Đề xuất quy trình thiết kế bài kiểm tra tiếng Anh theo định hướng giao tiếp cho sinh viên không chuyên

Proposing a process for designing communicative English tests for non-English-major students

ThS. LÊ THỊ THANH TRANG
Học viện Hành chính và Quản trị công; Email: lethanhtrangnapa.vn@gmail.com

Ngày nhận bài: 19/08/2026 Ngày phản biện: 25/08/2026 Ngày duyệt đăng: 20/9/2026
Chủ đề của bài viết thuộc lĩnh vực: Khoa học xã hội và Nhân văn

Tóm tắt: *Bài viết đề xuất quy trình thiết kế bài kiểm tra tiếng Anh theo định hướng giao tiếp cho sinh viên không chuyên trong học phần Tiếng Anh giao tiếp và hội nhập. Phương pháp là phân tích các tài liệu chuyên môn, hai nghiên cứu thực nghiệm và đề cương học phần. Kết quả cho thấy nguồn tài liệu đã nêu rõ nguyên tắc và các giai đoạn xây dựng, nhưng vẫn cần cách làm cụ thể ở cấp học phần để nối chuẩn đầu ra với hành vi giao tiếp, minh chứng, nhiệm vụ, cách chấm và cách dùng kết quả. Bài viết đề xuất bảy bước; mỗi bước nêu sản phẩm và câu hỏi kiểm tra, kèm ví dụ đặc tả cuối học phần. Điểm mới là chuỗi liên kết có thể kiểm tra ngược và hai vòng phản hồi: trước khi tổ chức để phát hiện lỗi, sau khi tổ chức để sửa bài kiểm tra bằng dữ liệu thực tế. Đây là đề xuất lý luận; cần phản biện chuyên gia, thử nghiệm nhiệm vụ và phân tích kết quả trước khi dùng cho quyết định quan trọng.*
Từ khóa: *thiết kế bài kiểm tra; định hướng giao tiếp; sinh viên không chuyên; chuẩn đầu ra; đặc tả bài kiểm tra.*

Abstract: *This article proposes a process for designing communicative English tests for non-English-major students in the English for Communication and Integration course. The methodology involves analyzing professional literature, two empirical studies, and the course syllabus. The findings show that the available literature clearly identifies the principles and stages of test development, but a more specific approach is needed at the course level to connect learning outcomes with communicative behaviors, evidence, tasks, scoring methods, and the use of results. The article proposes a seven-step process, with each step specifying outputs and guiding questions, accompanied by an example of an end-of-course test specification. The novelty lies in a traceable linkage chain and two feedback loops: one before administration to identify potential problems, and another after administration to revise the test based on empirical data. This is a theoretical proposal so expert review, task piloting, and results analysis are needed before the process is used for high-stakes decision-making.*
Keywords: *test design; communicative orientation; non-English-major students; learning outcomes; test specification.*

1. Đặt vấn đề

Kiểm tra, đánh giá là một phần của quá trình dạy học tiếng Anh ở đại học. Kết quả bài kiểm tra cuối học phần được dùng để xác nhận mức độ hoàn thành chuẩn đầu ra, tính điểm và cung cấp thông tin cho sinh viên, giảng viên, tổ bộ môn. Vì vậy, một bài kiểm tra tốt không chỉ cần đủ số câu hoặc đúng hình thức. Bài kiểm tra phải tạo được minh chứng phù hợp - tức những gì sinh viên thực sự làm và thể hiện trong bài - để người sử dụng kết quả đưa ra nhận định có căn cứ về năng lực tiếng Anh của sinh viên.

Trong bài viết này, bài kiểm tra theo định hướng giao tiếp được hiểu là bài kiểm tra yêu cầu người học tiếp nhận, lựa chọn hoặc sử dụng tiếng Anh để đạt một mục đích trong bối cảnh xác định. Nhiệm vụ cần thể hiện người giao tiếp, mục đích, thông tin cần xử lý và sản phẩm cần tạo. Cách hiểu này không loại bỏ câu hỏi có đáp án xác định. Câu hỏi lựa chọn hoặc trả lời ngắn vẫn phù hợp nếu yêu cầu người học hiểu và xử lý nghĩa trong một thông báo, email, hội thoại hay văn bản gần với việc sử dụng ngôn ngữ. Điểm cốt lõi là mỗi câu hỏi hoặc nhiệm vụ phải tạo ra đúng loại minh chứng mà chuẩn đầu ra yêu cầu.

