

XÂY DỰNG CÔNG CỤ ĐÁNH GIÁ NĂNG LỰC THIẾT KẾ CÔNG CỤ ĐÁNH GIÁ NĂNG LỰC CỦA SINH VIÊN SƯ PHẠM

Nguyễn Thị Diệu Linh và Đỗ Hương Trà
Khoa Vật lí, Trường Đại học Sư phạm Hà Nội

Tóm tắt. Bài viết giới thiệu việc xây dựng một công cụ để đánh giá (ĐG) năng lực thiết kế công cụ ĐG năng lực (NL) nhằm phục vụ việc bồi dưỡng NL này cho sinh viên sư phạm (SVSP). Công cụ được thiết kế nhằm đánh giá SVSP theo mô hình năng lực NL thiết kế công cụ ĐGNL gồm bốn thành tố, đó là: Xác định mục đích và mục tiêu của hoạt động ĐGNL; Lập kế hoạch xây dựng công cụ ĐG; Xây dựng câu hỏi/bài tập và công cụ cho điểm; Thử nghiệm và chỉnh sửa công cụ ĐG. Công cụ gồm bài kiểm tra trong đó có các tình huống đòi hỏi người được ĐG thực hiện tất cả các hành vi trong mô hình NL thiết kế công cụ ĐGNL và rubric đánh giá bài làm của họ. Độ giá trị và độ tin cậy của công cụ bước đầu được xác nhận. Kết quả nghiên cứu có thể sử dụng trong việc ĐG hoặc đào tạo SV sư phạm và bồi dưỡng GVPT về NLĐG.

Từ khóa: Công cụ đánh giá, đánh giá năng lực, bồi dưỡng năng lực, đào tạo sinh viên sư phạm, năng lực thiết kế công cụ đánh giá.

1. Mở đầu

Đánh giá NL có chức năng cung cấp cho người dạy và người học các thông tin phản hồi về những điểm mạnh, điểm yếu và mức độ năng lực hiện tại của người học. Từ đó, người dạy có cơ sở để cải tiến phương pháp dạy học nhằm định hướng người học đạt được mục tiêu. Tuy nhiên, do hầu như chỉ được tiếp xúc với ĐG truyền thống nên GV và SVSP gặp rất nhiều khó khăn khi ĐGNL, đặc biệt ở việc thiết kế công cụ ĐGNL học sinh [1].

Bài viết nằm trong một chuỗi các nghiên cứu của chúng tôi về NL thiết kế công cụ ĐGNL của SVSP. Trong bài viết trước, chúng tôi đã đề xuất khái niệm và cấu trúc của năng lực này [2]. Theo chúng tôi, năng lực thiết kế công cụ ĐGNL được hiểu là một năng lực thành phần của năng lực đánh giá. Đó là NL lựa chọn phương pháp ĐG thích hợp và xây dựng, phát triển các bài tập ĐG, quy trình chấm điểm phù hợp với mục đích ĐG. NL này có cấu trúc gồm 12 chỉ số hành vi (CSHV), mỗi CSHV được mô tả bởi 4 tiêu chí chất lượng.

Trong bài viết này, chúng tôi xây dựng công cụ ĐG năng lực thiết kế công cụ ĐGNL theo cấu trúc đã đề xuất. Công cụ này được xây dựng nhằm phục vụ việc bồi dưỡng NL thiết kế công cụ ĐGNL cho SVSP. Do đó, nó phải xác định được mức độ năng lực của SVSP tại thời điểm ĐG nhằm tìm hiểu tác động của các biện pháp phát triển NL này. Trên cơ sở đó, các NC tiếp theo sẽ xác định được các biện pháp hiệu quả trong bồi dưỡng NL thiết kế công cụ ĐGNL cho SVSP đáp ứng yêu cầu của dạy học phát triển NL ở trường phổ thông hiện nay.

Ngày nhận bài: 16/9/2019. Ngày sửa bài: 23/9/2019. Ngày nhận đăng: 30/9/2019.

Tác giả liên hệ: Nguyễn Thị Diệu Linh. Địa chỉ e-mail: linhntd@hnue.edu.vn

2. Nội dung nghiên cứu

2.1. Các công cụ đánh giá năng lực đánh giá của giáo viên trước đây

Việc xem xét các công cụ hiện có nhằm mục đích xác định mức độ phù hợp của các công cụ đó với mô hình NL thiết kế công cụ ĐGNL của GV trong nghiên cứu của chúng tôi. Tổng cộng, chúng tôi đã phân tích 15 công cụ được công bố trong các nghiên cứu sau năm 1990. Với từng công cụ, chúng tôi xác định các CSHV hoặc các kiến thức được ĐG, các mục đích mà công cụ hướng đến, và loại công cụ ĐG. Sau đó chỉ ra các phần của công cụ có thể được sử dụng để đánh giá NL thiết kế công cụ ĐGNL của GV theo mô hình NL thiết kế công cụ ĐGNL của nghiên cứu.

Mỗi mục đích ĐG phù hợp với một số loại công cụ ĐGNL. Với mục đích đánh giá NL, có một số công cụ chỉ giúp đánh giá một khía cạnh của NLĐG. Do đó, chúng tôi xem xét loại công cụ được sử dụng nhằm xác định các khía cạnh của NLĐG mà công cụ đó có thể đo được. Các công cụ hiện có được phân tích theo từng loại, gồm: 5 công cụ dưới dạng trắc nghiệm khách quan [3-7], 5 công cụ dưới dạng bảng hỏi [8-12], 1 công cụ là hỗn hợp của trắc nghiệm khách quan và bảng hỏi [13], 4 công cụ dưới dạng rubric [14-17].

