp và chia sẻ cộng đồng (PS), Sử dụng thiết bị và dụng cụ kỹ thuật (TE). Các chỉ số hành vi cũng được mã hóa tương ứng: thay A02, A03, A04, A05 lần lượt bằng IN01, IN02, IN03, IN04; thay A07, A08, A09, A10, A11, A12 lần lượt bằng PS01, PS02, PS03, PS04, PS05, PS06 và thay A13, A14, A15 lần lượt bằng TE01, TE02, TE03. Thực hiện phân tích EFA lại với mẫu nghiên cứu chính thức, kết quả được thể hiện qua Bảng 3.

Bảng 3. Kết quả EFA thang đo các thành tố trong cấu trúc năng lực STEM của học sinh

Chỉ số hành vi		Nhóm nhân tố		
Mã hóa trước	Mã hóa sau	1	2	3
A07	PS01	0,874		
A08	PS02	0,902		
A09	PS03	0,725		
A10	PS04	0,644		
A11	PS05	0,669		
A12	PS06	0,637		
A04	IN03		0,768	
A05	IN04		0,853	
A06	*		*	
A01	*		*	
A02	IN01		0,698	
A03	IN02		0,578	
A13	TE01			0,806
A14	TE02			0,840
A15	TE03			0,887
Eigenvalue		7,135	1,194	1,087
Phương sai trích		54,884	9,187	8,365
Cronbach's Alpha		0,910	0,840	0,903
Tổng phương sai trích =	72,436 %;			

Kết quả kiểm định Cronbach's Alpha cho thang đo cho thấy các thang đo này đều có độ tin cậy cao khi hệ số Cronbach's Alpha $> 0,8$ và hệ số tương quan biến tổng đều $> 0,3$ vì thế không có biến quan sát nào bị loại khỏi mô hình. Bên cạnh đó, hệ số KMO $> 0,5$ với $\text{Sig}_{\text{Bartlett's}} = 0,000 < 0,05$ cho thấy dữ liệu có điều kiện đủ để tiến hành phân tích nhân tố. Tiếp theo, thực hiện phân tích nhân tố theo phép trích Principal Component Analysis với phép quay không vuông góc Varimax. Giá trị tổng phương sai trích đều $> 50\%$: đạt yêu cầu. Và với kết quả thu được ta có thể nói rằng 3 nhân tố IN, PS, TE giải thích 72,436% biến thiên của dữ liệu của năng lực STEM của học sinh. Tất cả các hệ số tải nhân tố đều lớn hơn 0,5 và chênh lệch giữa các biến thành phần của hai nhân tố đều trên 0,3 nên các nhân tố đảm bảo được giá trị hội tụ và phân biệt khi phân tích EFA. Như vậy, với thang đo chính thức, sau khi phân tích nhân tố thì các nhân tố độc lập này được giữ nguyên, không bị tăng thêm hoặc giảm đi nhân tố.

3.4. Đánh giá mô hình cấu trúc năng lực STEM của học sinh bằng CB-SEM

3.4.1. Đánh giá độ tin cậy thang đo năng lực STEM của học sinh

Để kiểm định độ tin cậy của thang đo, cần thực hiện đánh giá thông qua hệ số Cronbach's Alpha và hệ số CR. Các kết quả Cronbach's Alpha ở Bảng 3 đã cho thấy các thang đo đạt yêu cầu về tính nhất quán nội tại cao. Bên cạnh đó, căn cứ vào Bảng 4, ta thấy hệ số tin cậy tổng hợp của các khái niệm đều lớn hơn 0,7, nhỏ nhất là nhân tố IN với $\text{CR}_{\text{IN}} = 0,842$. Như vậy, từ kết quả nghiên cứu, có thể thấy rằng những thang đo của các thành tố trong năng lực STEM của học sinh đạt yêu cầu về độ tin cậy.

Bảng 4. Kết quả đánh giá độ tin cậy và hội tụ của thang đo

Nhân tố	Độ tin cậy tổng hợp (CR)	Phương sai trích (AVE)
IN	0,842	0,572
PS	0,891	0,578
TE	0,904	0,758

3.4.2. Đánh giá giá trị hội tụ thang đo năng lực STEM của học sinh

Trong Bảng 3, cột hệ số tải của kết quả kiểm định giá trị hội tụ thể hiện hệ nhân tố chuẩn hóa của các biến thành phần của các thang đo. Các giá trị phương sai trích trung bình (AVE) cũng có hệ số đạt yêu cầu khi đều trên 0,5, nhỏ nhất là nhân tố AB với $\text{AVE}_{\text{IN}} = 0,572$. Như vậy, các thang đo đạt yêu cầu về giá trị hội tụ.

