

## DEVELOPMENT OF A STEM LESSON EVALUATION SCALE FOR SECONDARY EDUCATION IN VIETNAM

Kieu Phuong Thuy<sup>1</sup>, Pham Thi Binh<sup>2,\*</sup>  
and Hoang Thi Tham<sup>3</sup>

<sup>1</sup>Faculty of Information Technology, Hanoi National University of Education, Hanoi city, Vietnam;

<sup>2</sup>Faculty of Chemistry, Hanoi National University of Education, Hanoi city, Vietnam; <sup>3</sup>Department of Secondary Education, Department of Education and Training of Lao Cai, Lao Cai Province, Vietnam

\*Corresponding author: Pham Thi Binh,  
e-mail: ptbinh@hnue.edu.vn

Received January 4, 2025.

Revised January 23, 2025.

Accepted January 31, 2025.

**Abstract.** The lesson evaluation scale is an essential tool that supports teachers in analyzing and assessing lessons during professional development activities, contributing to the advancement of teachers' competencies. For STEM lessons, which are characterized by interdisciplinary integration, the use of a suitable evaluation tool is crucial. This paper presents the process of developing a STEM lesson evaluation scale and examines its reliability in assessing STEM lessons at secondary schools in Vietnam. The scale was developed based on theoretical research on STEM lessons, guidelines for implementing STEM education, professional development activities, lesson planning, and the practice of lesson evaluation in Vietnamese secondary schools. The development process includes gathering feedback from teachers regarding the unambiguity, clarity, and ease of understanding of the criteria and levels. The scale was then validated for reliability using Cronbach's Alpha and consensus level through the Intraclass Correlation Coefficient (ICC). The experiment involved 79 teachers who independently used the scale to evaluate three STEM lesson videos. The analysis results showed that the scale demonstrated high reliability (Cronbach's Alpha coefficients were all above 0.8 and one group at 0.675) and strong consensus (ICC coefficients above 0.8).

**Keywords:** scale, STEM lesson, STEM, Vietnam.

## XÂY DỰNG THANG ĐO ĐÁNH GIÁ BÀI HỌC STEM TRONG DẠY HỌC Ở TRƯỜNG PHỔ THÔNG VIỆT NAM

Kiều Phương Thùy<sup>1</sup>, Phạm Thị Bình<sup>2,\*</sup>  
và Hoàng Thị Thắm<sup>3</sup>

<sup>1</sup>Khoa Công nghệ Thông tin, Trường Đại học Sư phạm Hà Nội, thành phố Hà Nội, Việt Nam;

<sup>2</sup>Khoa Hóa học, Trường Đại học Sư phạm Hà Nội, thành phố Hà Nội, Việt Nam; <sup>3</sup>Phòng Giáo dục Trung học, Sở Giáo dục và Đào tạo Tỉnh Lào Cai, tỉnh Lào Cai, Việt Nam

\*Tác giả liên hệ: Phạm Thị Bình,  
e-mail ptbinh@hnue.edu.vn

Ngày nhận bài: 4/1/2025.

Ngày sửa bài: 23/1/2025.

Ngày nhận đăng: 31/1/2025.

**Tóm tắt.** Thang đo đánh giá bài học là công cụ quan trọng, hỗ trợ GV để phân tích, đánh giá bài học trong sinh hoạt chuyên môn nhằm phát triển chuyên môn cho giáo viên. Việc sử dụng một công cụ đánh giá phù hợp cho các bài học STEM có đặc trưng tích hợp và liên môn là rất cần thiết. Bài báo này trình bày quá trình xây dựng và phát triển thang đo đánh giá bài học STEM và kiểm tra độ tin cậy của thang đo trong việc đánh giá các bài học STEM tại trường phổ thông Việt Nam. Thang đo được xây dựng dựa trên các nghiên cứu lý thuyết về bài học STEM, hướng dẫn triển khai giáo dục STEM và sinh hoạt chuyên môn, xây dựng kế hoạch bài dạy và thực tiễn đánh giá bài học ở các trường phổ thông Việt Nam. Quá trình xây dựng và phát triển thang đo bao gồm việc thu thập ý kiến từ giáo viên về tính đơn nghĩa, độ rõ ràng và mức độ dễ hiểu của các tiêu chí và các mức độ. Thang đo sau đó được kiểm định độ tin cậy thông qua hệ số Cronbach's Alpha và mức độ đồng thuận qua hệ số tương quan nội nhóm (Intraclass Correlation Coefficient – ICC). Thử nghiệm được thực hiện với sự tham gia của 79 giáo viên, những người đã sử dụng thang đo để độc lập đánh giá 03 video bài học STEM. Kết quả phân tích cho thấy thang đo đạt độ tin cậy cao (hệ số Cronbach's Alpha đều trên 0,8 và có 1 nhóm là 0,675) và mức độ đồng thuận rất mạnh (hệ số ICC đều trên 0,8).

**Từ khóa:** thang đo, bài học STEM, STEM, Việt Nam.

## **1. Mở đầu**

Giáo dục STEM (Khoa học, Công nghệ, Kỹ thuật và Toán học) đã trở thành một xu hướng quan trọng trong hệ thống giáo dục toàn cầu, nhằm trang bị cho học sinh (HS) những kỹ năng cần thiết để đáp ứng yêu cầu của thế kỷ 21. Giáo dục STEM giúp HS phát triển năng lực giải quyết vấn đề, tư duy phản biện và sự sáng tạo của HS, đồng thời tạo điều kiện để áp dụng kiến thức liên ngành vào thực tiễn [1]-[4].

Ở Việt Nam, trong Chương trình Giáo dục phổ thông [5] và Công văn số 3089/BGDĐT-GDTrH [6], 909/BGDĐT-GDTH [7] hướng dẫn triển khai giáo dục STEM ở trường phổ thông cũng đều tiếp cận giáo dục STEM theo hướng tích hợp nhằm mục đích giúp HS áp dụng 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, qua đó phát triển năng lực và phẩm chất cho HS, khơi dậy hứng thú, đam mê học tập của HS trong các lĩnh vực STEM. Trong đó, giáo dục STEM được hướng dẫn triển khai với 3 hình thức: bài học STEM, trải nghiệm STEM và nghiên cứu khoa học kỹ thuật. Bài học STEM triển khai trong quá trình dạy học các môn học thuộc chương trình giáo dục phổ thông theo hướng tiếp cận tích hợp được xác định là hình thức tổ chức giáo dục STEM chủ yếu trong nhà trường phổ thông. Theo định hướng đó, trong những năm gần đây, nhiều nghiên cứu về giáo dục STEM ở Việt Nam tập trung vào đề xuất cách triển khai hoặc thiết kế bài học STEM trong các môn học [8]-[13].