Phân tích các công cụ trên cho thấy: chúng được thiết kế để đánh giá NLĐG của GV ở các khía cạnh khác nhau và được hướng đến các mục đích khác nhau. Các khía cạnh được các công cụ ĐG này gồm: kiến thức về ĐG của GV (sử dụng các phương pháp ĐG, sử dụng các hoạt động ĐG để phân loại học sinh; giải thích kết quả kiểm tra tiêu chuẩn hóa và ĐG lớp học, truyền đạt kết quả ĐG và tuân thủ các nghĩa vụ pháp lí và đạo đức trong ĐG) và mức độ thành thạo trong thực hành ĐG của GV. Phần lớn các công cụ ĐG được xây dựng dựa trên “Tiêu chuẩn về năng lực của GV trong đánh giá HS” năm 1990 của Mỹ. Trắc nghiệm khách quan nhiều lựa chọn đã được sử dụng để ĐG kiến thức về ĐG của GV. Đối chiếu các kiến thức về ĐG được đo bởi các công cụ đó với mô hình NL thiết kế công cụ ĐGNL của SVSP, cho thấy vẫn chưa có công cụ nào ĐG các kiến thức về: xác định mục tiêu ĐG, biên soạn câu hỏi/bài tập, thử nghiệm công cụ ĐGNL. Mặt khác, để đo mức độ thành thạo trong thực hành ĐG của GV, các công cụ sử dụng thang đo Likert và rubric. Với thang đo Likert, người trả lời tự ĐG nên kết quả ĐG bị ảnh hưởng nhiều bởi tính chủ quan. Kết quả ĐG thu được bởi thang đo Likert thường được dùng để thăm dò, không thể chỉ dựa vào đó để xác định NL người trả lời. Rubric là loại công cụ phù hợp để ĐGNL. Tuy nhiên, các rubric hiện có tập trung vào ĐG NL thiết kế các loại PPĐG riêng biệt hoặc chỉ đưa ra tiêu chí ĐG chung về việc sử dụng PPĐG, chưa tập trung vào NL thiết kế công cụ ĐGNL [14-16]. Do đó, chúng tôi sẽ thiết kế công cụ ĐG gồm các bài tập và rubric ĐG đầy đủ các CSHV trong mô hình NL thiết kế công cụ ĐGNL của SVSP.

2.2. Thiết kế công cụ đo đánh giá năng lực thiết kế công cụ đánh giá năng lực của sinh viên sư phạm

2.2.1. Thiết kế bài tập đánh giá năng lực thiết kế công cụ đánh giá năng lực

Công cụ ĐG năng lực thiết kế công cụ ĐGNL được xây dựng trong nghiên cứu này nhằm phục vụ việc bồi dưỡng NL thiết kế công cụ ĐGNL cho SVSP nên phải giúp xác định mức độ ở từng CSHV của SVSP. Để làm được điều đó, các bài tập phải ĐG đúng các CSHV và cũng phải có độ khó khác nhau sao cho mọi SV đều có cơ hội bộc lộ mức độ của bản thân ở các CSHV cần ĐG [18]. Do đó, chúng tôi sử dụng cách tiếp cận diễn dịch để xây dựng các công cụ này. Tức là, xuất phát từ mô hình của NL thiết kế công cụ ĐGNL của SVSP, chúng tôi xác định rằng: nhiệm vụ đưa ra trong đề bài luôn đòi hỏi SV thực hiện hành vi cần ĐG; yêu cầu đối với các nhiệm vụ phải bám sát các tiêu chí chất lượng; những thông tin cần cung cấp về tình huống trong đề bài phải đảm bảo cung cấp đầy đủ thông tin cần thiết để SV có thể bộc lộ được mức độ của bản thân ở các CSHV cần ĐG.

Dưới đây, chúng tôi trình bày cụ thể việc xây dựng các bài tập ĐG các CSHV trong mô hình NL thiết kế công cụ ĐGNL.

Đánh giá CSHV 1.1. Xác định mục đích sử dụng công cụ ĐG. Những SV đạt tiêu chí chất lượng ở mức cao của CSHV này sẽ xác định được tất cả các mục đích ĐG có thể hướng đến khi thiết kế công cụ ĐGNL trong một tình huống xác định và biết được mục đích nào cần ưu tiên. Do đó, đề SV thể hiện được CSHV này, đề bài cần đưa ra một tình huống thực tế trong đó xuất hiện nhu cầu ĐGNL học sinh. SV phải xác định các mục đích ĐG có thể đạt được trong tình huống đó. Cả hai bài: Bài 1 của đề số 1 và Bài 1.1 của đề số 2 đều được biên soạn để ĐG CSHV này.

Bài 1.1 của đề số 2 đã đưa ra câu hỏi trực tiếp để SV thể hiện mức độ ở CSHV này. Bài 1 của đề số 1 không hỏi trực tiếp mà đưa ra nhiệm vụ để SV bộc lộ mức độ của bản thân ở CSHV này qua quá trình thực hiện. Phân tích nhiệm vụ yêu cầu trong Bài 1.1 của đề số 1, có thể thấy khi viết thư trình bày ý kiến về việc tổ chức ĐGNLTN trong đợt thi cuối năm, SV cần phải nêu ra các mục đích mà ĐGNL có thể đạt được và các mục đích mà ĐG kiến thức, kỹ năng theo cách truyền thống có thể đạt được trong tình huống của đề bài. Để thuyết phục người đọc, trong lập luận của mình, người đó phải đưa ra các mục đích quan trọng hơn lên trước. Trên cơ sở so sánh các mục đích mà hai cách đánh giá này có thể đạt được, SV đưa ra ý kiến của bản thân.

Đánh giá CSHV 1.2. Xác định đặc điểm chính của tình huống sử dụng công cụ. Đề SV thể hiện CSHV này, đề bài cần đưa ra một tình huống thực tế trong đó mục đích ĐGNL đã được xác định. SV phải nêu được các đặc điểm chính của tình huống sử dụng công cụ. Tuy nhiên, cần phải đưa ra câu hỏi để biết SV đã xác định các đặc điểm của tình huống bằng cách nào, qua đó mới có thể xác định được mức độ của SV đạt được ở CSHV 1.2. Bài 2 của đề số 1, Bài 1.2 của đề số 2 đều được biên soạn để ĐG CSHV này. Bài 2 của đề số 1 có thể ĐG riêng biệt CSHV 1.2 bởi vì đề bài đã cung cấp mục đích ĐG. Do đó, nếu SV không làm được bài tập ĐG CSHV 1.1, thì vẫn có cơ hội thể hiện đúng mức độ ở CSHV 1.2 của bản thân.

