

## QUẢN LÝ CÔNG TÁC KHẢO THÍ THEO TIẾP CẬN ĐÀO TẠO DỰA TRÊN CHUẨN ĐẦU RA Ở BẬC ĐẠI HỌC

Nguyễn Thanh Vương

Trường Đại học Văn Lang  
Email: thanhvuonggd@gmail.com

### Article history

Received: 28/7/2023

Accepted: 23/8/2023

Published: 05/11/2023

### Keywords

Testing, testing management,  
outcomes-based education

### ABSTRACT

Improving testing and assessment quality is one of the important goals of every higher education institution in order to achieve accurate and objective evaluation of learners' performance. Based on the methods of synthesizing and analyzing documents, the article systematized domestic and international studies on the management of testing and assessment by approaching certain aspects of the organization of assessment and testing at the tertiary level. The result shows that very few studies have thoroughly and systematically addressed the issue of testing and assessment management in higher education institutions today, especially according to the outcomes-based education approach. Therefore, the author proposes some in-depth research directions to better understand the situation of this issue.

### 1. Mở đầu

Trong hơn hai thập niên trở lại đây, nhiều xu hướng và mô hình giáo dục hiện đại đã được Bộ GD-ĐT chủ trương áp dụng rộng rãi tại các cơ sở giáo dục đại học (GDĐH) ở Việt Nam. Một trong những phương thức tiếp cận đã trở nên chiến lược là Giáo dục dựa trên chuẩn đầu ra (Outcomes-based Education - OBE). Spady (1994) nhấn mạnh đặc điểm cốt lõi của tiếp cận OBE là hướng tới vai trò trung tâm của người học và yêu cầu xác định rõ các chuẩn đầu ra (CDR) về kiến thức, kỹ năng, thái độ mà người học cần đạt sau khi hoàn thành chương trình đào tạo (CTĐT); theo đó cơ sở GDĐH phải đo lường các CDR mà người học đạt được nhằm chứng minh chất lượng đào tạo của nhà trường (Bộ GD-ĐT, 2021).

Tại Việt Nam, đã có nhiều nghiên cứu hình thành các lý luận về hoạt động kiểm tra, đánh giá (KT, ĐG) kết quả học tập của người học, tuy nhiên chưa có nhiều công trình chuyên sâu nhằm giải quyết những vấn đề liên quan đến quản lý và đánh giá hiệu quả công tác khảo thí, đặc biệt theo hướng tiếp cận OBE. Trong bài báo này, tác giả tổng hợp, hệ thống hoá các nghiên cứu khác nhau ở trong nước và quốc tế có liên quan đến quản lý công tác khảo thí bậc đại học, từ đó gợi mở một số nội dung cần nghiên cứu chuyên sâu tại Việt Nam.

### 2. Kết quả nghiên cứu

#### 2.1. Sự hình thành và phát triển của quản lý công tác khảo thí

Ngày nay, công tác khảo thí ở mỗi quốc gia được triển khai đa dạng, đặc biệt trong công tác tuyển sinh đại học. Nhiều nước trên thế giới đã sử dụng ngân hàng câu hỏi trắc nghiệm khách quan để tuyển sinh đầu vào đại học, cụ thể: Tại Hoa Kỳ, các trường đại học không tổ chức thi tuyển mà dựa vào kết quả thi do các công ty độc lập tổ chức để xét tuyển như: kì thi SAT (Scholastic Assessment Test) do công ty ETS tổ chức và kì thi ACT (American College Testing) với hình thức câu hỏi hầu hết là trắc nghiệm. Tại Nhật Bản, Trung tâm Quốc gia về Tuyển sinh Đại học (The National Center Test for University Admissions) chịu trách nhiệm tổ chức kì thi tuyển sinh chung hằng năm cho hầu hết các trường đại học công lập và tư thục; thí sinh lựa chọn các môn thi theo 28 khối thi đều ở hình thức trắc nghiệm khách quan và được đo lường để đảm bảo tính công bằng, độ tin cậy và hợp lệ theo quy định của Bộ Giáo dục Nhật Bản. Thái Lan cũng tổ chức các kì thi đánh giá chung trên toàn quốc (như ONET, ANET) để lấy kết quả xét tuyển vào các trường đại học; mỗi kì thi đều có một hội đồng ra đề thi để mỗi thí sinh làm một đề thi riêng với khoảng 80% - 90% là câu hỏi trắc nghiệm.

