

# Thực trạng sử dụng các hoạt động kiểm tra đánh giá trong dạy học của giáo viên tiếng Anh

Nguyễn Việt Hùng\*

\*TS. Khoa Ngôn ngữ ứng dụng, Trường Quốc tế - ĐHQG Hà Nội  
Received: 15/04/2024; Accepted: 26/04/2024; Published: 10/5/2024

**Abstract:** Assessment in teaching in general and English teaching in particular is a major factor affecting teaching results. Therefore, the National Foreign Language Project 2020 (National Foreign Language Program 2020) has considered improving the foreign language testing and assessment competence for teachers as one of the pivotal tasks of the Project in the expansion duration from 2017- 2025. By means of a cross-sectional survey carried out in the summer of 2019, systematic random sampling with a sample size of 3103 Vietnamese English teachers at all 3 levels (primary school, middle school, secondary school) by regions (6 key areas in the country, including (1) Northern Mountain Region, (2) Northern Delta Region, (3) Central Region, (4) Central Highlands Region, (5) Southeast Region, and (6) Southwest Region), using quantitative statistical analysis tools, this article is to find out the perception of English teachers about knowledge of testing and assessment, and the actual implementation of assessment in daily teaching.

**Keywords:** Awareness, testing and assessment, teachers, primary school, middle school, secondary school

## 1. Mở đầu

Về mặt lý luận giáo dục ngoại ngữ, việc kiểm tra đánh giá đóng một vai trò quan trọng, và có một tác động rất lớn đối với quá trình dạy học, vì nó cung cấp thông tin để cải thiện việc dạy học liên tục (Lambert & Lines, 2000; Pierce, 2002; Brown, 2004). Việc nhận thức đúng về vai trò của KT-ĐG và hiểu rõ các sử dụng các hoạt động KT-ĐG khác nhau trong lớp học đối với GV TA là thực sự cần thiết. Vì nó không chỉ giúp cho GV đánh giá đúng người học mà còn có tác dụng xây dựng động cơ học tập cho người học (Martha & Barbara, 2014; Hoque, 2016). Từ những phát hiện đó, các nhà quản lý giáo dục có thể biết được GV TA cần các khóa tập huấn nào để nâng cao chất lượng dạy học, từ đó có kế hoạch cụ thể xác thực cho việc điều chỉnh chính sách về sử dụng nhân lực cũng như bồi dưỡng nguồn nhân lực cho phù hợp mà còn giúp cho các bên liên quan, đặc biệt là giáo viên có nhận thức đúng về những điểm mạnh, yếu của bản thân về năng lực khảo thí để chủ động tự bồi dưỡng và thay đổi trong cách dạy học của mình (Martha & Barbara, 2014; Kumaravadevelu, 2006; Johnson, 2011). Từ ý nghĩa đó nghiên cứu này nhằm mục đích trả lời 3 câu hỏi sau: (1) Thực trạng việc sử dụng các loại hình kiểm tra đánh giá của giáo viên TA các cấp như thế nào? (2) Giáo viên TA các cấp kiểm tra những nội dung gì trên lớp và theo hình thức nào? (3) Giáo viên TA các cấp có kinh nghiệm tham gia những hoạt động kiểm tra đánh giá nào?

## 2. Nội dung nghiên cứu

### 2.1. Một số vấn đề lý luận về kiểm tra đánh giá trong

### dạy học ngoại ngữ

Nghiên cứu về nhận thức của GV về các yếu tố tác động tới quá trình dạy học từ lâu đã được quan tâm rộng rãi trên thế giới (Borg, 2006, 2009). Năng lực và kiến thức về kiểm tra đánh giá của GV tiếng Anh ngày càng được quan tâm, và được xác định là yếu tố quan trọng ảnh hưởng tới kết quả dạy học của GV; vì lý do đó mà Đề án NNQG2020 đã xác định việc bồi dưỡng kiến thức kiểm tra đánh giá là một nhiệm vụ quan trọng trong giai đoạn điều chỉnh ĐA NNQG2020 giai đoạn 2017-2025 theo Quyết định 2080/QĐ-TTg của Thủ tướng Chính phủ (Chính phủ, 2008, 2017). Vì vậy nhận thức của GV tiếng Anh về kiểm tra đánh giá là một nhân tố khám phá nên được quan tâm và phân tích, để từ đó tìm ra GV TA đang thực sự nhìn nhận vai trò của KTĐG TA là gì, thực trạng GV đang áp dụng thực tế trên lớp như thế nào.