Đánh giá CSHV 1.3. Xác định các mục tiêu đánh giá của hệ thống câu hỏi/bài tập. Đề SV thể hiện CSHV này, đề bài cần đưa ra một tình huống thực tế trong đó trình bày mục đích ĐGNL và các đặc điểm của tình huống. SV phải xác định các yêu cầu về phạm vi nội dung kiến thức, cách diễn đạt và đặc biệt các mục tiêu phải được chỉ ra cụ thể như: các mức độ cần ĐG và thời lượng tối đa dành cho việc ĐG từng CSHV. Bài tập ĐG CSHV này được minh họa bằng Bài 1.3 trong đề số 2.

Đánh giá CSHV 2.1 Xác định loại minh chứng và số lượng minh chứng cần thu thập. Đề bài đánh giá CSHV này cần đưa ra mục tiêu ĐG của hệ thống câu hỏi/bài tập trong đó chỉ rõ mức độ cần ĐG và thời lượng tối đa dành cho việc ĐG từng CSHV. Đề bài yêu cầu SV xác định các loại minh chứng cần thu thập sao cho phù hợp với các mục tiêu đó. Có thể có nhiều loại minh chứng giúp ĐG một CSHV. Để được ĐG ở mức cao nhất, SV cần chọn các loại minh chứng tốn ít công sức, cần điều kiện đơn giản để thu thập và nhiều CSHV của NL cần ĐG cũng cần loại minh chứng đó. Nhờ đó, việc xác định loại và số lượng minh chứng mới đạt được đầy đủ các mục tiêu ĐG. Bài tập đánh giá CSHV này được minh họa bằng Bài 2.1 trong đề số 2.

Đánh giá CSHV 2.2 Lựa chọn các phương pháp đánh giá (PPĐG) hỗ trợ việc thu thập các loại minh chứng đó. Đề SV bộc lộ CSHV này, đề bài cần đưa ra các minh chứng cần thu thập và yêu cầu SV xác định các PPĐG thu thập các bằng chứng đó. SV đạt được mức cao nhất khi họ hiểu rõ ưu, nhược điểm, cách thực hiện từng loại PPĐG và vận dụng các kiến thức đó để chọn được PPĐG phù hợp với mục tiêu. Bài tập 2.2 trong đề số 2 được biên soạn để đánh giá CSHV này.

Đánh giá CSHV 3.1 Xác định đặc điểm thông tin cần sử dụng để biên soạn câu hỏi/bài tập. Để đánh giá CSHV này, đề bài cần cung cấp mục tiêu ĐG, các PPĐG từng CSHV và yêu cầu SV nêu các đặc điểm thông tin cần sử dụng để biên soạn câu hỏi/bài tập. SV đạt được mức 3

khi xác định được các thông tin có thể dùng để làm ý tưởng biên soạn câu hỏi/BT, khi đó họ đã biết và vận dụng được cách biến đổi loại thông tin đó thành các tình huống để đưa vào bài tập. SV đạt được mức cao nhất khi họ thể hiện sự sáng tạo bằng cách tìm ra những quy tắc mới, hợp lí cho bản thân. Các quy tắc này chỉ có thể được tìm ra khi họ suy ngẫm về quá trình làm bài của mình. Để ĐG được mức cao nhất này, đề bài không chỉ yêu cầu SV đưa ra quy tắc hợp lí khi xác định loại thông tin cần sử dụng mà còn yêu cầu họ trình bày quá trình suy ngẫm của bản thân. Bài tập đánh giá CSHV này được minh họa bằng Bài 3.1 trong đề số 2.

Đánh giá CSHV 3.2 Biên soạn câu hỏi/bài tập. Đề SV bộc lộ CSHV này, đề bài cần đưa ra một tình huống trong đó chỉ ra đầy đủ mục tiêu ĐGNL, gồm: ĐG mức độ nào. SV được yêu cầu xác định các đặc điểm chính của tình huống sử dụng công cụ. Tuy nhiên, cần phải đưa ra câu hỏi để biết SV làm thế nào để xác định được các đặc điểm đó. Hai bài: Bài 3 trong đề số 1 và Bài 3.2 đề số 2 đều đánh giá CSHV này. Ở phần a Bài 3 trong đề số 1, có thể thấy câu i) của bài là một tình huống thường gặp. Những SV biên soạn được bài tập đáp ứng phần a này sẽ đạt được mức 2. Câu ii) của bài là tình huống SV ít gặp, SV cần chú ý phối hợp ĐG các CSHV mới có thể làm được, do đó câu này ĐG năng lực thiết kế công cụ ĐGNL ở mức 3 trong trường hợp thiết kế công cụ ĐGNL tiến hành thí nghiệm và xử lí số liệu. Phần b của bài này nhằm ĐG mức 3 của CSHV 3.2 trong trường hợp thiết kế công cụ ĐGNL thiết kế PATN. Với yêu cầu SV nêu quá trình suy nghĩ khi biên soạn bài tập, đề bài nhằm ĐG mức 4 của CSHV này.

Đánh giá CSHV 3.3 Xây dựng công cụ để đánh giá các minh chứng thu được. Bài tập đánh giá CSHV này phải cung cấp câu hỏi/bài tập cần xây dựng công cụ ĐG câu trả lời/bài làm của HS. SV phải dựa trên kiến thức về các loại công cụ để lựa chọn các loại công cụ phù hợp và xây dựng các công cụ đó. Để đạt được mức cao nhất của CSHV này, người đó cần phân tích quá trình trả lời câu hỏi/giải bài tập và các mức độ được mô tả trong cấu trúc của NL cần ĐG để xây dựng được công cụ và tạo cơ hội học tập cho HS. Bài tập đánh giá CSHV này được minh họa bằng Bài 3.3 trong đề số 2.