Tại Việt Nam, từ những năm 90 của thế kỉ XX, vấn đề KT, ĐG trong nhà trường bắt đầu được nghiên cứu nhiều hơn, chủ yếu tập trung ở các hội thảo bàn về hình thức KT, ĐG trong các kì thi và nhấn mạnh tính ưu việt của phương pháp trắc nghiệm khách quan. Về mặt quản lý nhà nước, năm 2003, Cục Khảo thí và Kiểm định chất lượng (nay là Cục Quản lý chất lượng, Bộ GD-ĐT) được thành lập, khi ấy cụm từ "khảo thí" trở nên phổ biến và vai trò công tác tổ chức thi cử được chú trọng nhiều hơn. Nhằm hoàn chỉnh hệ thống tổ chức, trong Chỉ thị về nhiệm vụ của toàn ngành năm học 2004 - 2005, Bộ GD-ĐT (2004) yêu cầu các trường đại học và cao đẳng khẩn trương xây dựng và hoàn thiện về tổ chức, bộ máy và triển khai hoạt động của hệ thống khảo thí. Theo đó, các quy chế, thông tư cũng được Bộ ban hành qua từng giai đoạn để hướng

dẫn thực hiện KT, ĐG, tổ chức thi ở bậc đại học như: Quy chế đào tạo đại học, cao đẳng năm 2006, 2007, 2012, 2014, 2021; Hướng dẫn về xác định khối lượng kiến thức tối thiểu và năng lực người học đạt được sau khi tốt nghiệp năm 2015, 2021; Bộ tiêu chuẩn đánh giá chất lượng CTĐT ở các trình độ giáo dục năm 2016 và năm 2020 (đối với CTĐT từ xa). Nghiên cứu các văn bản trên cho thấy một bức tranh toàn cảnh về tiến trình phát triển công tác khảo thí trong GDĐH ở Việt Nam với những thay đổi, cập nhật qua các giai đoạn, đặc biệt theo Thông tư số 17/2021/TT-BGDĐT chỉ rõ cơ sở GDĐH cần phải tổ chức KT, ĐG và đo lường mức độ đạt được CĐR của người học để làm cơ sở cải tiến chất lượng CTĐT. Tuy nhiên, các văn bản trên chỉ mang tính chất quy định chung, thực tế triển khai còn phụ thuộc nhiều vào mức độ quan tâm và đầu tư của mỗi cơ sở GDĐH trong việc tổ chức, đánh giá và cải tiến công tác thi cử.

## 2.2. Các yêu cầu về tổ chức công tác khảo thí

Bộ Quy tắc trách nhiệm nghề nghiệp của Hội đồng Quốc gia về Đo lường trong giáo dục (1995) ở Hoa Kỳ yêu cầu người làm công tác khảo thí phải hiểu rõ các thủ tục, quy định và tiêu chuẩn trước khi thực hiện KT, ĐG người học, không thực hiện bất kỳ hành vi nào làm ảnh hưởng đến tính giá trị và hiệu lực của kết quả đánh giá. Hiệp hội Tâm lý Hoa Kỳ (American Psychological Association - APA, 2014) xây dựng Bộ tiêu chuẩn Khảo thí trong Giáo dục và Tâm lý cung cấp các tiêu chí và hướng dẫn cụ thể để triển khai hoạt động KT, ĐG, trong đó tính công bằng (fairness) là chủ đề xuyên suốt của bộ tiêu chuẩn. Nghiên cứu mô hình đánh giá kết quả học tập dựa trên CĐR tại một cơ sở GDĐH, Crysta (2008) đề nghị cần có hệ thống phản hồi kết quả học tập của người học đến các bên liên quan; việc đánh giá phải kết hợp với các yêu cầu giáo dục tổng quát của nhà trường; các công cụ đo lường cần được tiêu chuẩn hoá theo từng trình độ đào tạo khác nhau. Bộ quy tắc chất lượng do Cơ quan Đảm bảo chất lượng GDĐH của Anh chỉ rõ các hoạt động mang yếu tố rủi ro tiềm ẩn có thể ảnh hưởng đến kết quả thi cử là: các trường hợp đề thi, bài làm của người học được vận chuyển ra khỏi khuôn viên của cơ sở GDĐH; công tác coi thi của giám thị; công tác xác nhận danh tính của người dự thi tại phòng thi; KT, ĐG theo hình thức trực tuyến... (Quality Assurance Agency for Higher Education - QAA, 2018).