Tuy nhiên, các nghiên cứu xây dựng công cụ đánh giá với việc triển khai giáo dục STEM rất hạn chế. Nhiều nghiên cứu ở Việt Nam chủ yếu về xây dựng khung năng lực và công cụ đánh giá năng lực dạy học STEM của giáo viên (GV), sinh viên sư phạm [14]-[19]. Các nghiên cứu về đánh giá các hình thức giáo dục STEM gần đây có các nghiên cứu của tác giả NTT Trang (2021, 2022) đề xuất công cụ đánh giá kỹ năng tổ chức dạy học STEM cho sinh viên sư phạm [20] và công cụ đánh giá năng lực dạy học STEM [21], nghiên cứu của nhóm tác giả NV Biên (2024) [22] xây dựng công cụ đánh giá kế hoạch bài dạy STEM dựa trên việc sử dụng cây quyết định; và trên thế giới, có nhóm tác giả Gökçe, H. (2022) xây dựng công cụ đánh giá câu lạc bộ STEM [23].

Việc xây dựng công cụ đánh giá bài học STEM triển khai tại Việt Nam (được hiểu là các bài học thực hiện trên lớp học, không chỉ dừng lại ở kế hoạch bài dạy) hiện đang là một khoảng trống nghiên cứu cần được quan tâm. Một công cụ đánh giá đáng tin cậy và phù hợp với bối cảnh giáo dục Việt Nam sẽ đóng vai trò quan trọng trong việc hỗ trợ GV định hướng và tự đánh giá quá trình thiết kế, tổ chức bài học STEM. Bên cạnh đó, công cụ này còn có giá trị trong việc phân tích, thảo luận tại các buổi sinh hoạt chuyên môn nhằm phát triển năng lực nghề nghiệp cho GV, cũng như hỗ trợ tổ chức các cuộc thi bài học STEM, từ đó góp phần nâng cao chất lượng giáo dục STEM trong nhà trường.

Mục tiêu của nghiên cứu này là phát triển một thang đo đánh giá bài học STEM và kiểm tra độ tin cậy của nó trong việc đánh giá các bài học STEM ở trường phổ thông Việt Nam. Cụ thể, nghiên cứu tập trung vào trả lời các câu hỏi sau: (1) Thang đo đánh giá bài học STEM được xác định dựa trên cơ sở lý thuyết và thực tiễn nào?; (2) Thang đo đánh giá bài học STEM gồm những tiêu chí, mức độ nào?; (3) Thang đo đánh giá bài học STEM xây dựng có tin cậy không?

Nghiên cứu này đóng góp vào khung lý thuyết giáo dục STEM các tiêu chí cụ thể giúp làm rõ hơn các yếu tố cốt lõi trong bài học STEM và định hướng phát triển các công cụ và phương pháp đánh giá bài học STEM trong các nghiên cứu và ứng dụng tương lai. Đồng thời nghiên cứu cung cấp công cụ khoa học để GV phân tích, đánh giá và cải tiến bài học STEM trong thực tiễn dạy học ở trường phổ thông.

Phần tiếp theo của bài báo sẽ trình bày chi tiết về phương pháp nghiên cứu, quy trình phát triển thang đo và kết quả, thảo luận kiểm tra độ tin cậy của thang đo.

## 2. Nội dung nghiên cứu

### 2.1. Phương pháp nghiên cứu

Thang đo đánh giá được nhóm tác giả tự xây dựng dựa trên các nghiên cứu lí thuyết về giáo dục STEM cũng như thực tiễn triển khai giáo dục STEM ở trường phổ thông Việt Nam, hướng dẫn thiết kế kế hoạch bài dạy phát triển năng lực, hướng dẫn sinh hoạt chuyên môn từ Bộ Giáo dục và Đào tạo Việt Nam [6]-[7].

Để kiểm tra sự phù hợp độ tin cậy của thang đo này, chúng tôi tiến hành theo 2 phương pháp, thứ tự như sau:

Thứ nhất, lấy ý kiến của GV về thang đo. Mục đích để đánh giá tính đơn nghĩa, rõ ràng, dễ hiểu, sự phù hợp với thực tiễn, khả thi để đánh giá của các tiêu chí và các mức độ. Cụ thể đã lấy ý kiến phát biểu trực tiếp với 50 GV trong một hội nghị ở Lào Cai vào tháng 3/2024.

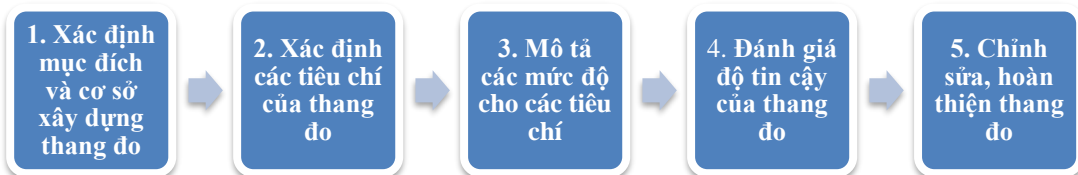
Thứ hai, phương pháp thử nghiệm được sử dụng để đánh giá độ tin cậy trong quá trình chuẩn hoá công cụ. 79 GV từ các lớp tập huấn về giáo dục STEM tại Trường Đại học Sư phạm Hà Nội đã sử dụng thang đo đánh giá độc lập 03 video bài học STEM ở trường phổ thông. Các video được quay lại từ các bài học STEM đã thực hiện trực tiếp ở trường phổ thông, đảm bảo chất lượng về kênh hình và tiếng để người xem quan sát được các hoạt động dạy và học.