Đánh giá CSHV 3.4 Xác định các bước cụ thể người đánh giá cần thực hiện để sử dụng công cụ. Đề SV bộc lộ CSHV này, đề bài cần cung cấp các thông tin: mục đích ĐG, đặc điểm của tình huống ĐG, đề câu hỏi/bài tập, công cụ để ĐG các minh chứng thu được, và yêu cầu SV xác định các bước cụ thể người ĐG cần thực hiện để sử dụng công cụ. Bài tập 3.4 trong đề số 2 dùng để đánh giá CSHV này.

Đánh giá CSHV 4.1 Thử nghiệm các công cụ đánh giá để xác định chất lượng và khả năng áp dụng. Thông tin cần cung cấp trong đề bài tập đánh giá CSHV này gồm: mục đích ĐG, đặc điểm của tình huống ĐG, câu hỏi/bài tập, công cụ để ĐG các minh chứng thu được, và các bước cụ thể người ĐG cần thực hiện để sử dụng công cụ. SV phải lên kế hoạch thử nghiệm trong đó trình bày cách chọn mẫu thử nghiệm, cách thức thu thập thông tin thử nghiệm. Để đạt được mức cao nhất của CSHV này, SV cần chỉ rõ cách thực hiện và các điểm cần lưu ý khi thử nghiệm để loại bỏ hoặc giảm thiểu các yếu tố ảnh hưởng đến sự sẵn sàng tham gia của đối tượng thử nghiệm. CSHV này được ĐG bằng Bài 4.1 trong đề số 2.

Đánh giá CSHV 4.2 Phân tích, đánh giá các thông tin phản hồi từ những người liên quan đến quá trình thử nghiệm. Đề ĐG CSHV này, đề bài cần đưa ra một tình huống thực tế trong đó cung cấp các thông tin: mục đích ĐG, đặc điểm của tình huống ĐG, câu hỏi/bài tập, công cụ để ĐG các minh chứng thu được, kế hoạch thử nghiệm cụ thể và các thông tin thu được từ những người liên quan đến quá trình thử nghiệm. Đề bài yêu cầu SV phân tích các thông tin phản hồi nhằm xác định mức độ phù hợp và những điểm cần chỉnh sửa của công cụ ĐG. Một số khía cạnh của CSHV này được ĐG qua Bài 4.1 trong đề số 2. CSHV này được ĐG đầy đủ qua Bài 4.1 trong đề số 2.

Đánh giá CSHV 4.3 Chính sửa công cụ đánh giá. Đề bài đánh giá CSHV này cũng cần đưa ra những thông tin tương tự đề bài đánh giá CSHV 4.2 và yêu cầu SV chỉnh sửa công cụ ĐG đó. Vì vậy, có thể ĐG cả CSHV 4.2 và 4.3 trong cùng một bài như Bài 4 trong đề số 1.

Kết quả ĐG bằng các bài tập trong đề số 1 không bị ảnh hưởng bởi các CSHV khác. Do đó, nó rất phù hợp để ĐG trong quá trình bồi dưỡng NL thiết kế công cụ ĐG NL. Nó giúp người dạy và người học xác định rõ các mức độ hiện tại của từng CSHV ở người học. Hơn nữa, nó cũng giúp người học thoải mái trong việc phát triển từng CSHV, không cần phải đợi để phát triển các CSHV theo tuần tự.

Đề số 1

1. Giả sử em đã trở thành GV. Năm nay, tổ Vật lí trường em đã tổ chức nhiều buổi trao đổi chuyên môn hơn trước. Trong buổi trao đổi chuyên môn gần đây, một số GV trong tổ chưa ủng hộ việc tổ chức ĐGNLTN trong đợt thi cuối năm cho khối 10, 11 và vẫn muốn ĐG kiến thức, kĩ năng theo cách truyền thống vì họ cho rằng: “Thi cuối năm để cho điểm tổng kết. Chỉ cần ĐG kiến thức cũng phân biệt được học sinh giỏi với học sinh kém rồi, không cần tổ chức ĐGNLTN vì như vậy sẽ tốn thời gian và công sức”. Em đang trong một chuyến công tác dài ngày, em hãy viết thư (dưới 300 từ) trình bày ý kiến của em.

2. Cô Hoa mới được giao dạy một lớp 11. Vào đầu năm học, cô Hoa muốn ĐG năng lực thực nghiệm của học sinh nhằm các mục đích sau: xác định PP dạy học phù hợp; phản hồi cho phụ huynh và cho học sinh và truyền đạt yêu cầu học tập. Nếu em giúp cô Hoa thiết kế được các công cụ ĐG đạt được các mục đích trên trong thời gian 1 tuần, em cần chú ý đến đặc điểm gì của tình huống này? Cần phải tìm hiểu những đặc điểm gì của HS lớp đó? Hãy sắp xếp các đặc điểm đó theo thứ tự ưu tiên. Để tìm hiểu từng đặc điểm em đã nêu, hãy chỉ rõ: Loại và số lượng dữ liệu thực tế cần tìm; Cách tìm dữ liệu đó; cách xử lí dữ liệu.”

3. a. Từ các PATN bài đo gia tốc rơi tự do (SGK vật lí lớp 10), hãy biên soạn bài tập để ĐGNL “Tiền hành thí nghiệm và xử lí số liệu”. Thời điểm ĐG là sau khi HS học xong bài ôn tập chương 1, chưa học bài thực hành “Đo gia tốc rơi tự do”. Hãy biên soạn các bài tập ứng với các tình huống sau:

i, nếu thời gian cho việc ĐGNL “Tiền hành thí nghiệm và xử lí số liệu” của mỗi HS có thể kéo dài tối đa 30’.

ii, nếu thời gian tối đa cho việc ĐGNL “tiền hành thí nghiệm” của mỗi HS chỉ có 10’.

b. Hãy soạn thảo càng nhiều càng tốt bài tập ĐG năng lực “thiết kế phương án thí nghiệm” từ các thông tin gợi ý sau:

- Ứng dụng kĩ thuật: phanh điện từ
- Bài tập tính toán: một viên bi được ném với vận tốc ban đầu là v theo hướng hợp với phương ngang một góc α . Hỏi viên bi chạm đất tại điểm nào.
- Sản phẩm: bình thủy điện
- Lời khuyên: mặc nhiều áo mỏng ấm hơn mặc một áo dày
- Hiện tượng: quá trình chuyển động của một vật khi bị rơi xuống nước kể từ mặt nước đến khi chạm đáy.