### 2.2. Phương pháp nghiên cứu.

Nghiên cứu được thực hiện dựa trên phương pháp nghiên cứu khảo sát thực tiễn, sử dụng bảng câu hỏi khảo sát trực tuyến (tại đường link: <http://ulis.vnu.edu.vn/phieu-hoi-giao-vien-ngoai-ngu-pho-thong-ve-hoat-dong-day>), đảm bảo các quy trình chuẩn hóa từ việc lựa chọn bảng hỏi, chuẩn hóa bảng hỏi, chọn mẫu, thu thập số liệu, và phân tích số liệu. Bảng câu hỏi có 12 câu hỏi, được thiết kế thành nhiều câu hỏi nhỏ (item) theo thang Likert 5 bậc nhằm phân tích sâu về nhận thức của GV tiếng Anh về kiến thức kiểm tra đánh giá, và thực tế thực hiện việc kiểm tra đánh giá trong việc dạy học hàng ngày. Bảng câu hỏi được điều tra thử nghiệm, kết quả phân tích độ tin cậy và độ

hiệu lực đều rất tốt. Về độ tin cậy Cronbach's Alpha ( $\alpha$ ), tất cả các câu hỏi gồm nhiều item nhỏ theo thang 5 bậc của bảng hỏi đều đạt giá trị lớn hơn 0.6. Về độ hiệu lực, hầu hết các câu hỏi gồm nhiều item nhỏ theo thang 5 bậc của bảng hỏi đều đồng cấu trúc vì phép kiểm định KMO and Bartlett's Test đều cho kết quả  $\text{sig} < 0.05$  và mỗi miền đều triết xuất được 1 factor, không có miền đo hay bảng hỏi nào triết xuất ra 2 factor trở lên.

Số liệu điều tra được xử lý bằng phần mềm SPSS for Window phiên bản 16.0. Các thông số được sử dụng trong nghiên cứu này bao gồm (1) các chỉ số thống kê mô tả như tần số tuyệt đối (các số đếm) và tần số tương đối (số phần trăm) đối với các biến dạng số và biến dạng chuỗi với các nhóm hạng không thứ bậc, điểm trung bình cộng, điểm trung vị (median), độ lệch chuẩn (SD); và (2) các chỉ số thống kê suy luận như so sánh giá trị trung bình (compare means) giữa 2 nhóm (T-Test), so sánh giá trị trung bình với 3 nhóm trở nên, phép phân tích phương sai một yếu tố (ANOVA). Đối với các câu có nhiều item nhỏ theo thang 5 bậc cần kiểm định (như Q2, Q7, Q10..), nếu phân phối giá trị tần suất và % không chuẩn hoặc kết quả kiểm định đồng nhất phương sai trong ANOVA không cho giá trị  $\text{sig} > 0.05$  thì sử dụng phép kiểm định Kruskal Wallis Test. Đối với các phép so sánh giá trị trung bình của 2 nhóm, phép kiểm định t về độc lập giữa hai mẫu (Independent Samples T Test) cho biết đối với một nhóm đơn thì trung bình của một nhóm chủ thể này có khác với trung bình của các nhóm chủ thể khác không. Các giá trị trung bình được coi là có ý nghĩa về mặt thống kê khi t-test của phân tích biến thiên có giá trị vượt ngưỡng thống kê với xác suất  $\text{sig} < 0.05$ . Đối với các biến cần kiểm định (như Q2, Q7, Q10), nếu phân phối giá trị tần suất và % không chuẩn thì sử dụng phép kiểm định Mann Whitney Test.

### 2.3. Quan niệm của giáo viên tiếng Anh về vai trò của kiểm tra đánh giá.

Ý kiến đánh giá của GV

Các hoạt động kiểm tra đánh giá theo thực tế của GV các cấp

Nội dung đánh giá về các hoạt động KT – ĐG thực tế của GV các cấp được thể hiện qua kết quả các câu hỏi 2, 3, 4, 5, 6 trong bảng câu hỏi dành cho GV. Các kết quả phân tích định lượng được kết hợp với kết quả định tính từ các cuộc phỏng vấn nhóm. Các thông tin bàn luận cụ thể được trình bày thành các phân nhỏ sau đây:

1) *Việc sử dụng các loại hình kiểm tra đánh giá của giáo viên các cấp*

Ý kiến đánh giá của GV về các loại hình kiểm tra đánh giá (Câu hỏi 2) được thể hiện qua bảng 2.1.

Bảng 2.1: Tần suất tự đánh giá về các loại hình KT – ĐG

	Các lựa chọn	2.2	2.3	2.4
Giá trị	Không sử dụng	48	18	51
	Ít khi sử dụng	239	89	173
	Thỉnh thoảng sử dụng	786	564	752
	Ưu tiên sử dụng	1672	1986	1730
	Rất ưu tiên sử dụng	358	446	397
	Tổng cộng	3103	3103	3103

Theo Bảng 2.1, một nét nổi bật dễ nhận thấy là nhóm ý kiến “ưu tiên sử dụng” luôn chiếm tỉ lệ cao nhất, và đều chiếm tỉ lệ trên 50% so với các nhóm ý kiến khác đánh giá về từng item trong câu hỏi 2. Nhóm ý kiến “thỉnh thoảng/có sử dụng” là nhóm có tỉ lệ cao thứ hai. Tiếp đó là các nhóm ý kiến “rất ưu tiên sử dụng”. Rõ ràng một thực tế là GV thường xuyên sử dụng các loại hình KT-ĐG được đưa ra. Theo thống kê có tới hơn 95% giáo viên tham gia khảo có cùng quan điểm như trên.