Ở Việt Nam, nhiều tác giả đã nghiên cứu sâu về các phương pháp, kỹ thuật và công cụ ứng dụng trong đổi mới công tác KT, ĐG. Trong đó, Nguyễn Công Khanh và Đào Thị Oanh (2016) chỉ ra rằng các phương pháp và công cụ đo lường kết quả học tập của người học phải dựa trên các nguyên tắc về tính phù hợp, khách quan, công bằng, toàn diện, công khai; đặc biệt, tác giả đề cập đến cách thức phản hồi kết quả KT, ĐG nhằm hỗ trợ, cải tiến chất lượng học tập của người học. Sái Công Hồng và cộng sự (2017) đưa ra quy trình triển khai KT, ĐG theo hướng tiếp cận năng lực người học gồm: lập kế hoạch, xây dựng công cụ đo lường; xây dựng tiêu chí KT, ĐG; thiết kế phương pháp đánh giá; tổ chức KT, ĐG; phân tích và sử dụng kết quả... Bên cạnh đó, một số ít nghiên cứu cũng đề cập tới cách thức tổ chức hoạt động khảo thí gắn với cách tiếp cận OBE trong cơ sở GDĐH. Theo Cẩn Thị Thanh Hương và Lê Đức Ngọc (2010), cần chuyên môn hoá bộ phận khảo thí trong quản lý chung hoạt động KT, ĐG tại cơ sở GDĐH.

Nguyễn Thanh Sơn (2015) cho rằng cần đảm bảo đội ngũ CBQL, giảng viên (GgV) hiểu rõ mục đích, nội dung, phương pháp tiến hành KT, ĐG theo năng lực; đặc biệt đề thi, kiểm tra phải được xây dựng theo đúng quy trình, hướng tới việc đạt được CĐR đã xác định. Dương Thế Việt (2020) đề xuất các giải pháp về KT, ĐG nhằm đảm bảo người học đạt được CĐR bao gồm: nâng cao nhận thức cho CBQL, GgV, người học về tổ chức KT, ĐG; xây dựng và phát triển CĐR cho từng môn học/học phần; xây dựng ngân hàng đề thi; cách thức tổ chức thi để đạt CĐR; tổ chức tranh tra, đánh giá việc kiểm tra, thi cử. Liên quan đến triển khai công tác kiểm tra, thi cử trực tuyến, Đinh Thành Việt và Trần Thị Hà Vân (2022) yêu cầu cần có các quy định, biểu mẫu và hướng dẫn người học thực hiện cam kết liêm chính học thuật cùng với công bố các biện pháp xử lý kỉ luật tương ứng; cần kiểm tra các điều kiện kĩ thuật, hệ thống quản lý học tập (LMS) phục vụ đánh giá trực tuyến, trong đó cần nhắc phần mềm giám thị trực tuyến (Online Proctoring).

## 2.3. Cách thức thiết kế và đánh giá chất lượng câu hỏi thi/đề thi

Kết quả đánh giá, thi cử được sử dụng với đa dạng mục đích như: đánh giá đỗ/trượt, khen thưởng hay dùng để định hướng học tập cho người học hoặc cải thiện quá trình giảng dạy của GgV. Tuy nhiên, muốn sử dụng kết quả thi cho những mục đích này, trước hết cần đảm bảo đề thi đáp ứng các tiêu chuẩn chuyên môn và kĩ thuật về đo lường. Một số chỉ số giúp đảm bảo tính giá trị của đề thi như: độ giá trị, độ tin cậy, độ phân biệt, độ khó,... Cụ thể:

(1) Độ giá trị (Validity) là “tương quan giữa điểm thi thực tế và điểm tiêu chí”, được sử dụng để trả lời câu hỏi: đề thi dự báo như thế nào cho những tiêu chí đánh giá. Độ giá trị gồm bốn loại: độ giá trị đồng thời, độ giá trị dự báo, độ giá trị nội dung và độ giá trị cấu trúc (Finch & French, 2019).