Các GV tham gia thử nghiệm sử dụng thang đo đề được trang bị lí thuyết, hướng dẫn về thiết kế và tổ chức bài học STEM; nhận thang đo và được giới thiệu về các tiêu chí, mức độ theo công cụ đánh giá. Sau đó GV sử dụng công cụ để độc lập đánh giá 1 video bài học STEM.

### 2.2. Kết quả nghiên cứu

#### 2.2.1. Tiến trình phát triển thang đo

Theo quy trình đánh giá năng lực nói chung, thang đo đánh giá bài học STEM được đề xuất xây dựng theo 4 bước như sau:



Hình 1. Quy trình xây dựng thang đo đánh giá bài học STEM

#### \* Bước 1. Xác định mục đích và cơ sở xây dựng thang đo

Mục đích xây dựng thang đo là đánh giá bài học STEM được triển khai ở trường phổ thông Việt Nam. Bài học STEM ở đây được hiểu là bài học thực hiện trong chương trình của các môn học theo hướng tiếp cận tích hợp các lĩnh vực STEM để giải quyết vấn đề. Trong bài học đó HS được yêu cầu tìm các giải pháp để giải quyết vấn đề thực tiễn (các vấn đề thực tiễn về đời sống xã hội, khoa học, công nghệ), chiếm lĩnh kiến thức, đáp ứng yêu cầu cần đạt của bài học.

Các căn cứ được xác định để xây dựng thang đo bài học STEM trong nghiên cứu này gồm:

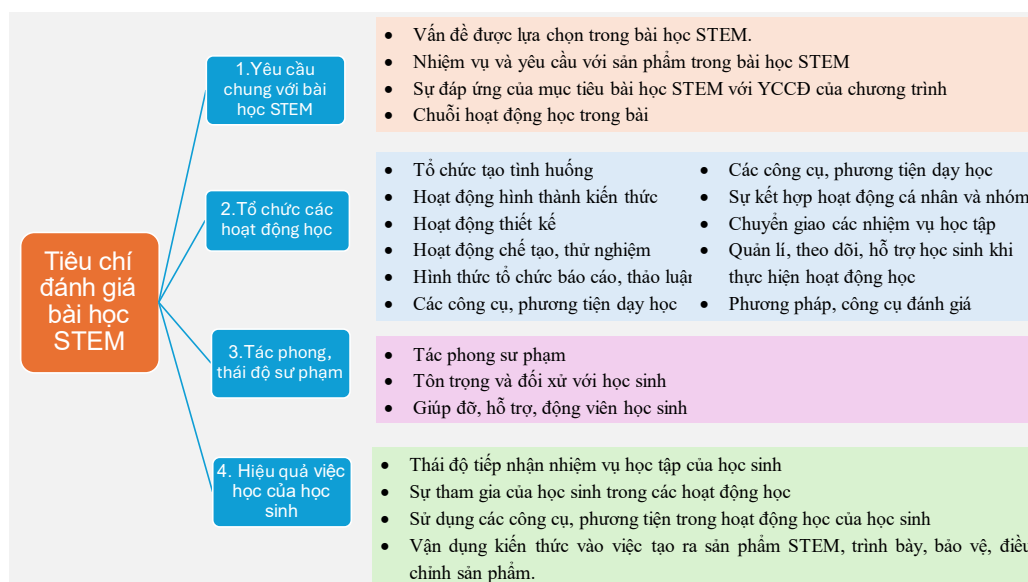
- Lí thuyết về giáo dục STEM [24].
- Đặc điểm và tiến trình thực hiện bài học STEM theo quy trình thiết kế kĩ thuật được hướng dẫn trong công văn 3089/BGDĐT-GDTrH [6] và công văn 909/ BGDĐT-GDTH [7].
- Hướng dẫn sinh hoạt chuyên môn theo công văn 5555/BGDĐT-GDTrH, cụ thể là hướng dẫn nội dung sinh hoạt tổ/nhóm chuyên môn về: Thiết kế tiến trình dạy học; Tổ chức dạy học và dự giờ; Phân tích, rút kinh nghiệm bài học [25].

- Hướng dẫn xây dựng kế hoạch bài dạy phát triển năng lực HS theo công văn 5512/BGDĐT-GDTrH [26].

Việc xây dựng thang đo đánh giá bài học STEM cũng căn cứ vào việc tham khảo nhiều phiếu đánh giá giờ dạy của các nhà trường, phòng giáo dục và đào tạo, sở giáo dục và đào tạo các địa phương ở Việt Nam và dựa vào kinh nghiệm của nhóm tác giả đều là các giảng viên bộ môn Phương pháp Giảng dạy có kinh nghiệm của Trường Đại học Sư phạm Hà Nội, đã tham gia viết sách bài học STEM, các tài liệu hướng dẫn, tập huấn GV về triển khai hoạt động giáo dục STEM, dạy học phát triển năng lực cũng như các hoạt động sinh hoạt chuyên môn nghiên cứu bài học với dạng bài học STEM ở nhiều trường phổ thông.

### \* Bước 2. Xác định tiêu chí của thang đo

Theo mục đích xây dựng thang đo và những căn cứ xác định ở trên, chúng tôi đã phát triển công cụ đánh giá bài học STEM dưới dạng thang đo, gồm các tiêu chí đánh giá và 5 mức độ đánh giá. Theo các mức độ này, các nhà trường có thể quy đổi thành điểm để sử dụng đánh giá bài học STEM. Thang đo được đề xuất gồm 4 nhóm, nhóm 1 được đề xuất dựa trên lí thuyết về giáo dục STEM và đặc điểm của bài học STEM [6]-[7], các nhóm 2, 3, 4 được phát triển theo cách tiếp cận phân tích, đánh giá một bài học trong công văn 5555/BGDĐT-GDTrH [25] hướng dẫn sinh hoạt chuyên môn ở Việt Nam cũng như trong các phiếu đánh giá bài học mà các nhà trường thường sử dụng trong thực tiễn. Cụ thể các nhóm tiêu chí như sau:



Hình 2: Cấu trúc thang đo đánh giá bài học STEM

### \* Bước 3. Mô tả các mức độ cho các tiêu chí

Mỗi tiêu chí được mô tả theo các yêu cầu thể hiện mức độ hoàn thiện từ thấp đến cao của vấn đề đề cập trong tiêu chí. Các yêu cầu được mô tả này căn cứ theo: hướng dẫn thực hiện về nội dung đối với bài học STEM; bản chất của giáo dục STEM; các yêu cầu chung với việc thiết kế kế hoạch bài dạy trong các môn học; mục tiêu của từng hoạt động trong tiến trình thực hiện bài học STEM theo quy trình thiết kế kĩ thuật; hướng dẫn tổ chức thực hiện hoạt động học, sử dụng phương pháp công cụ đánh giá trong dạy học theo mục tiêu phát triển phẩm chất năng lực HS; sự kết hợp các hình thức học tập trong các văn bản hướng dẫn của Bộ Giáo dục và Đào tạo Việt Nam và lí thuyết dạy học phát triển năng lực.