Các tài liệu chuyên môn đã cung cấp nền tảng khá đầy đủ. Brown và Abeywickrama (2019) trình bày các nguyên tắc về tính khả thi, độ tin cậy, độ giá trị, tính xác thực và tác

động đối với việc học. Hughes và Hughes (2020,) xem xây dựng bài kiểm tra là quá trình giải quyết vấn đề, bắt đầu từ mục đích rồi mới lựa chọn nội dung, kỹ thuật và cách chấm. Fulcher và Davidson (2007) nhấn mạnh cấu trúc năng lực cần đo và đặc tả. Council of Europe và Association of Language Testers in Europe (2026) trình bày chu trình rộng từ lập kế hoạch, xây dựng đặc tả, viết câu hỏi, tổ chức, chấm đến theo dõi và sửa đổi.

Hai nghiên cứu của Kang và Lam (2024) và Phuong và cộng sự (2025) không nghiên cứu trực tiếp việc xây dựng bài kiểm tra cuối học phần. Kang và Lam tìm hiểu kiến thức, quan niệm và cách thực hiện đánh giá thường xuyên của ba giảng viên tiếng Anh đại học tại Trung Quốc. Vì vậy, bài viết không dùng hai công trình này để nói rằng các quy trình thiết kế bài kiểm tra hiện có còn thiếu. Hai nghiên cứu giúp làm rõ một vấn đề khác: một nguyên tắc đúng chỉ tạo ra kết quả khi người thực hiện hiểu và áp dụng được trong điều kiện cụ thể. Kang và Lam cho thấy hiểu biết cơ bản chưa chắc dẫn đến nhiệm vụ lấy người học làm trung tâm, tiêu chí rõ và phản hồi theo quá trình. Phuong và cộng sự cho thấy giáo viên phải điều chỉnh cách đánh giá theo người học, nguồn lực và yêu cầu của cơ sở giáo dục. Mối liên hệ với bài kiểm tra cuối học phần nằm ở chỗ người biên soạn, người tổ chức và người chấm cũng phải chuyển quy định thành việc làm cụ thể. Do đó, quy trình cần nêu rõ sản phẩm phải hoàn thành, có phản biện đồng nghiệp, hướng dẫn chấm và bước kiểm tra điều kiện thực hiện.

Học phần Tiếng Anh giao tiếp và hội nhập tại Học viện Hành chính và Quản trị công được chọn làm bối cảnh minh họa. Đây là học phần bắt buộc thuộc khối kiến thức giáo dục đại cương, dành cho sinh viên đại học không chuyên Ngôn ngữ Anh. Học phần có 5 tín chỉ, gồm 45 tiết lý thuyết và 60 tiết thực hành; ngôn ngữ giảng dạy là tiếng Anh. Đề cương không nêu một mức đầu vào CEFR riêng và ghi rõ học phần không có học phần tiên quyết hoặc học trước (Học viện Hành chính và Quản trị công, 2025, tr. 1-2). Khi hoàn thành học phần, sinh viên được hướng đến bậc 3 của Khung năng lực ngoại ngữ 6 bậc dùng cho Việt Nam, được đề cương diễn giải ở khoảng A2+ đến B1- (tr. 4). Bài kiểm tra cuối học phần chiếm 60% điểm tổng kết; phần nghe, đọc và viết kéo dài 110 phút, còn phần vấn đáp kéo dài 10 phút (tr. 7-8).

Bài viết trả lời hai câu hỏi: (1) Cần dựa vào những nguyên tắc và yêu cầu nào khi thiết kế bài kiểm tra tiếng Anh theo định hướng giao tiếp cho sinh viên không chuyên? (2) Quy trình nào có thể chuyển chuẩn đầu ra thành nhiệm vụ, minh chứng, cách chấm và cơ chế cải tiến có căn cứ? Điểm mới không phải là gộp nhiều tài liệu thành một danh sách bước. Đề xuất tập trung vào chuỗi có thể truy vết từ chuẩn đầu ra đến hành vi giao tiếp, minh chứng, nhiệm vụ, cách chấm và cách sử dụng kết quả; quy định sản phẩm và điểm kiểm soát ở từng bước; đồng thời tạo vòng phản hồi trước và sau khi tổ chức. Đây là đề xuất khái niệm, chưa phải mô hình đã được thẩm định.