(2) Độ tin cậy (Reliability) phản ánh mức độ nhất quán trong kết quả của các lần thi. Độ tin cậy càng cao có nghĩa là các câu hỏi trong bài thi có xu hướng kết nối với nhau, khi đó người học trả lời đúng câu hỏi này có xu hướng trả lời đúng câu hỏi khác trong cùng nhóm. Phương pháp đo lường độ tin cậy thông thường là chỉ số tin cậy Cronbach's Alpha (Bai & Ola, 2017).

(3) Độ khó (Difficulty) của câu hỏi được sử dụng nhằm xác định liệu người thi có nắm được nội dung hoặc có khả năng thực hiện câu hỏi hay không, từ đó phân biệt những người thi có năng lực cao và người thi có năng lực thấp. Độ khó được tính dựa trên phần trăm người thi trả lời đúng câu hỏi.

(4) Độ phân biệt (Discrimination) của câu hỏi là khả năng phân biệt năng lực của người thi dựa trên việc họ trả lời đúng hay sai câu hỏi. Một câu hỏi có độ phân biệt cao khi người học trong nhóm giỏi (điểm thi cao) đều làm đúng và không có người học trong nhóm yếu (điểm thi thấp) trả lời đúng; ngược lại, độ phân biệt thấp khi câu hỏi quá khó hoặc quá dễ (Bai & Ola, 2017).

Việc xây dựng và phân tích giá trị các câu hỏi thi có hai cách tiếp cận chính, gồm Lí thuyết khảo thí cổ điển (Classical Test Theory - CTT) và Lí thuyết ứng đáp câu hỏi (Item Response Theory - IRT). Theo Muñoz (2010), ứng dụng của CTT là phân tích câu hỏi trắc nghiệm để làm tăng chất lượng của chúng thông qua phương pháp chuyên gia (hỏi ý kiến nhóm chuyên gia về mức độ tương thích và phù hợp giữa những câu hỏi thành phần với mục tiêu đánh giá) và phương pháp định lượng (phân tích các thuộc tính của câu hỏi để biết được những câu hỏi tốt và chưa tốt như: (1) Chỉ số mô tả sự phân bố trả lời của thí sinh về một câu hỏi cụ thể; (2) Chỉ số mô tả mối quan hệ giữa bài làm của thí sinh về một câu hỏi và tiêu chí đánh giá; (3) Chỉ số liên quan đến phương sai của câu hỏi). IRT có tính hiện đại hơn khi các thuộc tính của câu hỏi và năng lực của người học được tách riêng trong việc phân tích câu hỏi trắc nghiệm khách quan. Wu & Adams (2007) cho rằng ý tưởng chính của IRT là việc dự đoán xác suất trả lời đúng một câu hỏi dựa trên chỉ số về năng lực của người trả lời và độ khó của câu hỏi. Theo Sái Công Hồng và cộng sự (2017), sự khác nhau chủ yếu ở hai lí thuyết trên là để đánh giá đối tượng nào đó, CTT tiếp cận ở cấp độ một đề kiểm tra, còn IRT tiếp cận ở cấp độ một câu hỏi; nhóm tác giả cũng cho rằng, việc ứng dụng các lí thuyết khảo thí trong phân tích và đánh giá kết quả câu hỏi/đề thi là điều cần thiết, giúp GgV thấy được những điểm cần điều chỉnh, từ đó nâng cao chất lượng hoạt động KT, ĐG kết quả học tập người học.

#### **2.4. Xu hướng ứng dụng công nghệ thông tin trong công tác khảo thí**

Đẩy mạnh việc ứng dụng công nghệ thông tin vào hoạt động KT, ĐG là một trong những xu hướng mới về đo lường kết quả học tập hiện nay. Chủ đề này được một số tác giả quan tâm và đề xuất hướng các ứng dụng vào công tác tổ chức cũng như phân tích kết quả KT, ĐG.

