

CÁC BÀI THI ĐÁNH GIÁ NĂNG LỰC HỌC SINH TRUNG HỌC PHỔ THÔNG TẠI VIỆT NAM: MỘT NGHIÊN CỨU SO SÁNH

Phạm Hương Thảo¹,
Phùng Thừa Thảo¹,
Nguyễn Bá Tiến¹,
Trần Thị Thu Hương^{2,+}

¹Viện Đào tạo số và Khảo thí - Đại học Quốc gia Hà Nội;

²Trường Đại học Giáo dục - Đại học Quốc gia Hà Nội

+Tác giả liên hệ • Email: huong.tran@vnu.edu.vn

Article history

Received: 13/3/2025

Accepted: 08/5/2025

Published: 20/7/2025

Keywords

Standardized test,
competency-based
assessment, high school
student, testing and
assessment, comparative
study

ABSTRACT

In the context of modern education, the trend of learner competency-based assessment has become more and more popular. Instead of focusing merely on knowledge, this approach focuses on assessing students' ability to apply knowledge, skills and attitudes to solve practical problems, emphasizing students' thinking, creativity, problem solving and teamwork skills. In Vietnam, implementing competency-based tests for high school students is an important step in innovating assessment methods. However, there are currently very few studies on assessing and summarizing implementation experiences as well as the effectiveness of these assessment forms in Vietnam. Therefore, this study was conducted to compare and analyze competency-based tests currently used to assess high school students in Vietnam in a large-scale manner. The results of this study provide useful insights for educational managers, teachers and researchers in developing the student competency assessment system, thereby improving the quality of education.

1. Mở đầu

Trong bối cảnh thế giới biến đổi mạnh mẽ về KT-XH và công nghệ, giáo dục cũng không đứng ngoài xu hướng này. Cuộc Cách mạng công nghiệp lần thứ tư với sự phát triển của trí tuệ nhân tạo (AI), Internet vạn vật (IoT) và các công nghệ mới yêu cầu nguồn nhân lực có khả năng tư duy sáng tạo, giải quyết vấn đề phức tạp và thích ứng nhanh chóng. Đánh giá theo tiếp cận năng lực trở nên thiết yếu hơn bao giờ hết, đặc biệt khi nhiều quốc gia, bao gồm Việt Nam, chuyển từ chương trình học nội dung sang định hướng năng lực (Hoàng Hoà Bình, 2015). Phương pháp đánh giá nội dung, dù tập trung vào kiểm tra kiến thức, không phản ánh đầy đủ năng lực thực tế của người học. Ngược lại, đánh giá năng lực giúp đo lường toàn diện các phẩm chất, kĩ năng và thái độ của HS, khuyến khích sự chủ động và phát triển các năng lực tự học, tự đánh giá và hợp tác. Đây là yêu cầu tất yếu của giáo dục trong thời đại mới (Lê Thái Hưng và Nguyễn Thái Hà, 2021), giúp tạo ra những người học không chỉ có kiến thức mà còn đủ năng lực đối mặt với thách thức của thế kỉ XXI.

Tại Việt Nam, các bài thi đánh giá năng lực đang nhận được sự quan tâm lớn từ xã hội (Bùi Đức Nhân, 2023). Bộ GD-ĐT đã ban hành nhiều văn bản pháp lí hướng dẫn đổi mới phương thức kiểm tra, trong đó nhấn mạnh việc đánh giá năng lực HS. Những văn bản như Quyết định số 3538/QĐ-BGDĐT (Bộ GD-ĐT, 2014) và Công văn số 1496/BGDĐT-QLCL (Bộ GD-ĐT, 2022) thể hiện sự chuyển dịch từ đánh giá kiến thức sang đánh giá năng lực, tạo hành lang pháp lí cho các trường đại học triển khai kì thi đánh giá năng lực. Đề án thi đánh giá năng lực của Đại học Quốc gia (ĐHQG) Hà Nội là một bước tiến quan trọng, đánh giá toàn diện năng lực tư duy, giải quyết vấn đề và ngôn ngữ của HS. Nhiều trường đại học đã sử dụng kết quả kì thi này làm tiêu chí xét tuyển, mở rộng cơ hội cho HS. Hiện nay, Việt Nam có nhiều kì thi đánh giá năng lực do các trường đại học và tổ chức giáo dục tổ chức, bao gồm ĐHQG Hà Nội, ĐHQG TP. Hồ Chí Minh, Đại học Bách khoa Hà Nội và Bộ Công an. Dù kì thi tốt nghiệp THPT vẫn chủ yếu đánh giá kiến thức, Bộ GD-ĐT đã định hướng chuyển sang đánh giá năng lực với câu hỏi mở và vận dụng để đánh giá tư duy sáng tạo.