Khi đánh giá sẽ theo 5 mức từ thấp đến cao: 1. Rất không phù hợp; 2. Không phù hợp; 3. Phù hợp một phần; 4. Phù hợp; 5. Rất phù hợp. Trong đó mức 5 là mức đạt được tất cả các mô tả trong tiêu chí.

#### **\* Bước 4. Đánh giá độ tin cậy của thang đo**

Đánh giá sự phù hợp của các tiêu chí, mức độ rõ ràng, dễ hiểu của mô tả các tiêu chí thông qua việc xin ý kiến của GV phổ thông. Các GV được hỏi đều là GV đã hiểu và được trải nghiệm về bài học STEM. Các GV này sẽ đọc nội dung các tiêu chí và phản hồi về mức độ rõ ràng, dễ hiểu của từng tiêu chí. Căn cứ vào các phản hồi đó, các tiêu chí được sửa lại cho phù hợp.

Đánh giá độ tin cậy của công cụ bằng phương pháp thử nghiệm. Công cụ được giới thiệu cho GV, sau đó sử dụng để đánh giá bài dạy STEM ở trường phổ thông. Tính hệ số Cronbach's Alpha để và tính hệ số tương quan ICC (Intraclass Correlation Coefficient) về sự đồng thuận của những người sử dụng với công cụ từ đó khẳng định độ tin cậy của công cụ.

#### **\* Bước 5. Chỉnh sửa, hoàn thiện thang đo**

Từ kết quả đánh giá độ tin cậy của thang đo, hoàn thiện thang đo.

### **2.2.2. Kết quả đánh giá độ tin cậy của thang đo**

Thang đo được đề xuất đã được lấy ý kiến từ 50 GV phổ thông trong một hội nghị trực tiếp tại Lào Cai năm 2024. Những GV tham gia khảo sát đều đã được tập huấn và có kinh nghiệm triển khai bài học STEM tại trường phổ thông. Trong quá trình thu thập ý kiến, thang đo được gửi trực tiếp tới các GV, kèm theo phân giải thích chi tiết về mục đích, căn cứ phát triển và cách thức sử dụng thang đo. GV phản hồi trực tiếp về các tiêu chí và mô tả trong thang đo, tập trung làm rõ ý nghĩa và nội dung. Trên cơ sở các phản hồi này, chúng tôi đã điều chỉnh lại cách diễn đạt để đảm bảo tính dễ hiểu và chính xác. Đồng thời, một số ý kiến bổ sung đã được tích hợp vào thang đo, bao gồm các tiêu chí liên quan đến sự an toàn khi sử dụng phương tiện và công cụ, khả năng quản lý và điều hành hoạt động học tập đảm bảo thời gian, xử lý tình huống sư phạm phù hợp, cũng như việc hoàn thành tốt nhiệm vụ học tập của HS trong các hoạt động học tập.

Sau khi được chỉnh sửa lần thứ nhất dựa trên ý kiến góp ý từ GV, thang đo đã được mang đi thử nghiệm để đánh giá tính hiệu quả và phù hợp đối với các bài học STEM. Quá trình thử nghiệm được thực hiện trên 3 video bài học STEM đã thực hiện tại trường phổ thông, bao gồm: (1) Thuộc bảo vệ thực vật có nguồn gốc sinh học (Lớp 11); (2) Pha chế nước muối sinh lí (Lớp 8); và (3) Quạt Mandala (Lớp 6). Người tham gia thử nghiệm thang đo là 79 GV đã tham gia các lớp tập huấn "Bồi dưỡng GV tổ chức hoạt động STEM/STEAM" do Trường Đại học Sư phạm Hà Nội tổ chức vào năm 2024 (sau đây gọi tắt là các chuyên gia). Các GV này đều được trang bị kiến thức về giáo dục STEM và hình thức bài học STEM. Ba video được lựa chọn để thử nghiệm đều đảm bảo chất lượng cao về hình ảnh và âm thanh, nhằm đảm bảo quan sát đầy đủ các hoạt động học tập của HS. Các bài học STEM này được thiết kế và triển khai theo đúng hướng dẫn về tổ chức và triển khai bài học STEM tại Việt Nam.

Mỗi video bài học sẽ được các chuyên gia đánh giá độc lập theo cùng 1 thang đo, thang đo chia theo 4 nhóm tiêu chí (S1-S4). Chuyên gia bao gồm GV thuộc các lĩnh vực, chuyên môn có liên quan. Kết quả thu được là 79 phiếu đánh giá trong đó 22 phiếu cho bài 1, 34 phiếu cho bài 2 và 23 phiếu cho bài 3.

Bước đầu, toàn bộ kết quả tính toán sẽ được kiểm định độ tin cậy Cronbach's Alpha, sau đó kết quả từng bài học sẽ được tính toán giá trị trung bình, độ lệch chuẩn và độ đồng thuận.

Kết quả kiểm định Cronbach's Alpha theo từng nhóm tiêu chí thể hiện trong Bảng 1.

**Bảng 1. Kết quả kiểm định Cronbach's Alpha**

Group	Cronbach's Alpha	Hệ số tương quan biến-Tổng nhỏ nhất (CITC)
S1 (S1.1-S1.4)	0,806	0,499
S2 (S2.1-S2.10)	0,863	0,378
S3 (S3.1-S3.3)	0,810	0,611
S4 (S4.1-S4.4)	0,675	0,329

Kết quả cho thấy 3 nhóm tiêu chí có kết quả Cronbach's Alpha > 0,7, ngưỡng coi là đạt yêu cầu theo Nunnally, J. (1978) [27], riêng đối với tiêu chí S4, mức 0,675 > 0,6 cũng được coi là chấp nhận được khi đây là nghiên cứu thăm dò với đối tượng khảo sát nhỏ [28]. Ngoài ra, hệ số tương quan giữa biến và tổng tất cả đều > 0,3 [29], đảm bảo các biến quan sát đều có ý nghĩa đóng góp đáng kể trong giá trị biến tổng, cần được giữ lại trong tính toán.