3.4.3. Đánh giá giá trị phân biệt thang đo năng lực STEM của học sinh

Đánh giá tính phân biệt của thang đo, kết quả nghiên cứu tại Bảng 5 cho thấy tính phân biệt của thang đo đã đạt được vì cản bậc hai của các giá trị AVE (đường chéo in đậm) cao hơn so với các tương quan ngoài đường chéo. Ví dụ như thang đo PS có chỉ số AVE = 0,578 (Bảng 4) và cản bậc hai của giá trị này đạt 0,760 (Bảng 5). Như vậy, 0,760 cao hơn cả giá trị tương quan cùng hàng với nó (0,745) và cùng cột với nó (0,703).

Bảng 5. Hệ số tương quan giữa các khái niệm

Khái niệm	IN	PS	TE
IN	0,756		
PS	0,745	0,760	
TE	0,671	0,703	0,870
MSV	0,555	0,555	0,494

Ngoài ra, hệ số tương quan giữa các khái niệm đều nhỏ hơn 0,85 và hệ số MSV trong khoảng từ 0,494 đến 0,555 đều đạt yêu cầu khi nhỏ hơn AVE của các khái niệm. Như vậy, thang đo đạt yêu cầu về giá trị phân biệt.

3.4.4. Dánh giá sự phù hợp của mô hình năng lực STEM của học sinh

Bảng 6. Kết quả mức độ phù hợp mô hình

Chỉ số	Tiêu chí	Kết quả	Dánh giá
χ^2 / df	< 5,0	3,339	Phù hợp
GFI	> 0,9	0,952	Phù hợp
CFI	> 0,9	0,973	Phù hợp
TLI	> 0,9	0,953	Phù hợp
RMSEA	< 0,08	0,066	Phù hợp

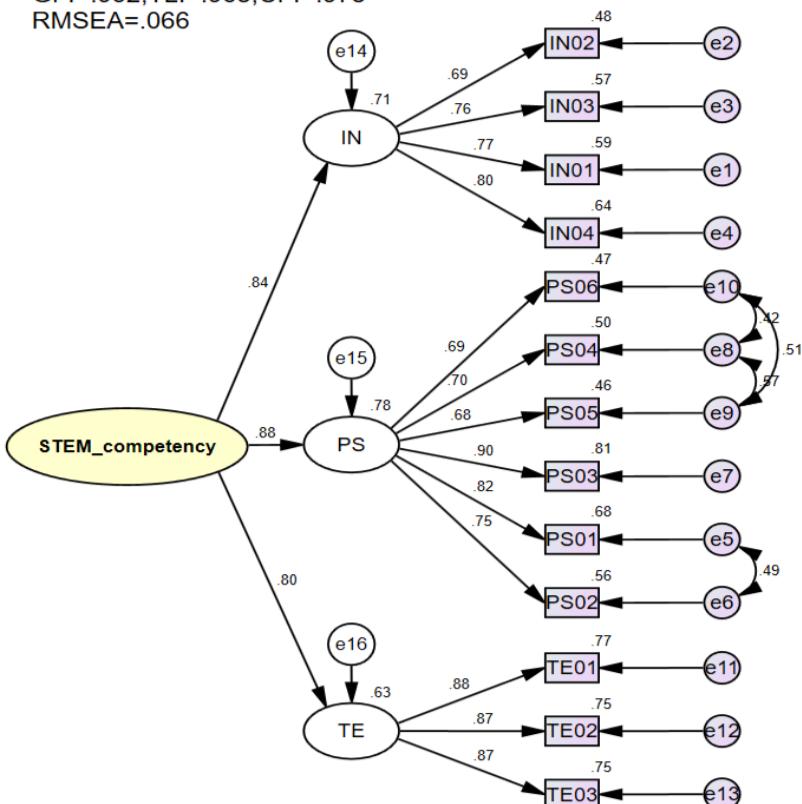
Mô hình năng lực STEM của học sinh được biểu diễn bằng mô hình CB-SEM bậc 2 với 58 bậc tự do, cho $\chi^2 = 193,679$ với giá trị $p = 0,000$. Kết quả kiểm định sự phù hợp của mô hình đo lường cho thấy đa số các chỉ số đạt yêu cầu về sự phù hợp của mô hình đo lường: $\chi^2 / df = 3,339 (< 5)$; TLI = 0,963 ($> 0,95$); CFI = 0,973 ($> 0,95$); GFI = 0,952 ($> 0,95$); RMSEA = 0,066 ($< 0,08$). Điều này cho thấy mô hình cấu trúc năng lực STEM của học sinh được tái cấu trúc qua phân tích CB-SEM là phù hợp với dữ liệu thực tiễn tại địa bàn TPHCM.