Mặc dù đã có một số nghiên cứu về đánh giá năng lực HS THPT tại Việt Nam, song vẫn còn tồn tại khoảng trống nghiên cứu đáng kể. Hiện chưa có một nghiên cứu đồng bộ, tổng thể nào được thực hiện nhằm so sánh một cách toàn diện các phương thức kiểm tra, đánh giá theo tiếp cận năng lực người học đang được triển khai tại các trường THPT trên cả nước, từ đó đưa ra những đánh giá khách quan, khoa học về ưu nhược điểm của từng phương thức, cũng như

đề xuất các giải pháp để hoàn thiện và nâng cao hiệu quả của công tác đánh giá năng lực HS. Bài báo này nhằm tổng quan và so sánh các bài thi đánh giá năng lực hiện hành, sử dụng phương pháp nghiên cứu định tính để phân tích các tài liệu nghiên cứu, bài báo khoa học trong và ngoài nước về các khái niệm liên quan đến đánh giá và tiếp cận năng lực. Các từ khóa như “năng lực”, “đánh giá năng lực”, “competence”, “competence-based assessment” sẽ được sử dụng để làm cơ sở phân tích. Đồng thời, bài báo cũng đối chiếu các đặc điểm của các bài thi từ các đơn vị tổ chức, tham khảo từ các nguồn dữ liệu uy tín như Science Direct, Scopus và Google Scholar. Nghiên cứu này không chỉ mang tính cấp thiết mà còn đóng góp vào việc xây dựng cơ sở lý luận cho khoa học giáo dục và khảo thí.

2. Kết quả nghiên cứu

2.1. Một số vấn đề lý luận

2.1.1. Khái niệm về đánh giá năng lực

“Đánh giá theo tiếp cận năng lực” là phương pháp đánh giá tập trung vào việc đo lường khả năng vận dụng kiến thức, kỹ năng và thái độ của người học để giải quyết các tình huống thực tế, thay vì chỉ kiểm tra mức độ ghi nhớ và tái hiện kiến thức. Glaser và cộng sự (2001) cho rằng đánh giá năng lực là quá trình thu thập và phân tích dữ liệu về cách người học áp dụng kiến thức và kỹ năng vào các tình huống thực tiễn. Nghiên cứu của OECD (2002), đánh giá năng lực không chỉ kiểm tra kiến thức mà còn đo lường khả năng sử dụng kiến thức đó trong ngữ cảnh thực tế. Škrinjaric (2022) khẳng định đánh giá năng lực phải phản ánh được khả năng của người học trong bối cảnh thực tiễn, không chỉ là sự ghi nhớ hay nhận diện thông tin. Như vậy, đánh giá theo tiếp cận năng lực nhằm kiểm tra khả năng vận dụng kiến thức vào thực tiễn, giúp người học phát triển tư duy và kỹ năng giải quyết vấn đề khác với đánh giá theo tiếp cận nội dung nhằm kiểm tra mức độ hiểu biết và ghi nhớ kiến thức được dạy trong chương trình học. Đối với cách tính điểm bài thi của đánh giá theo cách tiếp cận đánh giá năng lực ngoài việc sử dụng điểm số còn có cách tính điểm theo tiêu chí để ghi nhận mức độ thành thạo của người học, rất khác so với cách tính điểm chỉ có đúng sai của đánh giá theo cách tiếp cận nội dung.