Điểm số trung bình (Mean), độ lệch chuẩn (SD, Std. Deviation) và hệ số dao động dữ liệu (Coefficient of Variation: CV) cho từng tiêu chí được thể hiện trong Bảng 2.

**Bảng 2. Giá trị thống kê cho các bài học STEM**

Item	Bài học STEM 1 (n = 22)			Bài học STEM 2 (n = 34)			Bài học STEM 3 (n = 23)		
	Mean	Std. Deviation	CV	Mean	Std. Deviation	CV	Mean	Std. Deviation	CV
S1.1	4,730	0,550	11,64%	4,790	0,479	9,98%	4,220	0,671	15,92%
S1.2	4,320	0,780	18,06%	4,620	0,551	11,94%	3,901	0,668	17,08%
S1.3	4,410	0,734	16,65%	4,650	0,646	13,90%	4,220	0,795	18,86%
S1.4	4,270	0,767	17,96%	4,470	0,662	14,81%	3,870	0,757	19,56%
S2.1	2,136	0,351	16,44%	2,324	0,346	14,87%	1,978	0,281	14,21%
S2.2	2,159	0,284	13,15%	2,250	0,354	15,71%	1,826	0,357	19,55%
S2.3	2,159	0,419	19,42%	2,309	0,276	11,94%	1,717	0,295	17,17%
S2.4	2,136	0,384	17,96%	2,294	0,279	12,14%	2,022	0,384	18,98%
S2.5	2,159	0,358	16,59%	2,250	0,431	19,15%	1,761	0,395	22,44%
S2.6	4,410	0,666	15,11%	4,440	0,927	20,88%	4,480	0,511	11,41%
S2.7	4,410	0,734	16,65%	4,500	0,615	13,68%	4,000	0,739	18,46%
S2.8	4,270	0,703	16,44%	4,680	0,535	11,44%	3,870	0,548	14,16%
S2.9	4,270	0,767	17,96%	4,500	0,615	13,68%	4,480	0,665	14,86%
S2.10	4,360	0,790	18,09%	4,470	0,748	16,73%	3,780	0,671	17,75%
S3.1	4,640	0,658	14,19%	4,910	0,288	5,86%	4,740	0,449	9,47%
S3.2	4,770	0,528	11,07%	4,910	0,288	5,86%	4,780	0,422	8,82%
S3.3	4,590	0,666	14,51%	4,710	0,462	9,83%	4,700	0,470	10,02%
S4.1	4,410	0,590	13,39%	4,760	0,431	9,04%	3,960	0,638	16,13%
S4.2	4,680	0,477	10,18%	4,710	0,524	11,13%	4,610	0,499	10,83%
S4.3	9,090	1,019	11,21%	9,530	0,861	9,04%	8,170	1,029	12,59%
S4.4	9,090	1,342	14,76%	9,240	0,987	10,68%	8,090	1,411	17,45%

Kết quả Bảng 2 cho thấy các câu hỏi với thang điểm 5 đều có độ lệch chuẩn nhỏ hơn 1, độ dao động dữ liệu < 20% cho thấy kết quả đánh giá của chuyên gia là tập trung, ít có sự chênh lệch.

Thông kê ban đầu về phân bố điểm chấm của các chuyên gia theo từng kế hoạch bài dạy thể hiện trong các Bảng 3, 4 và 5.

**Bảng 3. Phân bố điểm bài học 1 (Bài Thuốc bảo vệ thực vật) (n = 22)**

Tiêu chí \ Điểm	2	3	4	5	6	8	10
S1.1		1	4	17			
S1.2		4	7	11			
S1.3		3	7	12			
S1.4	1	1	11	9			
S2.1		3	10	9			
S2.2		1	13	8			
S2.3	1	2	8	11			
S2.4	1	1	11	9			
S2.5		3	9	10			
S2.6		2	9	11			
S2.7		3	7	12			
S2.8		3	10	9			
S2.9	1	1	11	9			
S2.10	1	1	9	11			
S3.1		2	4	16			
S3.2		1	3	18			
S3.3		2	5	15			
S4.1		1	11	10			
S4.2			7	15			
S4.3						10	12
S4.4					2	6	14

**Bảng 4. Phân bố điểm bài 2 (Nước muối sinh lý) (n = 34)**

Tiêu chí \ Điểm	1	2	3	4	5	8	10
S1.1			1	5	28		
S1.2			1	11	22		
S1.3			3	6	25		
S1.4		1		15	18		
S2.1		1	1	7	25		
S2.2			4	9	21		
S2.3			1	11	22		

*Xây dựng thang đo đánh giá bài học STEM trong dạy học ở trường phổ thông Việt Nam*

S2.4			1	12	21		
S2.5		1	2	9	22		
S2.6	1	1	1	10	21		
S2.7			2	13	19		
S2.8			1	9	24		
S2.9			2	13	19		
S2.10		1	2	11	20		
S3.1				3	31		
S3.2				3	31		
S3.3				10	24		
S4.1				8	26		
S4.2			1	8	25		
S4.3						8	26
S4.4						13	21

**Bảng 5. Phân bố điểm bài 3 (Bài Quạt Madala) (n = 23)**

Tiêu chí	Điểm									
	1	2	3	4	5	6	8	10		
S1.1			3	12	8					
S1.2			6	13	4					
S1.3			5	8	10					
S1.4			8	10	5					
S2.1			4	16	3					
S2.2			11	9	3					
S2.3			14	8	1					
S2.4			6	10	7					
S2.5		2	9	10	2					
S2.6				12	11					
S2.7		1	3	14	5					
S2.8			5	16	2					
S2.9			2	8	13					
S2.10		1	5	15	2					
S3.1				6	17					
S3.2				5	18					
S3.3				7	16					
S4.1			5	14	4					
S4.2				9	14					
S4.3						2	17		4	
S4.4						2	12		6	



Dữ liệu từ các Bảng 3, 4, 5 cho thấy kết quả chấm điểm của các chuyên gia khá tập trung, không phân tán trên toàn bộ phổ điểm.