2. Phương pháp nghiên cứu

Bài viết sử dụng phương pháp phân tích tài liệu. Nguồn thứ nhất gồm ba sách chuyên môn và một sổ tay phát triển bài kiểm tra. Các phần được đọc tập trung vào nguyên tắc chất lượng, năng lực cần đo, đặc tả và các giai đoạn xây dựng bài kiểm tra: Brown và Abeywickrama (2019), Fulcher và Davidson (2007), Hughes và Hughes (2020), cùng Council of Europe và Association of Language Testers in Europe (2026). Nguồn thứ hai gồm Kang và Lam (2024) và Phuong và cộng sự (2025). Hai nghiên cứu này được dùng để xem xét vai trò của người thực hiện và điều kiện lớp học, không được coi là bằng chứng trực tiếp về thiết kế bài kiểm tra cuối học phần. Nguồn thứ ba là đề cương học phần Tiếng Anh giao tiếp và hội nhập, nhất là các trang 1-8, dùng để xác định người học, thời lượng, mục tiêu, chuẩn đầu ra, tỷ trọng và hình thức đánh giá.

Tác giả đọc và đối chiếu các nội dung liên quan đến mục đích, năng lực cần đo, minh chứng, đặc tả, nhiệm vụ, cách chấm, tổ chức và cải tiến. Sau đó, các quyết định được sắp xếp theo trình tự công việc ở cấp học phần; mỗi bước được gắn với sản phẩm và câu hỏi kiểm soát. Bài viết không sử dụng điểm kiểm tra, khảo sát, phỏng vấn hoặc quan sát lớp học. Phạm vi của bài viết là đề xuất khái niệm, không phải kiểm định hiệu quả hay xác nhận độ giá trị của một công cụ.

3. Kết quả nghiên cứu và thảo luận

3.1. Căn cứ đề xuất quy trình

3.1.1. Năng lực giao tiếp và yêu cầu chất lượng

Độ giá trị, hiểu đơn giản là mức độ điểm số phản ánh đúng điều cần đo, phụ thuộc vào mối nối giữa chuẩn đầu ra, minh chứng và cách hiểu điểm số. Nếu chuẩn yêu cầu trao đổi để giải quyết tình huống nhưng phần nói chỉ yêu cầu học thuộc một bài trình bày, minh chứng chưa phù hợp. Nếu bài đọc có nhiều kiến thức chuyên môn ngoài học phần, điểm số có thể chịu ảnh hưởng của kiến thức nền. Do đó, tên dạng bài không tự bảo đảm độ giá trị; người thiết kế phải giải thích được nhiệm vụ tạo ra minh chứng gì (Fulcher & Davidson, 2007; Hughes & Hughes, 2020).

Độ tin cậy là mức kết quả giữ được sự nhất quán trong những điều kiện tương đương; công bằng là tạo cơ hội hợp lý để sinh viên thể hiện năng lực. Sai lệch có thể đến từ chỉ dẫn mơ hồ, quá ít nhiệm vụ, chất lượng âm thanh khác nhau,

đáp án thiếu phương án tương đương, người chấm hiểu tiêu chí khác nhau hoặc điều kiện giữa các lớp không thống nhất. Có thể giảm sai lệch bằng đặc tả rõ, phản biện, làm thử, hướng dẫn chấm và lưu biên bản sự cố. Tính xác thực cũng phải cân bằng với tính khả thi: nhiệm vụ gần thực tế nhưng quá dài hoặc khó chấm sẽ không phù hợp với lớp đông (Brown & Abeywickrama, 2019).