2.1.2. Phương pháp tính kết quả thi

“Điểm thô” (raw score) là tổng điểm mà thí sinh nhận được dựa trên số câu hỏi trả lời đúng khác với điểm thang đo (scale score) - điểm số được chuyển đổi toán học từ điểm thô sang một bộ số khác để có thể so sánh được một cách công bằng, ví dụ như giữa các phiên bản hoặc các dạng bài thi khác nhau nhưng có cấu trúc tương tự. Khi báo cáo kết quả thi, một trong những quyết định quan trọng là lựa chọn phương thức báo cáo điểm - theo điểm thô hay điểm thang đo. Trước đây, hầu hết các kì thi đều sử dụng điểm thô để tính kết quả. Điểm thô phản ánh trực tiếp số lượng câu trả lời đúng của thí sinh mà không có bất kì điều chỉnh hay chuyển đổi nào. Tuy nhiên, điểm thô không luôn cung cấp bức tranh đầy đủ về năng lực của thí sinh vì nó không xem xét các yếu tố định tính như độ khó của câu hỏi hay hiệu suất của thí sinh so với những người tham gia kì thi khác. Để có được sự so sánh, báo cáo về điểm số các bài thi chuẩn hóa đều là điểm thang đo. Điểm số được báo cáo có được bằng cách điều chỉnh thống kê và chuyển đổi điểm số thô thành thang điểm chung để giải thích cho sự khác biệt về độ khó giữa các mã đề thi khác nhau. Đối với mã đề thi dễ hơn, thí sinh cần trả lời chính xác hơn các câu hỏi để có được một số điểm cụ thể. Đối với mã đề thi khó hơn, thí sinh có thể nhận được điểm số tương tự khi trả lời chính xác ít câu hỏi hơn.

Điểm thô dễ tính toán hơn so với điểm thang đo, tuy nhiên, để đảm bảo tính công bằng và chính xác khi so sánh điểm thi, đặc biệt khi thí sinh tham gia làm các mã đề khác nhau vào các ngày khác nhau, điểm thang đo là lựa chọn ưu việt. Điểm thang đo cho phép so sánh chính xác và đảm bảo rằng thí sinh làm đề thi khó hơn sẽ không thiệt và thí sinh làm đề thi dễ hơn sẽ không có lợi thế hơn những thí sinh làm đề khó (Kolen & Brennan, 2014). Trong khi đó, phương pháp tính điểm theo lý thuyết ứng đáp câu hỏi IRT (Item Response Theory - IRT equating) là một quy trình ba bước. Bước đầu tiên là ước tính các tham số của từng câu hỏi. Bước thứ hai là điều chỉnh các tham số này về thang đo IRT cơ sở thông qua phép biến đổi tuyến tính. Bước cuối cùng, việc so sánh điểm thi được thực hiện bằng cách áp dụng các phương pháp khác nhau, chẳng hạn như cân bằng theo phần trăm (Sansivieri et al., 2017).

2.2. So sánh cấu trúc các bài đánh giá năng lực học sinh trung học phổ thông tại Việt Nam hiện nay

Các kì thi đánh giá năng lực hiện nay ngày càng trở nên phổ biến, đóng vai trò quan trọng trong việc tuyển sinh và phân loại thí sinh một cách toàn diện. Tuy nhiên, mỗi bài thi có thể có cấu trúc và phương pháp tiếp cận khác nhau, tập trung vào các lĩnh vực năng lực riêng biệt. Để hiểu rõ hơn về đặc điểm của các bài thi này, cần phân tích và so sánh theo hai khía cạnh quan trọng: (1) So sánh cấu trúc nội dung của các bài thi, giúp nhận diện những điểm chung và sự khác biệt giữa chúng; qua đó có thể đánh giá xu hướng tập trung vào những năng lực nào. Đồng thời, phân tích số lượng tổ hợp/môn trong mỗi bài thi và phương pháp tính điểm (xem bảng 1, bảng 2); (2) Phân tích các loại câu hỏi xuất hiện trong từng bài thi, làm rõ sự khác biệt về tỉ lệ giữa câu hỏi trắc nghiệm khách quan (TNKQ)

và tự luận. Phần này cũng chỉ ra các dạng câu hỏi TNKQ phổ biến như điền từ vào ô trống, chọn đáp án ABCD, viết lại câu,... Từ các so sánh này, chúng ta có thể rút ra những nhận xét quan trọng về xu hướng ra đề, đặc biệt là sự ưu tiên giữa các dạng câu hỏi và mức độ đánh giá năng lực của từng bài thi (xem bảng 3). Trong bài viết này, các tác giả sẽ so sánh 03 bài thi phổ biến nhất hiện nay với bài thi tốt nghiệp THPT. Đó là bài thi Đánh giá năng lực HS THPT (HSA) của ĐHQG Hà Nội, bài thi đánh giá tư duy (TSA) của Đại học Bách khoa Hà Nội và bài thi Đánh giá năng lực (V-ACT) của ĐHQG TP. Hồ Chí Minh.