Để khẳng định sự đồng thuận trong đánh giá của chuyên gia đối với từng video giờ dạy, nghiên cứu sử dụng hệ số tương quan ICC (Intraclass Correlation Coefficient). Hệ số tương quan ICC được chọn thay vì sử dụng hệ số tương quan Kendall's W vì dữ liệu trong bài là chấm điểm - một dạng dữ liệu liên tục chứ không phải dữ liệu xếp hạng. Kết quả kiểm tra sự đồng thuận trong từng video giờ dạy thể hiện trong bảng 6:

**Bảng 6. Kết quả tính hệ số ICC các đánh giá của 3 bài học STEM**

Bài học	N	ICC	P	Mức độ đồng thuận
1	22	0,925	0,000	Rất mạnh
2	34	0,941	0,000	Rất mạnh
3	23	0,898	0,000	Rất mạnh

Kết quả tại Bảng 6 cho thấy, trong cả 3 video giờ dạy, kết quả đánh giá của các chuyên gia đều có sự đồng thuận > 0,8, là mức đồng thuận rất mạnh [30]. Điều này phù hợp với quan sát trực quan ban đầu về tỉ lệ phân bố điểm các chuyên gia tại các Bảng 4, 5, 6 ở trên. Kết quả này khẳng định thang đo đánh giá tổ chức dạy học bài học STEM đảm bảo độ tin cậy và thống nhất trong quá trình sử dụng. Có thể áp dụng trong nhiều giờ dạy khác nhau với chuyên gia đánh giá khác nhau.

### 2.2.3. Nội dung thang đo sau khi chỉnh sửa, hoàn thiện

Kết quả thử nghiệm cho thấy độ tin cậy của thang đo là đảm bảo, do đó nhóm tác giả giữ nguyên nội dung của thang đo sau khi chỉnh sửa theo ý kiến của GV. Chi tiết nội dung thang đo được trình bày trong Bảng 7.

**Bảng 7. Thang đo đánh giá bài học STEM sau chuẩn hóa**

Tiêu chí	Mức độ				
	1	2	3	4	5
<b>S1. Yêu cầu chung với bài học STEM</b>					
S1.1. Vấn đề cần giải quyết có tính thực tiễn, phù hợp với nội dung bài học, đối tượng HS và điều kiện của nhà trường.					
S1.2. Nhiệm vụ và yêu cầu với sản phẩm STEM cụ thể, rõ ràng, phù hợp với mục tiêu bài học và thể hiện được yêu cầu tích hợp các lĩnh vực STEM.					
S1.3. Mục tiêu bài học STEM đáp ứng được yêu cầu cần đạt của bài học của môn học chủ đạo và theo mục tiêu phát triển phẩm chất, năng lực cho HS.					
S1.4. Chuỗi các hoạt động học logic, khả thi, phù hợp theo tiến trình tổ chức hoạt động STEM và đối tượng HS, phân bố thời gian phù hợp.					
<b>S2. Tổ chức các hoạt động học</b>					
S2.1. Tổ chức tạo tình huống phù hợp, logic với vấn đề cần giải quyết, hấp dẫn; nêu được vấn đề cần giải quyết, nhiệm vụ, các yêu cầu cụ thể, rõ ràng với sản phẩm STEM.					

S2.2. Hoạt động hình thành kiến thức được tổ chức theo phương pháp, kỹ thuật dạy học tích cực trong đó HS là người tìm tòi, khám phá kiến thức; các kiến thức trong bài học chính xác.					
S2.3. Hoạt động thiết kế được tổ chức phù hợp, tích cực theo mục đích HS vận dụng kiến thức, tự đề xuất giải pháp phù hợp với nhiệm vụ và yêu cầu đặt ra với sản phẩm STEM.					
S2.4. Hoạt động chế tạo, thử nghiệm được tổ chức phù hợp, tích cực để HS tự lực ở mức độ phù hợp tạo ra sản phẩm STEM theo yêu cầu.					
S2.5. Hình thức tổ chức báo cáo, thảo luận phù hợp với nhiệm vụ, sản phẩm STEM và phù hợp để đánh giá được mức độ hiểu và vận dụng kiến thức, kỹ năng vào tạo ra sản phẩm STEM, giải quyết vấn đề của HS.					
S2.6. Các công cụ, phương tiện dạy học sử dụng phù hợp, đủ đơn giản, đảm bảo an toàn, tính khoa học, hiệu quả, thẩm mỹ,					
S2.7. Có sự kết hợp phù hợp, hiệu quả giữa làm việc cá nhân và làm việc nhóm trong tổ chức các hoạt động học của HS.					
S2.8. Chuyển giao các nhiệm vụ học tập rõ ràng, hấp dẫn.					
S2.9. Quản lý, theo dõi và có hỗ trợ phù hợp, kịp thời khi HS thực hiện hoạt động học để đảm bảo thời gian và hiệu quả.					
S2.10. Phương pháp, công cụ đánh giá sử dụng trong bài học phù hợp, hiệu quả để đánh giá được theo mục tiêu bài học và được đa số HS.					
<b>S3. Tác phong, thái độ sư phạm</b>					
S3.1. Tác phong sư phạm chuẩn mực, gần gũi, thân thiện với HS, tạo không khí vui vẻ, nhẹ nhàng, thoải mái, tích cực.					
S3.2. Tôn trọng và đối xử công bằng với HS, xử lý các tình huống sư phạm phù hợp, tích cực.					
S3.3. Giúp đỡ, hỗ trợ HS có khó khăn trong học tập kịp thời, đồng viên phù hợp để mỗi HS đều được phát triển năng lực học tập.					
<b>S4. Hiệu quả hoạt động học</b>					
S4.1. HS hào hứng, tích cực tiếp nhận nhiệm vụ học tập.					
S4.2. HS tích cực, chủ động thực hiện các hoạt động học tập, có tương tác tích cực với các bạn và GV, hoàn thành tốt nhiệm vụ học tập.					