##### **2.4.1. Ứng dụng công nghệ thông tin trong tổ chức kiểm tra, đánh giá**

Binkley và cộng sự (2012) mô tả chiến lược ứng dụng công nghệ thông tin làm cho hoạt động KT, ĐG truyền thống trở nên hiệu quả hơn khi đảm bảo sẵn sàng cung cấp các bài kiểm tra bất cứ khi nào, khả năng tái sử dụng câu hỏi, hay thuận tiện trong quản lí hệ thống bài kiểm tra; trên hết, phương thức này thay đổi cách thức tổ chức hoặc phát triển các định dạng mới để đánh giá các năng lực vốn khó thực hiện khi ở định dạng truyền thống như năng lực giao tiếp, làm việc nhóm, sáng tạo và đổi mới. Dolan & Burling (2012) chỉ ra cách thức áp dụng công nghệ số vào đánh giá quá trình (formative), thay vì trước đó chủ yếu sử dụng để đánh giá tổng kết (summative) thông qua hệ thống phản hồi của người học và hệ thống học tập trực tuyến. Trong bối cảnh đại dịch Covid-19, việc chuyển dịch hoạt động đào tạo từ hình thức trực tiếp sang trực tuyến, bao gồm việc tổ chức các hoạt động đánh giá, thi cử được Bilen & Matros (2020) chỉ ra những hạn chế của hình thức thi cử trực tuyến làm ảnh hưởng đến tính liêm chính trong học thuật, cụ thể: việc bố trí giám thị cho các kì thi trực tuyến không làm giảm tỉ lệ người học gian lận so với hình thức thi truyền thống; giám thị gặp khó khăn trong việc chứng minh có gian lận xảy ra vì chỉ thu được bằng chứng gián tiếp, đồng thời có nguy cơ phạt nhầm thí sinh; các điều kiện về kĩ thuật và áp lực tâm lí cũng ảnh hưởng đến kết quả thi của thí sinh...

##### **2.4.2. Ứng dụng công nghệ thông tin trong phân tích kết quả đánh giá, thi cử**

Nhằm hạn chế sự bất tiện của việc triển khai thủ công các thống kê trong thiết kế và phân tích chất lượng câu hỏi thi, nhiều phần mềm đã được thiết kế để hỗ trợ công tác này. Phổ biến nhất là phần mềm Quest được Hội đồng Nghiên cứu Giáo dục Úc (ACER) thiết kế đầu tiên vào năm 1983, cung cấp một phạm vi linh hoạt và toàn diện về các mô hình IRT bao gồm: (1) Thông số cơ bản của việc phân tích câu; (2) Các tham số liên quan đến độ khó, độ phân biệt, độ phỏng đoán, độ tin cậy của đề thi; (3) Sự phân bố độ khó của câu hỏi với năng lực của thí sinh; (4) Đường cong đặc tính câu hỏi. Một số phần mềm khác cũng được nghiên cứu và áp dụng như: TAP (Test Analysis Program) do nhóm tác giả của Đại học Ohio (Hoa Kỳ) thiết kế vào năm 2003; công cụ phân tích chất lượng câu hỏi dùng trong nội bộ trường để áp dụng vào công tác đào tạo lập trình máy tính do nhóm tác giả của Đại học Virginia State (Hoa Kỳ) phát triển năm 2017 (Bai & Ola, 2017). Các phần mềm này đều cung cấp các chức năng: (1) Phân tích bài kiểm tra của từng cá nhân và từng nhóm người học; (2) Phân tích câu hỏi bao gồm độ khó, hệ số tương quan point-biserial, độ phân biệt và các thông số thống kê khác; (3) Phân tích độ khó mỗi câu ở nhóm cao và nhóm thấp cho câu trả lời đúng và khả năng nhầm đáp án. Tại Việt Nam có phần mềm VITESTA dùng để đánh giá câu hỏi và

phân tích đề thi do Lâm Quang Thiệp và Edutech Việt Nam thiết kế theo IRT. Đặc điểm và tính năng chính của phần mềm VITESTA là ước lượng các tham số của câu hỏi và năng lực của thí sinh; trong tiến trình ước lượng, các tham số của câu hỏi trắc nghiệm theo CTT cũng được tính toán (Lâm Quang Thiệp, 2010).