Bảng 1. So sánh cấu trúc nội dung của các bài thi

Tên bài thi	Mục đích sử dụng kết quả	Mục tiêu đánh giá	Số môn thi/phần thi
Bài thi tốt nghiệp THPT (Bộ GD-ĐT)	Xét công nhận tốt nghiệp THPT và xét tuyển đầu vào đại học.	Đánh giá mức độ đáp ứng chuẩn kiến thức theo Chương trình giáo dục phổ thông 2018. Kiểm tra các năng lực chính: (1) Ghi nhớ và tái hiện kiến thức; (2) Vận dụng kiến thức vào bài tập cơ bản; (3) Vận dụng cao hơn (ở một số câu hỏi khó). Xét công nhận tốt nghiệp THPT và làm cơ sở xét tuyển đại học. Mức độ tư duy đánh giá: Tư duy tái hiện.	4 môn (bắt buộc 2 môn Ngữ văn, Toán, tự chọn 2 trong số các môn Ngoại ngữ, Lịch sử, Vật lý, Hóa học, Sinh học, Địa lý, Giáo dục kinh tế và pháp luật, Tin học, Công nghệ).
Bài thi Đánh giá năng lực (HSA- ĐHQG Hà Nội)	Xét tuyển đầu vào đại học của những cơ sở giáo dục đại học lựa chọn sử dụng kết quả bài thi này. Kết quả bài thi được công nhận tương đương với kết quả bài thi của ĐHQG TP. Hồ Chí Minh.	Đánh giá ba nhóm năng lực cốt lõi theo Chương trình giáo dục phổ thông 2018: (1) Giải quyết vấn đề và sáng tạo; (2) Giao tiếp và hợp tác; (3) Tự chủ và tự học. Ngoài ra, bài thi còn đánh giá các năng lực đặc thù: (4) Năng lực ngôn ngữ (tiếng Việt, tiếng Anh); (5) Năng lực tính toán, lập luận, tư duy logic; (6) Năng lực xử lý số liệu, tin học; (7) Năng lực khoa học tự nhiên/xã hội. Mức độ tư duy đánh giá: Tư duy tái hiện, tư duy suy luận, tư duy bậc cao.	3 phần thi bắt buộc (Toán học, Ngôn ngữ, Khoa học (tự chọn 3 trong tổng số 5 chủ đề: Vật lý, Hóa học, Sinh học, Lịch sử, Địa lý) hoặc Tiếng Anh).
Bài thi Đánh giá tư duy (TSA - Đại học Bách khoa Hà Nội)	Xét tuyển đầu vào đại học của những cơ sở giáo dục đại học lựa chọn sử dụng kết quả bài thi này.	Đánh giá tư duy nền tảng của HS, không kiểm tra kiến thức cụ thể của môn học. Bài thi tập trung vào ba nhóm năng lực chính: (1) Tư duy Toán học: Đánh giá khả năng tư duy định lượng, vận dụng Toán vào thực tế; (2) Tư duy Đọc hiểu: Đánh giá kỹ năng phân tích, lí giải, khái quát, tổng hợp văn bản; (3) Tư duy Khoa học/Giải quyết vấn đề: Đánh giá khả năng vận dụng kiến thức khoa học vào thực tiễn. Mức độ tư duy đánh giá: Tư duy tái hiện, tư duy suy luận, tư duy bậc cao.	3 phần thi bắt buộc (Toán, Đọc hiểu, Khoa học).
Bài thi Đánh giá năng lực (V-ACT- ĐHQG TP. Hồ Chí Minh)	Xét tuyển đầu vào đại học của những cơ sở giáo dục đại học lựa chọn sử dụng kết quả bài thi này. Kết quả bài thi được công nhận tương đương với kết quả bài thi của ĐHQG Hà Nội.	Đánh giá tư duy tổng quát của thí sinh, tránh học lệch, học tủ. Kiểm tra 4 nhóm năng lực quan trọng: (1) Tư duy logic; (2) Tư duy sáng tạo và giải quyết vấn đề; (3) Làm việc độc lập, làm việc nhóm; (4) Phân tích lập luận ngôn ngữ, số học, đại số, hình học, phân tích số liệu. Mức độ tư duy đánh giá: Tư duy tái hiện, tư duy suy luận, tư duy bậc cao.	3 phần thi bắt buộc (Ngôn ngữ (tiếng Việt và tiếng Anh), Phần 2. Toán học, tư duy logic và phân tích số liệu. Phần 3. Giải quyết vấn đề (thuộc các 5 lĩnh vực Hóa học, Vật lý, Sinh học, Địa lý và lĩnh vực Lịch sử, chính trị, xã hội).