Em hãy nêu quá trình suy nghĩ của mình để biên soạn được bài tập đó. Làm thế nào em có được cách biên soạn bài tập từ các thông tin trên. Với mỗi bài tập, trình bày ít nhất một PATN thỏa mãn yêu cầu của bài tập.

4. Trong file “Phụ lục câu 4” trình bày công cụ ĐG do một SV thiết kế, kèm theo là nội dung các hoạt động thử nghiệm và một số thông tin phản hồi từ những người liên quan đến quá trình thử nghiệm. Em có nhận xét gì về việc thử nghiệm công cụ ĐG NLTN của sinh viên trong tài liệu. Hãy phân tích các thông tin thu được trong thử nghiệm đó và xác định các yếu tố có thể ảnh hưởng đến độ chính xác và sự tối ưu của ĐG. Trên cơ sở đó, em hãy chỉnh sửa công cụ ĐG.

Tuy nhiên, đề bài tập ĐG riêng biệt từng CSHV khá dài và SV phải mất nhiều thời gian để nắm bắt thông tin cung cấp trong đề. Do đó, chỉ khi ĐGNL thiết kế những công cụ đánh giá NL phức tạp như NL thực nghiệm, SV có thể chưa đạt được các mức phù hợp ở một số CSHV mới cần phải sử dụng bài tập ĐG riêng biệt từng CSHV. Khi ĐGNL thiết kế công cụ ĐG những NL không quá phức tạp như: năng lực trao đổi thông tin, năng lực tự học, có thể sử dụng các bài tập ĐG tổng hợp các CSHV. Các bài tập tổng hợp này được thiết kế để giao cho sinh viên vào cuối kì học, khi hầu hết SV đều đạt yêu cầu ở các CSHV không quá khó. Mặt khác, khi chấm điểm bài làm, kết quả của những câu trước được coi là điều kiện đã cho của câu sau, người chấm chỉ ĐG trong phạm vi của từng câu mà không cần quan tâm đến kết quả làm câu trước. Dưới đây là đề bài tập ĐG tổng hợp nhiều CSHV (từ chỉ số hành vi 1.1 đến 4.1) khi SV thiết kế công cụ ĐGNL trao đổi thông tin khoa học.

Đề số 2

Vào đầu năm học, trường cô Mai cần chọn một nhóm học sinh có năng lực trao đổi thông tin khoa học. Cô Mai dạy lớp 10C và đang dạy bài *Chuyển động biến đổi đều* chương 1 sách giáo khoa Vật lí 10 Nâng cao. Trong vòng hai tuần tới, cô phải chọn ra trong lớp những học sinh có năng lực này. Em sẽ giúp cô Mai thiết kế công cụ ĐG năng lực trao đổi thông tin khoa học. Hãy trả lời lần lượt các câu hỏi sau:

1.1 Theo em, những ai có thể sử dụng kết quả của việc ĐG đó và sử dụng với các mục đích nào? (sắp xếp các mục đích theo thứ tự ưu tiên).

1.2 Để đạt được các mục đích trên, trước khi thiết kế công cụ ĐGNL em cần chú ý đến những đặc điểm gì của tình huống này. Cần phải tìm hiểu những đặc điểm gì của HS lớp đó? Để cô Mai có thể tìm hiểu được từng đặc điểm trên, em hãy nêu rõ: Loại dữ liệu em cần tìm; Cách tìm dữ liệu đó; cách xử lí dữ liệu đó.

1.3 Trong tình huống này, các câu hỏi/bài tập dùng để ĐG NL trao đổi thông tin khoa học cần phải đạt những yêu cầu nào. Chỉ rõ các mức độ cần ĐG và thời lượng tối đa dành cho việc ĐG từng CSHV của NL trao đổi thông tin khoa học.

2.1 Nêu rõ các loại và số lượng minh chứng (câu trả lời, bài viết, sản phẩm, hành động...) cần thu thập để ĐG từng CSHV.

2.2 Nêu rõ các phương pháp ĐG nên sử dụng để ĐG thu thập các loại minh chứng đã nêu ở câu 2.1.

3.1 Khi biên soạn câu hỏi/bài tập ĐG NL trao đổi thông tin khoa học, em cần sử dụng một số thông tin để tạo ra dữ kiện của đề bài. hãy mô tả những thông tin cần tìm để người khác có thể tìm giúp em. Hãy trình bày một số thông tin mà em tìm được. Nhìn lại quá trình suy nghĩ khi em xác định đặc điểm thông tin cần sử dụng để biên soạn câu hỏi/bài tập, em có rút ra quy tắc nào cho những lần sau? Nếu có, hãy trình bày quy tắc đó.

3.2 Từ các thông tin đã tìm được, em hãy soạn thảo câu hỏi/bài tập ĐG NL trao đổi thông tin khoa học. Em hãy nêu quá trình suy nghĩ của mình để biên soạn được các câu hỏi/bài tập đó. Nhìn lại quá trình suy nghĩ khi em biên soạn câu hỏi/bài tập, em có rút ra quy tắc nào cho những lần sau? Nếu có, hãy trình bày quy tắc đó.

3.3 Em hãy trình bày công cụ (thang đo, bảng kiểm, rubric...) để ĐG các minh chứng thu được.

3.4 Em hãy nêu các bước cụ thể để sử dụng công cụ ĐG sao cho cô Mai có thể thực hiện được.

4.1 Em hãy trình bày kế hoạch thử nghiệm công cụ ĐG trên để xác định chất lượng và khả năng áp dụng của nó, trong đó trình bày rõ: thời điểm thử nghiệm, cách chọn đối tượng thử nghiệm, cách thu thập thông tin phản hồi từ những người liên quan đến quá trình thử nghiệm, cách xử lí thông tin đó. Giải thích vì sao em lập kế hoạch thử nghiệm như vậy.