S4.3. HS sử dụng các công cụ, phương tiện thực hiện các hoạt động học an toàn.					
S4.4. HS vận dụng được kiến thức và tạo ra được sản phẩm STEM; trình bày, bảo vệ, có điều chỉnh phù hợp với sản phẩm STEM làm ra.					

Năm mức độ của các tiêu chí trong thang đo là: 1. Rất không phù hợp; 2. Không phù hợp; 3. Phù hợp một phần; 4. Phù hợp; 5. Rất phù hợp.

### 2.3. Thảo luận

Kết quả nghiên cứu đã xây dựng thành công một thang đo đánh giá bài học STEM, đáp ứng tính phù hợp với bối cảnh giáo dục phổ thông tại Việt Nam. Thang đo được phát triển dựa trên nền tảng lý thuyết về giáo dục STEM, các định hướng chính sách từ Bộ Giáo dục và Đào tạo, cũng như thực tiễn triển khai tại các trường phổ thông. Kết quả nghiên cứu không chỉ đóng góp vào khung lý thuyết hiện có mà còn cung cấp một công cụ đánh giá cụ thể, có tính khả thi cao, dễ sử dụng, nhằm hỗ trợ GV trong việc thiết kế, tổ chức và đánh giá hiệu quả các bài học STEM.

Việc kiểm định độ tin cậy với hệ số Cronbach's Alpha đạt trên ngưỡng 0,7 ở hầu hết các nhóm tiêu chí đã khẳng định tính nhất quán nội tại của thang đo, phù hợp với khuyến nghị của Nunnally, J. (1978) [27], nhóm tiêu chí S4 có độ tin cậy cũng có thể chấp nhận được theo khuyến nghị của Hair JF, Black (2023) [28] về ngưỡng độ tin cậy. Kết quả này có thể phản ánh sự phức tạp của các tiêu chí liên quan đến ứng dụng thực tiễn STEM trong giảng dạy, vốn cũng là một thách thức được đề cập trong nghiên cứu của Bybee (2013) [1].

Phân tích số liệu từ phiếu đánh giá GV cho thấy thang đo có khả năng phân biệt rõ ràng giữa các bài học STEM. Điều này góp phần khẳng định khả năng ứng dụng của công cụ, tương đồng với nghiên cứu của Honey, Pearson, & Schweingruber (2014) [31] về tầm quan trọng của các công cụ đánh giá thực tiễn trong giáo dục STEM. Tuy nhiên, mẫu thử nghiệm hạn chế cả về số lượng GV và bài học dẫn đến hạn chế nhất định trong khả năng tổng quát hóa.

### 3. Kết luận

Nghiên cứu này đã thành công trong việc phát triển và kiểm định bước đầu một thang đo đánh giá bài học STEM, đáp ứng yêu cầu về lý thuyết giáo dục STEM, hướng dẫn triển khai thực hiện giáo dục STEM trong nhà trường phổ thông của Bộ Giáo dục và Đào tạo, cũng như thực tiễn giáo dục phổ thông ở Việt Nam. Thang đo gồm 4 nhóm với 21 tiêu chí: S1. Yêu cầu chung với bài học STEM (gồm 4 tiêu chí); S2. Tổ chức các hoạt động học (gồm 10 tiêu chí); S3. Tác phong, thái độ sư phạm (gồm 3 tiêu chí); S4. Hiệu quả hoạt động học (gồm 4 tiêu chí). Kết quả nghiên cứu này đóng góp quan trọng vào việc xây dựng công cụ đánh giá bài học STEM, hỗ trợ GV trong việc đánh giá và cải tiến các bài học STEM, đồng thời thúc đẩy sự phát triển của giáo dục STEM trong nhà trường phổ thông tại Việt Nam.

Thang đo có độ tin cậy cao và khả năng ứng dụng thực tiễn, nhưng vẫn cần được mở rộng thử nghiệm với nhiều nhóm đối tượng và bối cảnh khác nhau. Việc tối ưu hóa các tiêu chí, đặc biệt là nhóm tiêu chí S4, là hướng đi quan trọng nhằm nâng cao chất lượng công cụ.

Trong tương lai, nghiên cứu này có thể mở ra cơ hội áp dụng thang đo vào việc thiết kế và triển khai các chương trình đào tạo GV STEM, cũng như góp phần xây dựng các tiêu chuẩn quốc gia về đánh giá bài học STEM. Những nỗ lực này sẽ góp phần nâng cao chất lượng dạy học, thúc đẩy khả năng tư duy sáng tạo và năng lực giải quyết vấn đề của HS trong bối cảnh giáo dục hiện đại.