Bảng 2. Cấu trúc bài thi và phương pháp chấm điểm

Tên bài thi	Cấu trúc bài thi	Phương pháp chấm điểm
THPT	Gồm 3 phần thi: (1) Ngữ văn (120 phút, tự luận); (2) Toán (90 phút, trắc nghiệm); (3) Các môn khác (50 phút, trắc nghiệm).	Thang điểm 10 điểm/môn. Cộng tổng điểm các môn + điểm học bạ.
HSA	Gồm 3 phần thi: (1) Toán học & xử lý số liệu (75 phút, 50 câu); (2) Ngôn ngữ (60 phút, 50 câu); (3) Khoa học hoặc Tiếng Anh (60 phút, 50 câu).	Thang điểm 150 điểm Cộng tổng điểm các phần.
TSA	Gồm 3 phần thi: (1) Toán học (60 phút, 40 điểm); (2) Đọc hiểu (30 phút, 20 điểm); (3) Khoa học (60 phút, 40 điểm).	Thang điểm 100 điểm Lí thuyết IRT điểm số có trọng số khác nhau tùy độ khó câu hỏi.
V-ACT	Gồm 3 phần thi: (1) Ngôn ngữ (40 câu); (2) Toán, Tư duy logic, Phân tích số liệu (30 câu); (3) Giải quyết vấn đề (50 câu).	Thang điểm 1200 điểm Lí thuyết IRT, điểm số có trọng số khác nhau tùy độ khó câu hỏi.

Bảng 3. Các loại câu hỏi xuất hiện trong từng bài thi

Tên bài thi	Câu hỏi	Hình thức thi, thời gian
THPT	Dạng câu hỏi: Trắc nghiệm 4 lựa chọn; Câu hỏi Đúng/Sai; Câu hỏi trả lời ngắn; Tự luận (chỉ có trong môn Ngữ văn). Không có câu hỏi thử nghiệm trong bài thi. Tính ứng dụng của đề thi: Thấp - Chủ yếu đánh giá kiến thức, ít kiểm tra tư duy bậc cao.	Trắc nghiệm (trừ Ngữ văn) trên giấy trong thời gian từ 50-120 phút.
HSA	Dạng câu hỏi: Trắc nghiệm khách quan 4 lựa chọn; Câu hỏi điền đáp án; Câu hỏi tình huống/đọc hiểu chừa câu nhiều câu hỏi. Có câu hỏi thử nghiệm trong bài thi. Tính ứng dụng của đề thi: Cao - Đánh giá kĩ năng tư duy và phân tích dữ liệu.	Trắc nghiệm trên máy tính trong thời gian 195 phút.
TSA	Dạng câu hỏi: Trắc nghiệm 4 lựa chọn; Câu hỏi điền đáp án; Câu hỏi Đúng/Sai; Câu hỏi chọn nhiều đáp án; Câu hỏi kéo thả đáp án. Không có câu hỏi thử nghiệm trong bài thi. Tính ứng dụng của đề thi: Cao - Đánh giá khả năng lập luận và giải quyết vấn đề thực tiễn.	Trắc nghiệm trên máy tính trong thời gian 150 phút.
V-ACT	Dạng câu hỏi: Trắc nghiệm 4 lựa chọn; Câu hỏi xử lý tình huống thực tế; Câu hỏi đọc hiểu, lập luận logic. Không có câu hỏi thử nghiệm trong bài thi. Tính ứng dụng của đề thi: Cao - Đánh giá khả năng tư duy logic, phân tích dữ liệu, giải quyết vấn đề.	Trắc nghiệm trên giấy trong thời gian 150 phút.