Chi-square=193.679;df=58;P=.000

Chi-square/df=3.339

GFI=.952;TLI=.963;CFI=.973

RMSEA=.066



Hình 3. Mô hình SEM của cấu trúc năng lực STEM của học sinh thể hiện bằng AMOS 24

Dựa trên trọng số của các đường dẫn trong mô hình, có thể thấy các thành tố năng lực PS, IN, TE có mối tương quan chặt chẽ với năng lực hợp phần STEM. Trong đó, thành tố PS có hệ số chuẩn hóa lớn nhất là 0,88, thể hiện thành tố này có ý nghĩa quan trọng hơn các thành tố khác trong khung cấu trúc. Các thành tố khác có hệ số ít hơn nhưng mức chênh lệch là không nhiều.

3.5. Công cụ đánh giá năng lực STEM sau khi chuẩn hóa

Từ kết quả của mô hình cấu trúc, công cụ đánh giá năng lực cho học sinh phổ thông trong quá trình học tập theo định hướng giáo dục STEM đã chuẩn hóa được trình bày qua Bảng 7. Công cụ này sẽ giúp học sinh có thể theo dõi và ghi nhận đường phát triển năng lực của bản thân trong quá trình học tập, từ đó không ngừng hoàn thiện và phát triển. Bên cạnh đó, giáo viên cũng có thể căn cứ vào công cụ này để mô tả bổ sung các mức độ thành thạo, chất lượng của chỉ số hành vi, tạo thành công cụ bảng kiểm quan sát đánh giá năng lực cho học sinh.

Bảng 7. Công cụ đánh giá năng lực STEM của học sinh trung học phổ thông tại TPHCM sau khi chuẩn hóa

Năng lực thành tố hành vi	Chỉ số	Mô tả	Mức độ						
			1	2	3	4	5	6	7
Thu thập và xử lý thông tin (IN)	IN01	Định vị thông tin: Lựa chọn đa dạng nguồn chứa thông tin cần tìm kiếm (sách báo, internet, người có chuyên môn...).	<input type="checkbox"/>						
	IN02	Thu thập thông tin: Ghi nhận thông tin từ nguồn chứa thông tin đã lựa chọn	<input type="checkbox"/>						
	IN03	Đánh giá thông tin: Đánh giá chất lượng nội dung và mức độ tin cậy của các nguồn thông tin đã thu thập trước khi sử dụng	<input type="checkbox"/>						
	IN04	Tổng hợp thông tin: Tổng hợp tất cả thông tin dự kiến sẽ sử dụng thành ý hiểu của bản thân	<input type="checkbox"/>						
Thực hiện giải pháp và chia sẻ công đồng (PS)	PS01	Lập kế hoạch thực hiện giải pháp: Xây dựng kế hoạch thực hiện giải pháp, trong đó đảm bảo các khâu về thiết kế giải pháp, chuẩn bị vật liệu, phân công công việc, cách thức tiến hành chế tạo, thử nghiệm...	<input type="checkbox"/>						
	PS02	Thực hiện giải pháp: Thực hiện mô hình, sản phẩm vật chất để thực hiện hóa giải pháp giải quyết vấn đề	<input type="checkbox"/>						
	PS03	Cải tiến giải pháp: Đề xuất một số phương án cải tiến giải pháp giải quyết vấn đề trong bối cảnh mới	<input type="checkbox"/>						
	PS04	Lựa chọn hình thức chia sẻ: Lựa chọn hình thức chia sẻ kết quả thực hiện giải pháp đến với cộng đồng	<input type="checkbox"/>						
	PS05	Trình bày kết quả: Trình bày kết quả thực hiện giải pháp theo hình thức chia sẻ đã lựa chọn	<input type="checkbox"/>						
	PS06	Phản biện: Phản biện các ý kiến về sản phẩm học tập	<input type="checkbox"/>						
Sử dụng thiết bị và dụng cụ kĩ thuật (TE)	TE01	Đảm bảo an toàn kĩ thuật: Xác định được các thông tin kĩ thuật và tiêu chí an toàn đối với từng thiết bị, dụng cụ trước khi sử dụng	<input type="checkbox"/>						
	TE02	Thao tác kĩ thuật: Thực hiện thao tác đúng kĩ thuật trong quá trình sử dụng các thiết bị, dụng cụ kĩ thuật	<input type="checkbox"/>						
	TE03	Bảo quản thiết bị, dụng cụ kĩ thuật: Kiểm tra tình trạng thiết bị trước và sau khi sử dụng để hạn chế tối đa tình trạng hư hỏng	<input type="checkbox"/>						