## TÀI LIỆU THAM KHẢO

- [1] Bybee RW, (2013). *The Case for STEM Education: Challenges and Opportunities*. NSTA Press, p. 1-145. <https://www.nsta.org/case-stem-education>.
- [2] Honey M & Kanter DE, (2013). *Design, Make, Play: Growing the Next Generation of STEM Innovators*. Routledge, p. 1-300
- [3] English LD, (2016). STEM education K-12: Perspectives on integration. *International Journal of STEM Education*, 3(1), 1-8. <https://doi.org/10.1186/s40594-016-0036-1>
- [4] Voogt J & Pareja Roblin N, (2012). 21st-century skills: Discussion paper. *European Journal of Education*, 48(2), 255-272. <https://doi.org/10.1111/ejed.12016>.
- [5] Bộ Giáo dục & Đào tạo, (2018). Chương trình Giáo dục phổ thông môn Hóa học (ban hành kèm theo Thông tư số 32/2018/TTBGDDT ngày 26/12/2018 của Bộ trưởng Bộ GD-ĐT).
- [6] Bộ Giáo dục & Đào tạo, (2020). Công văn số 3089/BGDĐT-GDTrH ngày 14/8/2020 về việc *Triển khai thực hiện giáo dục STEM trong giáo dục trung học*.
- [7] Bộ Giáo dục & Đào tạo, (2023). Công văn số 909/BGDĐT-GDTH ngày 8/3/2023 về việc *Triển khai thực hiện giáo dục STEM trong giáo dục trung học*.
- [8] NM Đức & ĐT Ngoan, (2019). Thiết kế chủ đề “Pin chanh” (Chương trình Hóa học Vô cơ lớp 12) theo định hướng giáo dục STEM. *Tạp chí Giáo dục*, (số Đặc biệt tháng 4/2019), 214-221.
- [9] NT Hằng, (2020). Thiết kế và tổ chức hoạt động học trải nghiệm trong dạy học chủ đề “Sinh trưởng và phát triển ở động vật” (Sinh học 11) theo định hướng giáo dục STEM. *Tạp chí Giáo dục*, (488), 24-30.
- [10] PT Bình, NT Thủy & NT Sông Hương, (2021). Thiết kế và tổ chức bài học STEM khi dạy học các bài về dung dịch, độ tan trong môn Hóa học lớp 8. *Tạp chí Khoa học Trường Đại học Sư phạm Hà Nội*. DOI: 10.18173/2354-1075.2021-0208 2021, 66(4E), 283-294.
- [11] ĐH Ngọc, LH Hoàng & TT Ninh, (2021). Phát triển năng lực tìm hiểu thế giới tự nhiên dưới góc độ hóa học thông qua dạy học STEM phần Phi kim Hóa học 11. *Tạp chí Khoa học, Trường Đại học Sư phạm Hà Nội*, 66(4E), 34-45.
- [12] ĐTT Thư & PTB Đào, (2021). Phát triển năng lực vận dụng kiến thức, kỹ năng cho học sinh trung học phổ thông qua dạy học chủ đề STEM tích hợp nội dung giáo dục bảo vệ môi trường trong môn Hóa học. *Tạp chí Khoa học Trường Đại học Sư phạm Hà Nội*, 66(4E), 271-282.
- [13] ĐV Sơn, NT Linh & PT Bình, (2024). Tổ chức hoạt động giáo dục STEM trong dạy học mạch nội dung “Chất có ở xung quanh ta” (Khoa học Tự nhiên 6) nhằm phát triển năng lực vận dụng kiến thức, kỹ năng cho học sinh. *Tạp chí Giáo dục*, 24(2), 20-26. ISSN: 2354-0753.
- [14] NT Thủy, NV Biên & DX Quý, (2019). Đề xuất mô hình bồi dưỡng năng lực dạy học tích hợp STEM cho giáo viên khoa học tự nhiên theo hướng nghiên cứu bài học. *Kỷ yếu Hội thảo quốc tế lần thứ nhất về đổi mới đào tạo giáo viên*, 107-115.
- [15] HT Chien, (2019). Designing a competencies framework for STEM teaching for pre-teachers of chemistry in the University of Education for meeting the new demands of current teacher training. *Tạp chí Khoa học Giáo dục và Công nghệ*, 8(2), 89-9. <https://doi.org/10.25073/0866-773X/306>.
- [16] NT Thùy Trang, ĐT Oanh, PT Bình, (2020). Đề xuất khung năng lực dạy học tích hợp STEM cho sinh viên Sư phạm Hóa học. *Tạp chí Khoa học Trường Đại học Sư phạm Hà Nội*, 65(4), 177-184.
- [17] NTT Trang, ĐT Oanh, PT Bình, TT Ninh, LV Dung, MT Hùng Anh & NM Đức, (2021). Practical investigation of STEM teaching competence of pre-service chemistry teachers in Vietnam. *Journal of Physics: Conference Series*, 35(2021)012069, 1-11. DOI:10.1088/1742-6596/1835/1/012069.

- [18] TT Thùy, HT Thu Trà, ĐP Anh & PT Thủy, (2022). Phát triển công cụ tự đánh giá năng lực STEAM của giáo viên phổ thông. *Tạp chí Khoa học Giáo dục Việt Nam*, 18(10), 21-26.
- [19] NC Thắng, TL Huyền, TH Phương & NB Hậu, (2024). Đề xuất khung năng lực dạy học STEM của sinh viên ngành sư phạm tin học. *Tạp chí Giáo dục*, 24(19), 25-29.
- [20] NTT Trang, (2021). Công cụ đánh giá kỹ năng tổ chức dạy học STEM cho sinh viên Sư phạm Hóa học thông qua dạy học vi mô. *Tạp chí Khoa học Giáo dục Việt Nam*, (46), 13-19.
- [21] NTT Trang, (2021). Thiết kế bộ công cụ đánh giá năng lực dạy học STEM cho sinh viên Sư phạm Hóa học. *Tạp chí Khoa học - Trường Đại học Sư phạm, Đại học Huế*, 58(2), 103-112.
- [22] NV Bien, NTT Thủy & NT Khuyên, (2024). Xây dựng rubrics đánh giá kế hoạch bài dạy STEM của giáo viên và sinh viên sư phạm. *Journal of Educational Research*, 15(2), 123-139.
- [23] Gökçe H, (2022). STEM club evaluation scale: Validity and reliability study. *International Journal of STEM Education*, 7(2), 145-160.
- [24] NV Biên, TD Hải (đồng Chủ biên), TM Đức, NV Hạnh, CC Thơ, NA Thuần, ĐV Thược & TB Trình, (2019). *Giáo dục STEM trong nhà trường phổ thông*, Nhà xuất bản Giáo dục Việt Nam.
- [25] Bộ Giáo dục & Đào tạo, (2014). Công văn số 5555/BGDĐT-GDTrH ngày 8/10/2014 về việc *Hướng dẫn sinh hoạt chuyên môn về đổi mới phương pháp dạy học và kiểm tra, đánh giá; tổ chức và quản lý các hoạt động chuyên môn của trường trung học/trung tâm giáo dục thường xuyên qua mạng*.
- [26] Bộ Giáo dục & Đào tạo, (2012). Công văn số 5512/BGDĐT-GDTrH ngày 18/12/2020 về việc *Xây dựng và tổ chức thực hiện kế hoạch giáo dục của nhà trường*.
- [27] Nunnally J, (1978). *Psychometric Theory*. McGraw-Hill.
- [28] Hair JF, Black WC, Babin BJ & Anderson RE, (2014). *Multivariate Data Analysis* (7th International ed.). Pearson Education Limited.
- [29] DeVellis RF, (2016). *Scale Development: Theory and Applications* (Vol. 26). Sage, Thousand Oaks.
- [30] Cicchetti DV, (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 6(4), 284-290.
- [31] Honey M, Pearson G & Schweingruber H, (2014). *STEM Integration in K-12 Education: Status, prospects, and an agenda for research*. National Academies Press.