3.1.2. Yêu cầu của học phần

Đề cương xác định tám chuẩn đầu ra học phần, được viết tắt là CLO, gồm ngữ pháp; từ vựng và ngữ âm; nghe; nói; đọc; viết; trình bày và đánh giá quan điểm; tự chủ và trách nhiệm. Bài kiểm tra cuối học phần đánh giá CLO1 đến CLO7; CLO8 chủ yếu được thể hiện trong quá trình học. Tỷ trọng là 7,5% cho CLO1, 7,5% cho CLO2, 20% cho từng CLO3, CLO4, CLO5 và CLO6, cùng 5% cho CLO7 (Học viện Hành chính và Quản trị công, 2025). Cấu trúc này cho thấy ngữ pháp, từ vựng và ngữ âm cần được kiểm tra trong ngữ cảnh hoặc tích hợp vào tiêu chí của nhiệm vụ, nhưng vị trí tính điểm phải rõ để tránh tính trùng.

Thời lượng chính thức gồm nghe 20 phút, đọc 45 phút, viết 45 phút và vấn đáp 10 phút (Học viện Hành chính và Quản trị công, 2025). Đây là ràng buộc của quy trình. Phần nói không thể có quá nhiều nhiệm vụ; nghe và đọc cần đủ mẫu để giảm ảnh hưởng của một văn bản; viết phải phù hợp 45 phút; việc chấm nhiều lớp đòi hỏi đáp án và tiêu chí rõ. Vì vậy, định hướng giao tiếp cần phối hợp câu hỏi có đáp án xác định với nhiệm vụ tạo sản phẩm, không nên đồng nhất giao tiếp với một loại câu hỏi duy nhất.

3.1.3. Khoảng trống và phần bổ sung của đề xuất

Từ Brown và Abeywickrama, đề xuất kế thừa năm yêu cầu chất lượng và biến chúng thành câu hỏi kiểm tra ở từng bước. Từ Fulcher và Davidson, đề xuất kế thừa cách xác định năng lực và đặc tả, sau đó nối thành một chuỗi: chuẩn đầu ra - hành vi giao tiếp - minh chứng - nhiệm vụ - cách chấm - cách dùng kết quả. Từ Hughes và Hughes, đề xuất kế thừa cách bắt đầu từ mục đích và bổ sung các hồ sơ cần chuyển giao giữa người biên soạn, người phản biện, người tổ chức và người chấm. Từ sổ tay năm 2026, đề xuất rút gọn chu trình rộng thành bảy bước có thể thực hiện ở tổ bộ môn, nhưng vẫn giữ việc kiểm tra trước khi tổ chức và xem lại sau khi có kết quả. Hai nghiên cứu của Kang và Lam (2024) và Phuong và cộng sự (2025) không cung cấp các bước thiết kế. Chúng được dùng để nhắc rằng quy trình chỉ có thể vận hành khi người thực hiện hiểu nhiệm vụ và khi yêu cầu phù hợp với điều kiện thật.

So với một đặc tả chỉ ghi phần kiểm tra, số câu và số điểm, đề xuất có ba phần bổ sung. *Thứ nhất,* mỗi chuẩn đầu ra được nối với hành vi giao tiếp và loại minh chứng cần nhìn thấy; nhờ đó, người đọc có thể kiểm tra ngược vì sao một nhiệm vụ được đưa vào. *Thứ hai,* mỗi bước không chỉ nêu việc cần làm mà còn nêu kết quả phải hoàn thành và câu hỏi kiểm tra chất lượng. Cách này giúp phát hiện trường hợp tài liệu đã viết đúng nhưng người tổ chức hoặc người chấm chưa hiểu thống nhất. *Thứ ba,* sau khi kiểm tra, thông tin về câu hỏi, thời gian, điểm chấm và sự cố được đối chiếu với đặc tả để xác định phần cần sửa. Ba phần này là sự cụ thể hóa cho cấp học phần, không phải một lý thuyết mới thay thế các quy trình đã có.

3.2. Quy trình bảy bước

Quy trình được sắp xếp theo thứ tự công việc, nhưng nhóm thiết kế có thể quay lại bước trước khi phát hiện vấn đề. Mỗi bước phải tạo ra một kết quả cụ thể và được kiểm tra trước khi chuyển sang bước tiếp theo.

Bước 1, nhóm thiết kế xác định mục đích, người học, thời lượng, nguồn lực và quyết định sẽ được đưa ra từ điểm số. Kết quả cần có là bản mô tả ngắn, trong đó nói rõ bài kiểm tra được dùng và không được dùng cho việc gì.