### **2.5. Triển khai chấm điểm và phản hồi kết quả thi**

Theo APA (2014), công việc chấm điểm bài thi cần được thiết lập thành các quy trình, tiêu chuẩn và theo tiêu chí (rubrics); nếu thực hiện chấm điểm bằng máy tính thì độ chính xác của thuật toán và các bước thực hiện phải được ghi nhận và lưu trữ. Sadler (2009) yêu cầu phải có các hướng dẫn, tiêu chí cụ thể để tăng tính khách quan và thuận tiện trong việc chấm điểm và phản hồi kết quả đến người học; tuy vậy, Sadler cũng cho rằng trên thực tế việc chấm điểm dựa trên tiêu chí vẫn chủ yếu phụ thuộc vào quyết định chủ quan của các GgV. Kết quả khảo sát 162 GgV của Norton và cộng sự (2012) ở Anh cho thấy, phần lớn GgV không được đào tạo, huấn luyện về công tác chấm điểm cũng như xu hướng đồng ý cao rằng việc chấm điểm là một quá trình chủ quan mà dựa vào nền tảng chuyên môn của họ hơn là dựa vào một khoa học chính xác.

Về việc phản hồi kết quả học tập, Race (2001) chỉ ra phản hồi nhằm cung cấp cho người học sự khác biệt giữa mục tiêu mong đợi với thành tích đạt được nhằm cải thiện hoạt động học tập và xây dựng sự tự tin của người học trong suốt quá trình học. Hermesen và cộng sự (2016) đánh giá phản hồi trực tiếp (face-to-face) sẽ mang lại hiệu quả cao hơn so với việc phản hồi chung chung, không trực tiếp. Bộ tiêu chuẩn của APA (2014) yêu cầu khi thông tin về điểm kiểm tra được công bố, những người chịu trách nhiệm về kết quả đánh giá phải cung cấp các giải thích phù hợp với người học; các giải thích phải mô tả bằng ngôn ngữ đơn giản về nội dung bài kiểm tra, mức điểm, độ chính xác/độ tin cậy của điểm và cách sử dụng kết quả. Bộ Quy tắc chất lượng của Anh (QAA, 2018) yêu cầu công tác phản hồi kết quả đánh giá phải giúp người học hiểu được điểm mạnh và điểm hạn chế của bản thân trước đây và thông tin giúp người học nhận ra cách thức cải thiện hiệu quả học tập trong tương lai; đồng thời tạo cơ hội để người học tham gia đối thoại trên cơ sở ý kiến phản hồi nhận được.

### **2.6. Đề xuất nội dung nghiên cứu về quản lý công tác khảo thí tại các cơ sở giáo dục đại học ở Việt Nam**

Tại Việt Nam có khá nhiều nghiên cứu, đề tài liên quan đến các hoạt động/phương thức KT, ĐG nói chung (đứng ở góc độ đơn vị quản lý đào tạo, khoa chuyên môn, GgV), tuy nhiên, các nghiên cứu thực tiễn về quản lý công tác khảo thí hiện nay còn khá ít. Kết quả điều tra tại 5 cơ sở GDĐH của Trần Đức Hiếu (2014) cho thấy, việc triển khai quản lý “tổ chức thi kết thúc học phần”, “tổ chức coi thi”, “tổ chức chấm thi” ở các trường được đánh giá thực hiện tốt; trong khi đó công tác “bồi dưỡng kỹ năng đánh giá kết quả học tập cho CBQL và GgV”, “chỉ đạo lựa chọn các hình thức KT, ĐG phù hợp CĐR” còn nhiều hạn chế.