Bảng 4. So sánh phương thức tính kết quả của các bài thi

Tiêu chí	HSA	TSA	V-ACT
Tổng điểm tối đa	150 điểm	100 điểm	1200 điểm
Cách tính điểm	Điểm từng phần cộng lại, mỗi phần tối đa 50 điểm, điểm thô	Tổng điểm từ 3 phần, không quy đổi theo trọng số, sử dụng lí thuyết hồi đáp	Tổng điểm quy đổi theo thang 1200, phân bổ theo độ khó, sử dụng lí thuyết hồi đáp
Điểm từng phần *	- Toán & Xử lý số liệu: 50 câu/50 điểm - Ngôn ngữ & Văn học: 50 câu/50 điểm - Khoa học hoặc Tiếng Anh: 50 câu/50 điểm	- Tư duy Toán học: 40 câu/40 điểm - Tư duy Đọc hiểu: 20 câu/20 điểm - Tư duy Khoa học: 40 câu/ 40 điểm	- Sử dụng ngôn ngữ/60 câu/600 điểm - Toán học: 30 câu/300 điểm - Tư duy khoa học: 30 câu/300 điểm
Cách phân bổ điểm	Tính tổng từ ba phần thi, mỗi câu có trọng số khác nhau	Không có trọng số giữa các phần, điểm tổng hợp từ câu hỏi trắc nghiệm.	Điểm tính theo độ khó câu hỏi, câu khó có trọng số cao hơn
Thang điểm chuẩn hóa	Không chuẩn hóa, tổng điểm là thang 150	Không chuẩn hóa, tổng điểm là thang 100	Chuẩn hóa theo thang 1200 để phù hợp đánh giá năng lực
Công thức tính điểm	Điểm = Σ (Điểm từng phần)	Điểm = Σ (Điểm từng phần)	Điểm = Σ (Điểm từng phần)

Bảng 4 cho thấy, bài thi HSA được tính bằng tổng số các câu làm đúng của cả bài thi, mỗi phần có thang điểm 50, thí sinh làm đúng cả 3 phần sẽ có điểm 150, không quy đổi hay chuẩn hóa. Điểm của bài thi TSA không có trọng số giữa các phần, điểm tổng được tính trực tiếp từ các câu hỏi. Bài thi V-ACT thì điểm có sự chuẩn hóa theo độ khó câu hỏi, giúp đánh giá sát hơn về năng lực HS. Tuy nhiên, cả 02 bài TSA và V-ACT đều không công bố rõ cách tính điểm hay quy đổi điểm về điểm 100 và 1200.

2.3. Thảo luận

Hiện nay, kì thi tốt nghiệp THPT tại Việt Nam là bắt buộc đối với tất cả thí sinh, với kết quả thi chủ yếu được sử dụng để xét công nhận tốt nghiệp và làm cơ sở xét tuyển vào các trường đại học và cao đẳng. Trong khi đó, các kì thi đánh giá năng lực của ĐHQG Hà Nội, ĐHQG TP. Hồ Chí Minh và bài thi đánh giá tư duy của Đại học Bách khoa Hà Nội được thiết kế để đánh giá toàn diện năng lực của thí sinh, phục vụ cho việc xét tuyển vào các chương trình đào tạo đại học của các trường này và các trường đại học khác mà trường có thể tự quyết định lựa chọn. Một yếu tố quan trọng của các bài thi đánh giá năng lực là đảm bảo sự phân hóa hợp lý giữa các mức độ nhận thức của thí sinh. Các bài thi cần có sự cân bằng giữa các câu hỏi kiểm tra khả năng ghi nhớ, tái hiện kiến thức với các câu hỏi đánh giá khả năng suy luận, vận dụng và sáng tạo. Điều này không chỉ giúp đo lường năng lực toàn diện của thí sinh mà còn bảo đảm tính công bằng và khách quan của bài thi. Các kì thi đánh giá năng lực tại Việt Nam như HSA (ĐHQG Hà Nội), TSA (Đại học Bách khoa Hà Nội) hay V-ACT (ĐHQG TP. Hồ Chí Minh) đều có sự phân hóa độ khó, nhưng mức độ này có sự khác biệt giữa các bài thi. Một số kì thi áp dụng lí thuyết IRT để điều chỉnh điểm số theo độ khó, giúp kết quả có giá trị đo lường cao hơn. Tuy nhiên, hiện nay chưa có tiêu chuẩn chung để chuẩn hóa độ khó của các bài thi này, dẫn đến khó khăn trong việc so sánh kết quả giữa các kì thi. Bên cạnh đó, quy trình và chủ thể thiết kế ngân hàng câu hỏi, ra đề thi đóng vai trò quan trọng trong việc đảm bảo chất lượng và tính khách quan của bài thi. Để nâng cao chất lượng các bài thi, cần thành lập cơ quan độc lập chịu trách nhiệm kiểm định đề thi theo các tiêu chuẩn quốc tế về đánh giá giáo dục. Các tiêu chí như độ tin cậy, độ giá trị, tính công bằng và khả năng dự đoán cần được áp dụng nghiêm ngặt hơn. Việc công khai quy trình xây dựng đề thi cũng sẽ giúp tăng cường sự tin tưởng của xã hội vào kết quả thi. Về phương pháp tính điểm và đảm bảo độ tin cậy, hiện nay, cách tính điểm của các bài thi đánh giá năng lực tại Việt Nam chưa đồng nhất. Một số bài thi sử dụng điểm thô, trong khi một số bài thi khác lại sử dụng điểm chuẩn hóa theo độ khó của câu hỏi. Việc áp dụng lí thuyết IRT là một bước tiến quan trọng, giúp đảm bảo rằng điểm số phản ánh chính xác năng lực của thí sinh. Tuy nhiên, cần có hướng dẫn chi tiết và thống nhất về cách tính điểm để tránh nhầm lẫn và đảm bảo công bằng giữa các kì thi. Ngoài ra, một trong những vấn đề quan trọng trong tổ chức các kì thi đánh giá năng lực là lựa chọn hình thức thi phù hợp. Hiện nay, một số bài thi như HSA của ĐHQG Hà Nội và TSA của Đại học Bách khoa Hà Nội được tổ chức trên máy tính, trong khi một số bài thi khác vẫn duy trì hình thức thi trên giấy. Việc lựa chọn hình thức thi, loại câu hỏi thi (TNKQ hay trắc nghiệm tự luận) ảnh hưởng lớn đến độ tin cậy của kết quả, tốc độ trả kết quả cũng như khả năng phân tích kết quả thi trong các kì đánh giá diện rộng.