Bước 2, từng chuẩn đầu ra được chuyển thành hành vi có thể quan sát và loại minh chứng cần thu. Kết quả là một sơ đồ đối chiếu để bảo đảm mỗi chuẩn đều có minh chứng và không có nhiệm vụ nằm ngoài mục tiêu.

Bước 3, nhóm thiết kế viết đặc tả, tức bản mô tả chi tiết về các phần kiểm tra, văn bản, chủ đề, nhiệm vụ, thời gian, tỷ trọng điểm và cách chấm. Bản đặc tả phải đúng đề cương, bao phủ đủ nội dung và có thể tổ chức trong điều kiện thật.

Bước 4, câu hỏi, chỉ dẫn, đáp án và tiêu chí chấm được viết đồng thời để kiểm tra ngay xem nhiệm vụ có tạo ra đúng minh chứng hay không.

Bước 5, một người khác làm thử và phản biện; nhiệm vụ mới nên được thử với người học gần với nhóm mục tiêu khi điều kiện cho phép. Kết quả của bước này là phiếu nhận xét và bản đã sửa, không chỉ là sự đồng ý bằng lời.

Bước 6, nhóm phụ trách thống nhất cách tổ chức, chấm và báo cáo điểm giữa các lớp. Cần có hướng dẫn cho cán bộ tổ chức, người chấm và cách ghi sự cố để giảm khác biệt không liên quan đến năng lực của sinh viên.

Bước 7, nhóm thiết kế xem lại chất lượng câu hỏi, thời gian làm bài, phân bố điểm, mức chênh giữa người chấm và phản hồi của người tham gia. Mọi thay đổi phải nêu lý do và được lưu theo phiên bản để lần sau có thể kiểm tra lại.

3.2.1. Xác định mục đích, đối tượng và ràng buộc

Nhóm thiết kế xác định bài kiểm tra dùng để xác nhận hoàn thành học phần, tính điểm và cung cấp thông tin ở mức nào. Bản mô tả cần nêu nhóm sinh viên, trình độ mục tiêu, số lớp, thời lượng, thiết bị, nhân lực và thời gian chấm. Cần giới hạn cách diễn giải: điểm bài kiểm tra cho biết mức thực

hiện các chuẩn được lấy mẫu trong điều kiện cụ thể, không đại diện cho toàn bộ năng lực tiếng Anh trong mọi tình huống. Với học phần này, thời lượng và tỷ trọng đã được phê duyệt; mức đầu vào lại không được quy định, nên nhiệm vụ phải hướng đến đầu ra nhưng chỉ dẫn cần rõ và nội dung phải nằm trong phạm vi đã học.

3.2.2. Chuyển chuẩn đầu ra thành hành vi và minh chứng

CLO3 về nghe có thể gồm hiểu ý chính, chi tiết và hàm ý trong hội thoại hoặc bài nói. CLO4 về nói có thể gồm trình bày ý kiến, hỏi để làm rõ, phản hồi và xử lý tình huống. CLO6 về viết cần chỉ rõ loại văn bản, người nhận, mục đích và nội dung bắt buộc. Sau hành vi là minh chứng. Khả năng tương tác không thể chỉ lấy từ độc thoại; khả năng viết email không thể chỉ lấy từ câu chọn đáp án ngữ pháp. Với CLO1 và CLO2, nhóm thiết kế phải chỉ rõ phần điểm lấy từ câu hỏi trong ngữ cảnh và phần điểm lấy từ sản phẩm nói, viết.

3.2.3. Xây dựng đặc tả bài kiểm tra

Đặc tả là bản mô tả chi tiết để những người khác nhau có thể cùng hiểu bài kiểm tra sẽ đo gì và được làm như thế nào. Đặc tả cần quy định loại và độ dài văn bản, chủ đề, đặc điểm ngôn ngữ, số lần nghe, số mẫu nhiệm vụ, dạng câu hỏi, thời lượng, tỷ trọng và quy tắc chấm. Phần nói và viết cần nêu tình huống, lượng thông tin, mức hỗ trợ, thời gian chuẩn bị và tiêu chí. Phần đọc và nghe phải kiểm soát kiến thức nền, tốc độ và chất lượng âm thanh. Đặc tả phải đủ rõ để các nhóm giảng viên tạo được những phiên bản có cấu trúc tương đương, nhưng không quá cứng đến mức chỉ dùng được một chủ đề.