Thực hiện theo cách biên soạn trên, chúng tôi đã biên soạn được 4 đề kiểm tra, trong đó có hai đề ĐG từng CSHV riêng biệt và hai đề ĐG tổng hợp nhiều CSHV của NL thiết kế công cụ ĐGNL.

2.2.2. Thiết kế rubric đánh giá năng lực thiết kế công cụ đánh giá năng lực

Việc sử dụng xây dựng “rubric” từ việc quy nạp rất nhiều sản phẩm sẽ khá mất nhiều công sức và thời gian, nó phù hợp với nghiên cứu trên quy mô lớn. Trong giới hạn quy mô của nghiên cứu, chúng tôi chọn cách tiếp cận diễn dịch để thiết kế “rubric” dùng để chấm điểm bài làm của SV với các đề kiểm tra.

Mỗi tiêu chí chất lượng trong cấu trúc của NL thiết kế công cụ ĐGNL là những mô tả tổng quát về các biểu hiện của những SV đạt tiêu chí đó. Do đó, việc xây dựng “rubric” của mỗi bài kiểm tra là sự cụ thể hóa của các tiêu chí chất lượng trong cấu trúc của NL thiết kế công cụ ĐGNL, là những suy luận từ tiêu chí chất lượng trong tình huống của bài kiểm tra. Với cách tiếp cận này, để đưa ra những mô tả của từng mức độ, trước hết chúng tôi đóng vai SV để làm đề kiểm tra. Trên cơ sở nhìn lại quá trình tư duy khi giải từng bài tập đó, chúng tôi xác định các điểm SV có thể gặp khó khăn, sau đó nhóm các khó khăn đó theo từng tiêu chí chất lượng trong mô hình NL thiết kế công cụ ĐGNL. Kết hợp với những biểu hiện của SV quan sát được trong quá trình dạy học môn “Kiểm tra đánh giá trong giáo dục” trước đây, chúng tôi đưa ra những biểu hiện của từng nhóm khó khăn nói trên, từ đó xác định được các mô tả của các mức độ trong rubric. Các mức độ này được chúng tôi mô tả theo trình tự từ trên xuống phù hợp với cách tiếp cận diễn dịch [19].

Các mô tả trong rubric được thiết kế theo đề xuất của Brookhart (2013) về “Các đặc điểm mong muốn của mô tả mức độ của sự thực hiện cho “Rubrics”, gồm có: mô tả theo những gì được quan sát được; HS và GV đều hiểu đúng nghĩa những gì mô tả; mô tả một cách liên tục từ mức này sang mức khác cho một tiêu chí, không có miễn trừ; khác nhau từ cấp độ này đến cấp độ khác (có thể kết hợp các ví dụ về công việc với các mô tả ở mỗi mức); mô tả song song từ mức độ này sang mức độ khác cho cùng một khía cạnh. Những đặc điểm này giúp đảm bảo mỗi biểu hiện của SV luôn xếp được vào một mức độ tương ứng. Điều đó giúp người ĐG xác định chính xác mức độ đạt được của SV. Bảng 1 dưới đây là “rubric” ĐG của bài tập 3 của đề số 1.

Bảng 1. “Rubric” đánh giá một số chỉ số hành vi của năng lực thiết kế công cụ đánh giá năng lực

CSHV	Mức độ 1	Mức độ 2	Mức độ 3	Mức độ 4
Biên soạn câu hỏi/bài tập	1. Biên soạn được một số câu hỏi/bài tập ĐG NL tiến hành TN và xử lý số liệu ở phần a câu 3.	2. Biên soạn các câu hỏi/bài tập để ĐG đầy đủ các CSHV của NL tiến hành TN và xử lý số liệu ở câu 3. Biên soạn được ít nhất một câu hỏi/BT ĐG NL thiết kế PATN.	3. Thêm vào so với st2: Biên soạn được một số câu hỏi/BT ĐG NL thiết kế PATN theo một số cách khác nhau (từ ứng dụng kỹ thuật, từ bài tập...) với độ khó khác nhau.	4. Từ mỗi thông tin gợi ý, biên soạn được câu hỏi/bài tập ĐG NL thiết kế PATN. Biên soạn được ít nhất 1 câu hỏi/bài tập ĐG NL thiết kế PATN theo cách khác (không dựa vào các loại thông tin trong đề bài) và nêu được cách biên soạn bài tập rút ra từ quá trình biên soạn bài tập mới này.

2.2.3. Thử nghiệm công cụ

Để kiểm tra độ giá trị nội dung, các bài tập và “rubric” đã soạn thảo được hai chuyên gia về đánh giá và hai đồng nghiệp cùng dạy môn “Kiểm tra đánh giá trong giáo dục” xem xét sự phù

hợp của các bài tập với mô hình NL thiết kế công cụ ĐGNL của chúng tôi. Các chuyên gia được yêu cầu trả lời bộ câu hỏi được đề xuất bởi Moskal và Leydens (2000) như sau: Các tiêu chí đánh giá của “rubric” có phản ánh tất cả các khía cạnh dự định được đo lường trong các bài tập hay không? (Hiệu lực nội dung); Có phải tất cả các tiêu chí quan trọng liên quan đến phương pháp đánh giá được đánh giá thông qua rubric? (Hiệu lực cấu trúc); và các tiêu chí chấm điểm có ĐG các năng lực liên quan đến khả năng thành công trong tương lai không? (Giá trị tiêu chuẩn). Các chuyên gia về đánh giá và chuyên gia môn học đã đưa ra một số đề nghị chỉnh sửa nhỏ về cách sử dụng thuật ngữ và cách diễn đạt để người đọc hiểu đúng ý tác giả. Trên cơ sở tổng hợp các ý kiến của chuyên gia, chúng tôi thảo luận và chỉnh sửa lại đề bài. Các chuyên gia đã thống nhất rằng công cụ sau khi chỉnh sửa đã đảm bảo độ giá trị nội dung [20].