2.4. Khuyến nghị

Trên cơ sở kết quả nghiên cứu, chúng tôi đề xuất một số khuyến nghị sau:

- *Đối với các đơn vị tổ chức thi:* cần xây dựng bộ công cụ đánh giá năng lực chuẩn hóa, bao gồm các câu hỏi, bài tập, tiêu chí đánh giá, đảm bảo tính khách quan, khoa học và công bằng của các kì thi. Trong đó, một trong các công việc cần triển khai là chuẩn hóa mức độ nhận thức của câu hỏi. Các bài thi cần có sự phân bố hợp lý giữa các mức độ tư duy theo thang Bloom (từ ghi nhớ, hiểu, vận dụng đến phân tích, đánh giá và sáng tạo). Cuối cùng, cần công khai kết quả đánh giá và báo cáo phân tích kết quả để HS, phụ huynh, và xã hội có thể hiểu rõ hơn về năng lực của HS, từ đó có những định hướng phù hợp trong học tập và phát triển.

- *Đối với các đơn vị sử dụng kết quả thi, các đơn vị sử dụng kết quả đánh giá:* cần sử dụng một cách linh hoạt, kết hợp với các thông tin khác (như học bạ, kết quả phỏng vấn, bài luận,...) để có cái nhìn đầy đủ về năng lực của HS. Ngoài ra, các cơ sở giáo dục đại học cũng cần xây dựng tiêu chí xét tuyển rõ ràng, minh bạch, dựa trên năng lực thực tế của HS, đảm bảo công bằng và khách quan trong công tác tuyển sinh.

- *Đối với thí sinh:* thí sinh cần chủ động ôn tập và rèn luyện các năng lực cần thiết, không chỉ kiến thức mà còn cả kĩ năng tư duy, giải quyết vấn đề, làm việc nhóm, và năng lực ngôn ngữ. Ngoài ra, cần tìm hiểu kĩ về các kì thi đánh giá năng lực, bao gồm cả mục tiêu, nội dung, hình thức, và quy trình thi, để có sự chuẩn bị tốt nhất và cần tự tin và bình tĩnh khi thi, tránh căng thẳng và lo lắng quá mức, để có thể phát huy tối đa năng lực của bản thân.

- *Đối với cơ quan quản lí chất lượng giáo dục:* thứ nhất, việc thành lập một hội đồng kiểm định độc lập là cần thiết, một tổ chức trung gian có thể chịu trách nhiệm thẩm định độ khó và độ tin cậy của đề thi, đảm bảo tính công

bằng giữa các bài thi; thứ hai, áp dụng một phương pháp tính điểm chuẩn hóa cho tất cả các bài thi đánh giá năng lực để tăng tính so sánh và tin cậy của kết quả thi; thứ ba, cải thiện cơ sở hạ tầng thi trên máy.