3.2.4. Biên soạn đồng thời nhiệm vụ và tài liệu chấm

Câu hỏi, chỉ dẫn, đáp án và hướng dẫn chấm được viết cùng nhau. Nếu người biên soạn không nêu được đáp án hoặc minh chứng mong đợi, nhiệm vụ có thể chưa đủ rõ. Với viết và nói, tiêu chí phải bám chuẩn đầu ra và tránh trừ một lỗi ở nhiều tiêu chí khi không có quy định. Văn bản đầu vào được rà soát về nguồn, độ dài, kiến thức nền và quyền sử dụng. Bối cảnh giao tiếp chỉ có ý nghĩa khi làm rõ mục đích; không cần thêm một câu chuyện dài nhưng không ảnh hưởng đến việc người học phải làm.

3.2.5. Phản biện và thử nghiệm

Người phản biện đối chiếu từng câu hỏi với đặc tả, tự làm bài, kiểm tra đáp án và ghi thời gian. Phiếu phản biện cần hỏi: nhiệm vụ đo hành vi nào; minh chứng có đủ không; chỉ dẫn có rõ không; có kiến thức ngoài mục tiêu không; điểm và thời gian có tương xứng không. Khi có thể, nhiệm vụ mới được thử với nhóm gần đối tượng mục tiêu và ghi lại thời gian, câu khó hiểu, đáp án bất ngờ, khó khăn khi chấm. Đây là bước kiểm soát chất lượng, không phải thủ tục sau khi bài kiểm tra đã hoàn tất (Council of Europe & Association of Language Testers in Europe, 2026).

3.2.6. Chuẩn hóa tổ chức, chấm và báo cáo

Hướng dẫn tổ chức quy định thiết bị, âm lượng, số lần nghe, thời gian, cách xử lý gián đoạn, hỗ trợ hợp lý và bảo mật. Phần vấn đáp cần thống nhất cách giới thiệu nhiệm vụ, thời gian chuẩn bị, câu hỏi gợi mở và cách ghép cặp. Người chấm làm quen với đáp án, tiêu chí và bài mẫu trước khi chấm. Với nói và viết, tổ bộ môn nên chấm chung một số bài để thống nhất cách hiểu mức điểm, sau đó kiểm tra chéo một tỷ lệ phù hợp. Báo cáo kết quả cần qua bước kiểm tra nhập điểm; nếu điều kiện cho phép, sinh viên nên nhận thông tin theo phần hoặc tiêu chí thay vì chỉ một điểm tổng.

3.2.7. Phân tích kết quả và sửa đổi

Sau khi sử dụng, nhóm thiết kế xem xét câu hỏi quá dễ, quá khó, có đáp án ngoài dự kiến; phân bố điểm từng phần; thời gian; chênh lệch giữa người chấm; sự cố và phản hồi. Không phải lớp học nào cũng có điều kiện phân tích thống kê phức tạp, nhưng thống kê mô tả, đối chiếu đáp án và biên bản tổ chức vẫn hữu ích. Việc sửa phải truy ngược đúng mắt xích. Báo cáo sau kiểm tra ghi thay đổi, lý do, người duyệt và phiên bản áp dụng.

3.3. Minh họa đặc tả theo chuẩn đầu ra học phần

Với CLO1 về ngữ pháp, tỷ trọng là 7,5%. Sinh viên cần dùng được các cấu trúc cốt lõi để hiểu hoặc tạo một thông điệp rõ. Minh chứng có thể đến từ câu hỏi đặt trong ngữ cảnh và từ tiêu chí sử dụng ngôn ngữ ở phần viết hoặc nói. Đặc tả phải chỉ rõ điểm ngữ pháp được tính ở đâu.

Với CLO2 về từ vựng và ngữ âm, tỷ trọng là 7,5%. Sinh viên cần dùng từ phù hợp với chủ đề; khi nói, các em cần dùng trọng âm, nối âm và ngữ điệu ở mức giúp người nghe hiểu được ý. Có thể thu minh chứng qua câu hỏi từ vựng trong ngữ cảnh, tiêu chí ngôn ngữ ở bài viết và tiêu chí về mức dễ hiểu ở bài nói. Nơi tính điểm cũng phải được ghi rõ.