Các bài tập trên được thử nghiệm trên 21 SV Khoa Vật lí, Trường Đại học Sư phạm Hà Nội nhằm chỉnh sửa những điểm có thể gây hiểu lầm, đánh giá sự phù hợp giữa thời gian dự kiến và thực tế để hoàn thành bài tập; xác định mức độ năng lực mà bài tập có thể đánh giá được. Việc thử nghiệm được tiến hành trong khi SV đang học môn “Kiểm tra đánh giá trong giáo dục”. Trước khi SV làm đề kiểm tra, chúng tôi cung cấp bản cấu trúc của NL thiết kế công cụ ĐGNL nhằm giúp SV hiểu yêu cầu cần đạt. Kết quả là có hai thuật ngữ gây nhầm lẫn cho một số SV. Sau khi thảo luận chúng tôi đã thay đổi hai thuật ngữ này bằng các từ tương đương.

Phân tích sản phẩm của SV và thảo luận chuyên sâu giữa các giảng viên sẽ đảm bảo sự phù hợp giữa các mô tả trong rubric ban đầu được xây dựng dựa trên suy luận và kinh nghiệm của người nghiên cứu với thực tiễn [21]. Do đó, trước khi thử nghiệm, chúng tôi đối chiếu “rubric” với 21 bài làm của SV thu được từ việc thử nghiệm bài tập nêu trên. Với các bài làm của SV không có các đặc điểm trong mô tả rubric ban đầu, chúng tôi thảo luận để đưa ra mô tả mới.

Theo Reddy (2011), ngoài những phản hồi của giảng viên về sự rõ ràng, đầy đủ và khả năng áp dụng khái quát của “rubric”, cần phải lấy phản hồi từ SV về các rubric và thực hiện sửa đổi. Chúng tôi lấy phản hồi từ 68 sinh viên, sau khi họ học xong môn “Kiểm tra đánh giá trong giáo dục”. Các sinh viên được yêu cầu đưa ra phản hồi liên quan đến ngôn ngữ, sự hiểu biết, sự rõ ràng, khả năng sử dụng các rubric để tự đánh giá. Có 93% SV cho rằng các rubric là dễ hiểu và dễ sử dụng. Có năm sinh viên đã cho rằng cần thiết phải thảo luận về các phiếu tự đánh giá trong lớp học kèm theo các bài mẫu ở một số mức độ để giúp hiểu rõ hơn về sự khác biệt về trình độ. Với các phản hồi này, chúng tôi cũng thảo luận để chỉnh sửa một số mô tả [21].

Để kiểm tra độ tin cậy của “rubric”, chúng tôi thử nghiệm trên 7 SV đã học xong học phần “Kiểm tra đánh giá trong giáo dục” từ học kì trước. Nhằm thử nghiệm được đầy đủ các mức độ năng lực khác nhau trong đề kiểm tra, các SV được chọn có điểm thi của học phần này trải từ giỏi đến trung bình. Với những câu có sự khác nhau trong đánh giá của hai người chấm về mức độ đạt được của SV, chúng tôi trao đổi cùng họ để chỉnh sửa rubric. Sau đó, hai người chấm tiến hành chấm lại câu đó trên bài làm của SV khác. Rubric được chỉnh sửa sau khi thảo luận kết quả chấm 4 bài làm của SV. Sử dụng “rubric” này để chấm bài làm của ba SV còn lại, kết quả hai người chấm là trùng nhau. Theo tài liệu về độ tin cậy giữa những người đánh giá, đối với các đánh giá trong lớp học, hai người đánh giá là đủ để đưa ra mức độ chấp nhận của độ tin cậy. Trong nghiên cứu này, chúng tôi cho hai người đánh giá độc lập chấm điểm từng bài làm của SV. Những người đánh giá không biết kết quả của người ĐG khác và tên của SV. Cỡ mẫu trong nghiên cứu này nằm trong phạm vi chấp nhận được là 6-12 mẫu cho mỗi khóa học và phương pháp đánh giá. Như vậy, với việc thử nghiệm trên phạm vi hẹp, bước đầu công cụ đã đảm bảo độ tin cậy [22, 23].

3. Kết luận

Bài viết đã cung cấp một công cụ mới có thể hỗ trợ nghiên cứu và phát triển NL thiết kế công cụ ĐGNL cho SVSP. Những bằng chứng nhằm ban đầu chứng tỏ tính hợp lệ và độ tin cậy của công cụ cũng được trình bày trong bài viết. Với cách cung cấp thông tin về tình huống trong các bài kiểm tra, có thể thấy công cụ được thiết kế trong NC của chúng tôi đã đảm bảo cho các mức độ của NL thiết kế công cụ ĐGNL được bộc lộ đầy đủ. Do đó, công cụ này có thể mở ra những nghiên cứu mới về thực trạng NL thiết kế công cụ ĐGNL của SVSP hoặc GV, cũng như những NC về đào tạo SVSP và bồi dưỡng cho GV vật lí năng lực thiết kế công cụ ĐGNL. Công cụ này cũng có thể được sử dụng để thu thập các dữ liệu thực nghiệm nhằm xây dựng đường phát triển NL của năng lực thiết kế công cụ ĐGNL. Với việc bồi dưỡng năng lực thiết kế công cụ ĐGNL cho SVSP các môn khác, công cụ cũng có thể được sử dụng nếu thay đổi một số chi tiết liên quan đến môn vật lí trong đề kiểm tra. Thêm vào đó, bài viết đã trình bày khá chi tiết các giai đoạn xây dựng công cụ nên các nhà nghiên cứu và các giảng viên có thể sử dụng và cải tiến công cụ này trong các NC nhằm nâng cao NLĐG của giáo viên.