Có thể thấy, các nghiên cứu về công tác khảo thí hiện chủ yếu tiếp cận ở phương diện kỹ thuật hoặc công cụ triển khai riêng lẻ; phương diện tổ chức hay các giải pháp theo từng khía cạnh của công tác khảo thí mà chưa kết nối thành một công tác thống nhất, mang tính khái quát để thấy được hiệu quả tổ chức/phối hợp thực hiện trong công tác khảo thí tại các cơ sở GDĐH, từ đó có thể đưa ra những giải pháp khoa học, khả thi nhằm phát triển công tác này. Để nghiên cứu chuyên sâu về quản lý công tác khảo thí hiện nay tại các cơ sở GDĐH ở Việt Nam theo tiếp cận OBE, tác giả đề xuất 03 nội dung nghiên cứu chính như sau:

*Thứ nhất*, nghiên cứu thực trạng triển khai đào tạo theo tiếp cận OBE tại các cơ sở GDĐH, trong đó tập trung tìm hiểu hoạt động thiết kế mục tiêu và CĐR theo từng cấp độ, hoạt động thiết kế và triển khai nội dung đào tạo, hoạt động giảng dạy - học tập và hoạt động KT, ĐG hướng đến việc đạt được các CĐR.

*Thứ hai*, nghiên cứu thực trạng triển khai công tác khảo thí theo tiếp cận OBE tại các cơ sở GDĐH xoay quanh ba chức năng và nhiệm vụ mà công tác khảo thí được quy định hiện nay như: hoạt động xây dựng ngân hàng câu hỏi và kiểm soát chất lượng câu hỏi thi, đề thi; hoạt động tổ chức kiểm tra, thi cử, công bố kết quả; và hoạt động sử dụng kết quả kiểm tra, thi cử để cải tiến chất lượng.

*Thứ ba*, nghiên cứu về quản lý công tác khảo thí theo tiếp cận OBE tại các cơ sở GDĐH theo quy trình PDCA (Lập kế hoạch (Plan) - Thực hiện (Do) - Kiểm tra (Check) - Đưa vào ứng dụng (Act)). Trong đó, tìm hiểu chuyên sâu về: xác định mục tiêu/chiến lược cho công tác khảo thí; quản lý hoạt động xây dựng ngân hàng câu hỏi và phân tích chất lượng đề thi, câu hỏi thi; quản lý hoạt động tổ chức đánh giá, thi cử; quản lý việc sử dụng kết quả đánh giá, thi cử để cải tiến chất lượng; và quản lý các điều kiện hỗ trợ cho công tác khảo thí.

## **3. Kết luận**

Thông qua việc tìm hiểu các nghiên cứu trong nước và nước ngoài, bài viết đã cho thấy còn thiếu nhiều nghiên cứu về lý luận và thực tiễn quản lý công tác khảo thí tại các cơ sở GDĐH ở Việt Nam, đặc biệt gắn với tiếp cận OBE theo xu hướng đào tạo hiện nay. Do đó, tác giả đề xuất các hướng nghiên cứu chuyên sâu để tìm hiểu về thực trạng

cơ sở GDDH triển khai đào tạo theo OBE, thực trạng triển khai công tác khảo thí theo tiếp cận OBE và quản lý công tác khảo thí theo tiếp cận OBE. Những định hướng này giúp bổ sung cơ sở lý luận, đề xuất các giải pháp nâng cao chất lượng quản lý công tác khảo thí nói riêng và chất lượng đào tạo nói chung ở mỗi cơ sở GDDH.