3. Kết luận

Việc đánh giá năng lực HS tại các trường THPT ở Việt Nam đã được xem xét qua nhiều phương thức khác nhau, bao gồm các kì thi do Bộ GD-ĐT tổ chức, các kì thi đánh giá năng lực của các trường đại học, cùng các hình thức đánh giá trong lớp học. Các phương thức này đã được so sánh dựa trên nhiều tiêu chí như mục tiêu, nội dung, hình thức, quy trình và độ tin cậy của kết quả. Đồng thời, chúng tôi đã chỉ ra các điểm mạnh cần phát huy và những hạn chế cần khắc phục trong từng phương pháp. Để cải thiện công tác đánh giá năng lực HS và thúc đẩy đổi mới giáo dục tại Việt Nam, các giải pháp cụ thể đã được đề xuất. Tuy nhiên, phương pháp nghiên cứu chủ yếu dựa trên định tính, thiếu sự kết hợp với phương pháp định lượng, điều này hạn chế khả năng cung cấp cái nhìn toàn diện về vấn đề. Thêm vào đó, nguồn dữ liệu chủ yếu dựa vào tài liệu có sẵn, chưa khai thác đầy đủ dữ liệu thực tế từ các trường THPT và các đơn vị tổ chức thi. Trong tương lai, việc kết hợp phương pháp định tính và định lượng sẽ giúp thu thập và phân tích dữ liệu một cách đầy đủ hơn. Đồng thời, nghiên cứu sẽ mở rộng phạm vi và tăng cường thu thập dữ liệu thực tế từ các trường THPT, so sánh với các bài thi đánh giá năng lực quốc tế, từ đó cung cấp những đánh giá chính xác và khách quan hơn.

Tài liệu tham khảo

- Bộ GD-ĐT (2014). *Quyết định số 3538/QĐ-BGDĐT ngày 09/9/2014 phê duyệt phương án thi tốt nghiệp trung học phổ thông và tuyển sinh đại học, cao đẳng từ năm 2015*.
- Bộ GD-ĐT (2022). *Công văn số 1496/BGDĐT-GDTrH ngày 19/4/2022 về việc triển khai thực hiện chương trình giáo dục trung học năm học 2022-2023*.
- Bùi Đức Nhân (2023). Xu hướng đổi mới đánh giá kết quả học tập của học sinh hướng đến hình thành năng lực. *Tạp chí Khoa học, Trường Đại học Thủ đô Hà Nội*, 69, 20-27.
- Đại học Bách Khoa Hà Nội (2025). *Giới thiệu về kì thi đánh giá tư duy TSA*. <https://ts.hust.edu.vn/tin-tuc/gioi-thieu-ve-ky-thi-danh-gia-tu-duy-tsa>
- Đại học Quốc gia Thành phố Hồ Chí Minh (2024). *Cấu trúc, đề cương chi tiết bài thi Đánh giá năng lực (HSA) năm 2025 của Đại học Quốc gia Thành phố Hồ Chí Minh*. <https://vnuhcm.edu.vn/>
- Glaser, R., Chudowsky, N., & Pellegrino, J. W. (2001). *Knowing what students know: The science and design of educational assessment*. National Academies Press.
- Hoàng Hòa Bình (2015). Năng lực và đánh giá theo năng lực. *Tạp chí Khoa học, Trường Đại học Sư phạm Thành phố Hồ Chí Minh*, 6(71), 21-31.
- Kolen, M. J., & Brennan, R. L. (2014). *Test Equating, Scaling, and Linking: Methods and Practices*. Springer-Verlag, New York.
- Lê Thái Hưng, Nguyễn Thái Hà (2021). Xu thế kiểm tra, đánh giá năng lực người học trên nền tảng công nghệ. *Tạp chí Khoa học Giáo dục Việt Nam*, 42, 1-6.
- OECD (2002). *Definition and Selection of Competencies: Theoretical and Conceptual Foundation*. Neuchatel, Switzerland: Swiss Federal Statistical Office.
- Sansivieri, V., Wiberg, M., & Matteucci, M. (2017). A Review of Test Equating Methods with a Special Focus on IRT-Based Approaches. *Statistica*, 77(4), 329-352. <https://doi.org/10.6092/issn.1973-2201/7066>
- Škrinjarić, B. (2022). Competence-based approaches in organizational and individual context. *Humanities and Social Sciences Communications*, 9(1). <https://doi.org/10.1057/s41599-022-01047-1>
- Trung tâm Khảo thí, Đại học Quốc gia Hà Nội (2024). *Cấu trúc, đề cương chi tiết bài thi Đánh giá năng lực (HSA) năm 2025 của Đại học Quốc gia Hà Nội*. <https://vnu.edu.vn>