TÀI LIỆU THAM KHẢO

- [1] Nguyễn Thị Diệu Linh, Đỗ Hương Trà, 2016. *Từ khảo sát thực tiễn đến đề xuất giải pháp phát triển năng lực thiết kế công cụ đánh giá năng lực trong bồi dưỡng và đào tạo giáo viên vật lí*. Tạp chí Khoa học Trường Đại học Sư phạm Hà Nội, Vol. 61, số 8B, tr 213-225.
- [2] Nguyễn Thị Diệu Linh, Đỗ Hương Trà, 2018. *Cấu trúc năng lực thiết kế công cụ đánh giá năng lực và việc bồi dưỡng năng lực thiết kế công cụ đánh giá năng lực cho sinh viên sư phạm*. Tạp chí Khoa học Trường Đại học Sư phạm, Đại học Đà Nẵng. Số 29b(3), tr. 47-54.
- [3] Plake, B. *et al.*, 1993. *Assessment competencies of teachers: a national survey*. Educational Measurement: Issues and Practice, 12(4),10-39. doi:10.1111/j.1745-3992.1993.tb00548.x.
- [4] Campbell, C. *et al.*, 2002. *Psychometric analysis of an assessment literacy instrument: applicability to pre-service teachers*. Paper presented at the annual meeting of the Mid-Western Educational Research Association, Columbus.
- [5] Mertler, C.A., 2003. *Pre-service versus in-service teachers' assessment literacy: does classroom experience make a difference?*. Annual meeting of the Mid-Western Educational Research Association, Columbus.
- [6] Mertler, C.A., & Campbell, C., 2005. *Measuring teachers' knowledge & application of classroom assessment concepts: development of the assessment literacy inventory*. In Annual meeting of the American Educational Research Association, Montreal.
- [7] Daniel, L.G. & King, D.A., 1998. *Knowledge and use of testing and measurement literacy of elementary and secondary teachers*. The Journal of Educational Research, 91(6), pp. 331-344.
- [8] Kershaw IV, I., 1993. *Ohio vocational education teachers' perceived use of student assessment information in educational decision-making*. Ohio State University.

- [9] Zhang, Z. & Burry-stock, J.A., 1997. *Assessment practices inventory: a multivariate analysis of teachers' perceived assessment competency*. Annual meeting of the American Educational Research Association, Chicago.
- [10] Brown, G. L., 2002. *Teachers' Conceptions of Assessment*. PhD thesis in the University of Auckland, USA.
- [11] Jarr, K.A., 2012. *Education practitioners' interpretation and use of assessment results*. University of Iowa. Retrieved from <http://ir.uiowa.edu/etd/3317>.
- [12] DeLuca, C. & Klinger, D. A., 2010. *Assessment literacy development: Identifying gaps in teacher candidates' learning*. *Assessment in Education: Principles, Policy & Practice*, 17, 419-438. doi:10.1080/0969594X.2010.516643.
- [13] DeLuca, C. et al., 2016. *Approaches to Classroom Assessment Inventory: A New Instrument to Support Teacher Assessment Literacy*. *Educational Assessment*, 21(4), 248-266, DOI: 10.1080/10627197.2016.1236677.
- [14] McMorris, R. F., Boothroyd, R. A., 1993. *Tests That Teachers Build: An Analysis of Classroom Tests in Science and Mathematics*. *Applied Measurement in Education*. DOI: 10.1207/s15324818ame0604_5.
- [15] Sato, M., Chung, R., Darling-Hammond, L., 2008. *Improving Teachers' Assessment Practices Through Professional Development: The Case of National Board Certification*. *American Educational Research Journal*. DOI: 10.3102/0002831208316955.
- [16] Koh, K.H., 2011. *Improving teachers' assessment literacy through professional development*. *Teaching Education*, 22, 255-276.
- [17] Nguyễn Thị Việt Nga, 2016. *Hình thành cho SVSP kỹ năng ĐG năng lực khoa học của HS theo quan điểm PISA trong dạy HS học ở trường phổ thông*, LATS giáo dục học, ĐHSPHN.
- [18] Messick, S., 1995. *Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning*. *American Psychologist*, 50(9), 741-749. <http://dx.doi.org/10.1037/0003-066X.50.9.741>.
- [19] Brookhart, S. M., 2013. *How to create and use rubrics for formative assessment and grading*. ASCD.
- [20] Moskal, B.M. and Leydens, J.A., 2000. *Scoring rubric development: validity and reliability*. *Practical Assessment, Research & Evaluation*, Vol. 7, pp. 71-81.
- [21] Reddy, M.Y., 2011. *Design and development of rubrics to improve assessment outcomes*. *Quality Assurance in Education*. Vol. 19, No. 1, pp. 84-104.
- [22] Bonett, D.G., 2002. *Sample size requirements for estimating intraclass correlations with desired precision*. *Statistics in Medicine*, Vol. 21, pp. 1331-5.
- [23] Walter, S.D., Eliasziw, M. and Donner, A., 1998. *Sample size and optimal designs for reliability studies*. *Statistics in Medicine*, Vol. 17, No. 1, pp. 101-110.

ABSTRACT

A tool to measure competence of designing assessment competency of teachers and pre-service teachers

Nguyen Thi Dieu Linh and Do Huong Tra

Faculty of Physics, Hanoi National University of Education

The purpose of this article is to introduce the construction of a reliable tool to measure the competence of designing competency assessment tool to foster this competence for teachers and pre-service teachers. The paper is a part of a series of researches on the competence of designing competency assessment tool for teachers and pre-service teachers. The tool is designed to measure teachers and pre-service teachers according to the competency model of the competence of designing competency assessment tool with four elements, namely: Determine the purposes and objectives of the competency assessment activities; Plan the development of the assessment tool; Develop assessment tools; and Trial and edit the assessment tool. The tool consists of a test in which some situations require teachers or pre-service teachers to perform all the behaviors in the model of competence of designing competency assessment tool and rubric to evaluate their work. The process of developing the tool presented in the article also includes preliminary validity check and reliability check. The research results can be used in the assessment or training process of pre-service teachers and develop teachers of assessment competence.

Keywords: Measurement tools, competency assessment, develop competence, training pre-service teacher fostering, competence of designing competency assessment tool.