### Tài liệu tham khảo

- American Psychological Association (2014). *Standards for educational and psychological testing*. American Educational Research Association.
- Asean University Network (2020). *Guide to AUN-QA Assessment at Program level version 4.0*. Bangkok: Chulalongkorn University.
- Bai, X., & Ola, A. (2017). A Tool for Performing Item Analysis to Enhance Teaching and Learning Experiences. *Issues in Information Systems*, 18(1), 128-136. [https://doi.org/10.48009/1\\_iis\\_2017\\_128-136](https://doi.org/10.48009/1_iis_2017_128-136)
- Bilen, E., & Matros, A. (2020). Online Cheating Amid Covid-19. *Journal of Economic Behavior & Organization*, 182, 196-211. <https://doi.org/10.1016/j.jebo.2020.12.004>
- Binkley, M., Erstad, O., Herman, J., Raizen, S., Ripley, M., Miller-Ricci, M. & Rumble, M. (2012). Defining Twenty-First Century Skills. In: Griffin, P., McGaw, B., Care, E. (eds), *Assessment and Teaching of 21st Century Skills*, Springer, 17-66.
- Bộ GD-ĐT (2004). *Quyết định số 38/2004/QĐ-BGDĐT ngày 02/12/2004 về việc ban hành quy định tạm thời về kiểm định chất lượng trường đại học*.
- Bộ GD-ĐT (2021). *Thông tư số 17/2021/TT-BGDĐT ngày 22/6/2021 quy định về chuẩn chương trình đào tạo; xây dựng, thẩm định và ban hành chương trình đào tạo các trình độ của giáo dục đại học*.
- Cần Thị Thanh Hương, Lê Đức Ngọc (2010). Quy trình đánh giá thành quả học tập theo cách tiếp cận chuẩn đầu ra. *Tạp chí Giáo dục*, 251, 15-20.
- Crystal, B. (2008). *Community College of Philadelphia Learning outcomes Assessment Model*. Assessment Task Force of the Academic Master Plan Community College of Philadelphia, USA.
- Dolan, R. P., & Burling, K. S. (2012). Computer-based testing in higher education, In C. Secolsky (Ed.). *Handbook on Measurement, Assessment, and Evaluation in Higher Education* (1st ed), Routledge, 321-335.
- Dương Thế Việt (2020). Giải pháp quản lý kiểm tra, đánh giá kết quả học tập học phần tiếp cận chuẩn đầu ra. *Tạp chí Quản lý giáo dục*, số đặc biệt tháng 10, 47 - 53.
- Đình Thành Việt, Trần Thị Hà Vân (2022). *Phát triển, bảo đảm chất lượng chương trình đào tạo và đo lường đánh giá chuẩn đầu ra*. NXB Thông tin và Truyền thông.
- Finch, W. H., & French, B. F. (2019). *Educational and Psychological Measurement*. NY: Routledge, 171-189.
- Hermesen, S., Frost, J., Renes, R.J., & Kerkhof, P. (2016). Using Feedback through digital technology to disrupt and change habitual behaviour: A critical review of current literature. *Computers in Human Behaviour*, 57, 61-74.
- Lâm Quang Thiệp (2010). *Đo lường trong Giáo dục - Lý thuyết và ứng dụng*. NXB Đại học Quốc gia Hà Nội.
- Muñiz, J. (2010). Test theories: Classical theory and item response theory. *Papeles del Psicólogo*, 31(1), 57-66.
- Norton, L. S., Norton, B., & Sadler, I. (2012). Assessment, marking and feedback: understanding the lecturers' perspective. *Practitioner Researching Higher Education*, 6(2). 3-24. <https://doi.org/10.2307/j.ctv13xprqb>
- Nguyễn Công Khanh, Đào Thị Oanh (2016). *Giáo trình Kiểm tra, đánh giá trong giáo dục*. NXB Đại học Sư phạm.
- Nguyễn Thanh Sơn (2015). Đổi mới kiểm tra, đánh giá kết quả học tập của sinh viên theo hướng tiếp cận năng lực nhằm đáp ứng chuẩn đầu ra. *Journal of Science*, 9(1), 35-40.
- Quality Assurance Agency (2018). *The revised UK quality code for higher education*. <https://www.qaa.ac.uk/quality-code>
- Race, P. (2001). *Using feedback to help students to learn*. New York: The Higher Education Academy.
- Sadler, D. R. (2009). Indeterminacy in the use of preset criteria for assessment and grading. *Assessment & Evaluation in Higher Education*, 34(2), 159-179. <https://doi.org/10.1080/02602930801956059>
- Sái Công Hồng, Lê Thái Hưng, Lê Thị Hoàng Hà, Lê Đức Ngọc (2017). *Giáo trình kiểm tra đánh giá trong dạy học*. NXB Đại học Quốc gia Hà Nội.
- Spady, W. D. (1994). *Outcomes Based Education: Critical Issues and Answers*. Arlington, VA: American Association of School Administration.
- Trần Đức Hiếu (2014). Thực trạng quản lý đánh giá kết quả học tập của sinh viên trong đào tạo theo học chế tín chỉ ở các trường đại học hiện nay. *Tạp chí Giáo dục*, 348, 9-11.
- Wu, M. L., & Adams, R. J. (2007). *Applying the Rasch Model to Psychosocial Measurement: A practical approach*. Melbourne: Educational Measurement Solutions.