Với CLO3 về nghe, tỷ trọng là 20%. Bài kiểm tra cần cho thấy sinh viên có thể hiểu ý chính, chi tiết và một số hàm ý trong hội thoại, phỏng vấn hoặc bài nói. Nên dùng ít nhất hai văn bản để kết quả không phụ thuộc quá nhiều vào một chủ đề. Trong 20 phút, các câu hỏi phải yêu cầu xử lý nghĩa; số lần nghe và chất lượng âm thanh phải giống nhau giữa các lớp.

Với CLO4 về nói, tỷ trọng là 20%. Sinh viên cần trình bày ý kiến, hỏi để làm rõ, phản hồi, diễn đạt lại và xử lý một tình huống quen thuộc. Trong khoảng 10 phút, có thể kết hợp một phần nói cá nhân ngắn với một phần tương tác. Người hỏi và người chấm cần dùng cùng hướng dẫn và cùng tiêu chí.

Với CLO5 về đọc, tỷ trọng là 20%. Sinh viên cần tìm thông tin, hiểu quan điểm và mục đích của người viết, đồng thời nhận ra quan hệ giữa các ý. Nên dùng ít nhất hai văn

bản và các câu hỏi tập trung vào việc hiểu nghĩa. Nội dung phải vừa sức, hạn chế kiến thức chuyên môn ngoài học phần và phù hợp thời lượng 45 phút.

Với CLO6 về viết, tỷ trọng là 20%. Sinh viên cần tạo một văn bản có mục đích rõ, hướng đến một người đọc cụ thể, có bố cục và ngôn ngữ phù hợp. Nhiệm vụ nên cung cấp đủ dữ kiện, nói rõ người viết, người nhận và thông tin bắt buộc. Thời gian là 45 phút; hướng dẫn chấm cần mô tả riêng các mặt như hoàn thành nhiệm vụ, tổ chức ý và sử dụng ngôn ngữ.

Với CLO7 về trình bày và đánh giá quan điểm, tỷ trọng là 5%. Sinh viên cần nêu được lý do và nhận ra hoặc đánh giá một quan điểm khác biệt một cách hợp lý. Minh chứng có thể đến từ câu suy luận ở phần đọc hoặc yêu cầu giải thích ở phần viết và nói. Đặc tả phải ghi rõ phần nào tạo ra điểm CLO7 để tránh cộng điểm lặp lại.

3.4. Thảo luận

3.4.1. Mối liên hệ với năng lực đánh giá của giảng viên

Kang và Lam (2024, tr. 1, 16) nghiên cứu đánh giá thường xuyên, nhưng bốn việc họ xem xét - thiết kế, sử dụng, hiểu kết quả và truyền đạt kết quả - cũng có mặt trong một bài kiểm tra cuối học phần. Nếu người biên soạn hiểu nguyên tắc nhưng không viết được nhiệm vụ phù hợp, hoặc nếu người chấm hiểu tiêu chí khác nhau, chất lượng thiết kế sẽ không đi vào thực tế. Vì vậy, nghiên cứu này hỗ trợ việc bổ sung bước phản biện, làm thử và thống nhất cách chấm; nó không phải nguồn tạo ra cấu trúc bảy bước.

Phuong và cộng sự (2025, tr. 6, 13-18) cho thấy giáo viên Việt Nam điều chỉnh hoạt động đánh giá theo người học, nguồn lực và yêu cầu của cơ sở giáo dục. Một mặt, bài kiểm tra có thể làm cho giảng viên và sinh viên tập trung nhiều hơn vào những nội dung được tính điểm. Mặt khác, thông tin từ lớp học giúp nhóm thiết kế nhận ra một nhiệm vụ quá xa lạ, quá dài hoặc khó tổ chức. Do đó, quy trình cần giữ chuẩn đầu ra ổn định nhưng cho phép điều chỉnh cách thực hiện dựa trên dữ liệu. Hai nghiên cứu chỉ giúp giải thích điều kiện triển khai; chúng không chứng minh rằng quy trình đề xuất sẽ làm tăng kết quả học tập.