Qua quá trình xây dựng và chuẩn hóa công cụ, kết quả nghiên cứu cho thấy không phải chỉ số hành vi nào cũng phù hợp để đưa vào công cụ đánh giá. Ngoài việc, các chỉ số hành vi trong công cụ đánh giá được diễn đạt định tính một cách chính xác về mặt nội dung thì còn phải thỏa mãn những tiêu chí định lượng về độ tin cậy và độ giá trị. Một số tiêu chí đánh giá nếu chỉ được rút ra từ quá trình nghiên cứu lí thuyết mà thiếu công đoạn kiểm chứng định lượng thì việc đánh giá sẽ mang nhiều tính chủ quan, khó đảm bảo tính chính xác, khách quan khi người đánh giá đưa ra những kết quả và nhận định về năng lực của người học.

4. Kết luận

Như vậy, nghiên cứu đã phát triển và chuẩn hóa công cụ đánh giá năng lực cho học sinh phổ thông trong quá trình học tập theo định hướng giáo dục STEM. Công cụ sau khi chỉnh sửa bao gồm 3 thành tố của năng lực STEM bao gồm: Thu thập và xử lý thông tin (IN), Thực hiện giải pháp và chia sẻ cộng đồng (PS), Sử dụng thiết bị và dụng cụ kĩ thuật (TE) được đánh giá dưới dạng thang Likert (7 mức độ). Điểm trung bình của tất cả các chỉ số hành vi càng cao thì cho thấy khả năng đáp ứng việc học tập theo định hướng giáo dục STEM của học sinh càng cao. Thông qua quy trình nghiên cứu được thực hiện nghiêm ngặt gồm ba giai đoạn: thiết kế câu hỏi, phát triển thang đo và đánh giá thang đo, kết quả phân tích CB-SEM đã cho thấy công cụ được xây dựng đã đảm bảo các tiêu chí về độ tin cậy nhất quán nội tại và độ giá trị.

❖ **Tuyên bố về quyền lợi:** Các tác giả xác nhận hoàn toàn không có xung đột về quyền lợi.

TÀI LIỆU THAM KHẢO

- Arikan, S., Erktin, E., & Pesen, M. (2020). Development and validation of a STEM competencies assessment framework. *International Journal of Science and Mathematics Education*, 20(1), 1-24. doi:10.1007/s10763-020-10132-3
- Baartman, L. K., Bastiaens, T. J., Kirschner, P. A., & Vleuten, C. P. (2006). The wheel of competency assessment: Presenting quality criteria for competency assessment programs. *Studies in Educational Evaluation*, 32(2), 153-170. doi:10.1016/j.stueduc.2006.04.006
- Bainton, T. (2001). *Information literacy and academic libraries: the SCONUL approach*. United Kingdom: IFLA Council and general conference.
- Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best Practices for Developing and Validating Scales for Health, Social, and Behavioral Research: A Primer. *Front Public Health*, 6: 149, 1-18. doi:10.3389/fpubh.2018.00149
- Boon Ng, S. (2019). Exploring STEM competences for the 21st century. *Progress Reflection No. 30 On Current and Critical Issues in Curriculum, Learning and Assessment*. UNESCO International Bureau of Education.
- Bryant, F. B., & Yarnold, P. R. (1995). Principal-components analysis and exploratory and confirmatory factor analysis. In L. G. Grimm, & P. R. Yarnold, *Reading and understanding multivariate statistics* (pp. 99-136). Washington DC: American Psychological Association.
- Bundy, A. (2004). *Australian and New Zealand Information Literacy Framework*.