3.4.2 Sự đánh đổi và hướng thẩm định

Thiết kế bài kiểm tra luôn có sự đánh đổi. Nhiều nhiệm vụ hơn giúp bao phủ nội dung tốt hơn nhưng làm bài dài hơn. Câu trả lời mở cho nhiều thông tin hơn nhưng tốn thời gian chấm và dễ có chênh lệch giữa người chấm. Vì vậy, quy trình không đưa ra một công thức duy nhất mà yêu cầu nhóm thiết kế ghi rõ ưu tiên, giới hạn và lý do lựa chọn. Với học phần này, một phương án có thể xem xét là lấy mẫu cả bốn kỹ năng; dùng một số câu hỏi có đáp án xác định cho nghe và đọc; dùng nhiệm vụ có mục đích rõ cho viết và nói; đồng thời ghi rõ điểm ngữ pháp, từ vựng và phát âm được tính ở đâu.

Giới hạn lớn nhất của bài viết là mới phân tích một số tài liệu do một tác giả thực hiện; chưa có bài kiểm tra hoàn chỉnh và chưa có dữ liệu từ sinh viên. Bước tiếp theo nên gồm: mời khoảng 5-7 chuyên gia đối chiếu quy trình với đề cương; để giảng viên làm thử và thử nhiệm vụ mới với nhóm sinh viên gần đối tượng mục tiêu; xem câu nào quá dễ, quá khó hoặc không phân biệt được mức làm bài; kiểm tra thời gian, phân bố điểm và câu trả lời ngoài dự kiến; so sánh mức thống nhất giữa người chấm; thu phản hồi của sinh viên và giảng viên. Chỉ sau những bước này mới có đủ căn cứ để dùng bài kiểm tra cho các quyết định quan trọng.

4. Kết luận

Đóng góp của đề xuất là chuỗi có thể truy vết từ chuẩn đầu ra đến hành vi, minh chứng, nhiệm vụ, cách chấm và cách sử dụng kết quả; mỗi bước có sản phẩm và điểm kiểm soát; hai vòng phản hồi giúp kiểm tra mức sẵn sàng trước khi tổ chức và dùng bằng chứng sau kiểm tra để sửa đổi. Đây là sự cụ thể hóa các quy trình chuyên môn cho cấp học phần, không phải quy trình thay thế hoặc mô hình đã được xác nhận. Nghiên cứu tiếp theo cần hoàn thiện nhiệm vụ, thẩm định chuyên gia, thử nghiệm với sinh viên, kiểm tra sự thống nhất giữa người chấm và phân tích dữ liệu sau kiểm tra./.

TÀI LIỆU THAM KHẢO

1. Brown, H. D., & Abeywickrama, P. (2019). *Language assessment: Principles and classroom practices* (3rd ed.). Pearson Education, tr.27-49, 57-88.

2. Council of Europe, & Association of Language Testers in Europe. (2026). *Manual for language test development and examining: For use with the CEFR*. Council of Europe. https://rm.coe.int/new-revised-manual-for-language-test-development-and-examining-joint-p/48802cc30d, tr.9-10, 31-45, 50-64, 67-97., 9-10, 14-29, 31-97.3. Fulcher, G., & Davidson, F. (2007). *Language testing and assessment: An advanced resource book*. Routledge, tr.36-63.

4. Học viện Hành chính và Quản trị công. (2025). *Đề cương chi tiết học phần Tiếng Anh giao tiếp và hội nhập (Mã học phần ISF0013)* (Tài liệu nội bộ), tr. 1-2.

5. Hughes, A., & Hughes, J. (2020). *Testing for language teachers* (3rd ed.). Cambridge University Press. https://doi.org/10.1017/9781009024723, tr.8-10, 63-77.

6. Kang, M., & Lam, R. (2024). Understanding university English instructors' assessment literacy: A formative assessment perspective. *Language Testing in Asia, 14*, Article 52. https://doi.org/10.1186/s40468-024-00323-y, tr.1, 16.

7. Phuong, H. Y., Tran, N. B. C., Nguyen, T. T. L., Lam, T. D., Bui, N. Q., & Le, T. T. (2025). From tests to tasks: How Vietnamese EFL teachers navigate washback through formative assessment practices. *Language Testing in Asia, 15*, Article 45. https://doi.org/10.1186/s40468-025-00392-7.