- Devellis, R., & Thorpe, C. T. (2021). *Scale Development Theory and Applications* (5th ed.). Sage Publications.
- English, L. D. (2016). Advancing integrated STEM learning through engineering design: sixth-grade students' design and construction of earthquake resistant buildings. *The Journal of Educational Research*, 110(3), 255-271. doi:10.1080/00220671.2016.1264053
- Fornell, C., & Larcker, D. F. (1981). Evaluating Structural Equation Models with Unobservable Variables and Measurement Error. *Journal of Marketing Research*, 18(1), 39-50. doi:10.1177/002224378101800104
- General Secretary. (2013). *On fundamental and comprehensive renovation of education and training, meeting the requirements of industrialization and modernization in the conditions of socialist-oriented market economy and international integration* (Vols. 29-NQ/TW). Ha Noi.
- Gerbing, D. W., & Anderson, J. C. (1988). An Updated Paradigm for Scale Development Incorporating Unidimensionality and Its Assessment. *Journal of Marketing Research*, 25(2), 186-192. doi:10.2307/3172650
- GSA. (2016). *STEM education framework*. New York. Retrieved from https://www.nyas.org/media/13051/gsa_stem_education_framework_dec2016.pdf
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate Data Analysis* (8th ed.). Cengage Learning, EMEA.
- Hoang, H. B. (2015). Competence and assessment by competence [Nang luc va danh gia theo nang luc]. *Ho Chi Minh City University of Education Journal of Science*, 6(71), 21-32. doi:10.544607/hcmue.js.0.6(71).667.658(2015)
- Jang, H. (2015). Identifying 21st Century STEM Competencies Using Workplace Data. *Journal of science education and technology*, 25, 284-301. doi:10.1007/s10956-015-9593-1
- Joppe, M. (2000). *The research process*. Retrieved from <https://www.uoguelph.ca/hftm/research-process>
- Kline, R. B. (2016). *Principles and Practice of Structural Equation Modeling* (4th ed.). The Guilford Press.
- Marope, M., Griffin, P., & Gallagher, C. (2017). *Future Competences and the Future of Curriculum*. Geneva: UNESCO-IBE. Retrieved from http://www.ibe.unesco.org/sites/default/files/resources/future_competences_and_the_future_of_curriculum.pdf
- Ministry of Education and Training – MOET. (2018). General education program [Chuong trinh giao duc pho thong tong the]. Hanoi. Retrieved from <https://data.moet.gov.vn/index.php/s/LETzPhj5sGKnDii#pdfviewer>
- Nunnally, J. C., & Bernstein, I. H. (1994). The Assessment of Reliability. In *Psychometric Theory* (Vol. 3, pp. 248-292). New York: McGraw-Hill.
- Sen, C., Ay, Z. S., & Kiray, S. A. (2018). STEM skills in the 21st-century education. In M. Shelley, & S. A. Kiray, *Research Highlights in STEM Education* (pp. 81-101). ISRES Publishing.
- Shulman, L. (2008). Assessment of teaching or assessment for teaching? In D. H. Gitomer (Ed.), *Measurement Issues and Assessment for Teaching Quality* (pp. 234-243). Washington DC: Sage Publications.
- Trilling , B., & Fadel, C. (2009). *21st Century Skills: Learning for Life in Our Times*. San Francisco: Jossey - Bass .
- Watkins, J., Spencer, K., & Hammer, D. (2014). Examining Young Students' Problem Scoping in Engineering Design. *Journal of Pre-College Engineering Education Research (J-PEER)*, 4(1). doi:10.7771/2157-9288.1082
- Winter, G. (2000). A Comparative Discussion of the Notion of 'Validity' in Qualitative and Quantitative Research. *The Qualitative Report*, 4(3), 1-14. doi:10.46743/2160-3715/2000.2078

**DEVELOP AND STANDARDIZE A STEM COMPETENCY ASSESSMENT TOOL
FOR HIGH SCHOOL STUDENTS IN HO CHI MINH CITY**

Tạ Thành Trung, Trần Thị Xuân Quynh, Nguyễn Phuong Uyen, Nguyễn Thành Nga*

Ho Chi Minh City University of Education, Vietnam

*Corresponding author: Nguyễn Thành Nga – Email: nganthanh@hcmue.edu.vn

Received: March 31, 2022; Revised: April 04, 2022; Accepted: April 05, 2022

ABSTRACT

Nowadays, competency assessment is one of the assessment trends across the globe. Accordingly, the development and standardization of competency assessment tools attract much attention from educators. For STEM education, a standardized and practical STEM competency assessment tool plays a vital role in guiding development in the future. This study aims to develop and standardize a tool for self-assessing STEM competencies of students in Ho Chi Minh. To guarantee the reliability and validity of the tool, the research process was strictly carried out in three stages, including questionnaire design, scale development, and scale evaluation. Quantitative analysis results show that the tool is reliable, valid, and appropriate to evaluate the STEM competency of students in the region. Thereby, this tool is suitable for students to self-assess their competency or a competency framework to support teachers in formulating other assessment tools in STEM learning.

Keywords: CB-SEM; competency assessment; high school students; STEM competency