Để phân tích sâu hơn về đánh giá của GV về các loại hình kiểm tra đánh giá, các số liệu thống kê mô tả đối với từng câu hỏi nhỏ (item) được sử dụng gồm giá trị trung bình ( $\bar{X}$ ), độ lệch chuẩn (SD), trung tuyến (Median), mode, độ tin cậy ( $\alpha$ ) và tương quan (r) giữa các item với câu hỏi chính. Các số liệu mô tả về câu hỏi 2 (Q2) được thể hiện qua Bảng 4 dưới đây:

2) *Nội dung kiểm tra đánh giá trên lớp của giáo viên các cấp*

Một nét nổi bật dễ nhận thấy là nhóm ý kiến “ưu tiên đánh giá” luôn chiếm tỉ lệ trên 50% so với các nhóm ý kiến khác đánh giá về từng item trong câu hỏi 3. Các nhóm ý kiến “thỉnh thoảng/có đánh giá” là nhóm có tỉ lệ cao thứ hai. Tiếp đó là các nhóm ý kiến “rất ưu tiên đánh giá” là nhóm có tỉ lệ cao thứ ba. Rõ ràng Bảng 5 đã cho thấy một thực tế là GV thường xuyên KT – ĐG các nội dung nghe, nói, đọc, viết, ngữ âm, từ vựng, ngữ pháp và văn hóa như bảng hỏi đã đưa ra. Theo thống kê có tới hơn 90% giáo viên tham gia khảo có cùng quan điểm như trên. Có thể tính ra con số trung bình chung về tỉ lệ các mức độ đồng ý đối với các nội dung về vai trò của KT – ĐG là rất cao 92,19% (bao gồm 65,67% đồng ý, và 26,42% hoàn toàn đồng ý), chỉ có 7,89% phản. Để phân tích sâu hơn về đánh giá của GV về các loại hình kiểm tra đánh giá, các số liệu thống kê mô tả đối với từng câu hỏi nhỏ (item) được sử dụng gồm giá trị trung bình ( $\bar{X}$ ), độ lệch chuẩn (SD), trung tuyến (Median), mode, độ tin cậy ( $\alpha$ ) và tương quan (r) giữa các item với câu hỏi chính. Các số liệu mô tả về câu hỏi 3 (Q3) được thể hiện qua Bảng 7 dưới đây:

3) *Các hình thức kiểm tra đánh giá trên lớp của giáo viên các cấp*

Ý kiến đánh giá của GV về các hình thức kiểm tra

đánh giá (Câu hỏi 4) được thể hiện qua Bảng 2.2.

Bảng 2.2: Tần suất tự đánh giá về các hình thức KT-ĐG

	Các lựa chọn	4.1	4.2	4.3	4.4
Giá trị	Không bao giờ	247	13	173	712
	Ít khi	634	78	503	1040
	Thỉnh thoảng	1044	439	877	921
	Thường xuyên	946	1656	1227	361
	Rất thường xuyên	232	917	323	69
	Tổng cộng	3103	3103	3103	3103

Theo Bảng 2.2, một nét nổi bật dễ nhận thấy là nhóm ý kiến “thường xuyên”, và nhóm ý kiến “thỉnh thoảng/có đánh giá” thường chiếm tỉ lệ cao nhất so với các nhóm ý kiến khác đánh giá về từng item trong câu hỏi 4.

Để phân tích sâu hơn về đánh giá của GV về các hình thức kiểm tra đánh giá, các số liệu thống kê mô tả đối với từng câu hỏi nhỏ (item) được sử dụng gồm giá trị trung bình ( $\bar{X}$ ), độ lệch chuẩn ( $SD$ ), trung tuyến (Median), mode, độ tin cậy ( $\alpha$ ) và tương quan ( $r$ ) giữa các item với câu hỏi chính.

4) Số lượng các bài kiểm tra trên lớp của giáo viên các cấp

Ý kiến đánh giá của GV về số lượng các bài kiểm tra trên lớp (Câu hỏi 5) được thể hiện qua Bảng 10.

Để phân tích sâu hơn về đánh giá của GV về các số lượng các bài kiểm tra trên lớp, các số liệu thống kê mô tả đối với từng câu hỏi nhỏ (item) được sử dụng gồm giá trị trung bình ( $\bar{X}$ ), độ lệch chuẩn ( $SD$ ), trung tuyến (Median), mode, độ tin cậy ( $\alpha$ ) và tương quan ( $r$ ) giữa các item với câu hỏi chính.

5) Kinh nghiệm tham gia những hoạt động kiểm tra đánh giá của giáo viên các cấp

Ý kiến đánh giá của GV về kinh nghiệm tham gia những hoạt động kiểm tra đánh giá (Câu hỏi 6) cho thấy nhóm ý kiến “tham gia các hoạt động KT-ĐG cấp trường” chiếm tỉ lệ cao nhất so với các nhóm ý kiến khác đánh giá về từng item trong câu hỏi 6, nghĩa là đa số GV chỉ tham gia các hoạt động KT-ĐG ở cấp trường, ít tham gia đến các cấp khác ngoài trường. Và tiếp theo là nhóm ý kiến “chưa tham gia các hoạt động KT-ĐG” chiếm tỉ lệ cao thứ hai. Tiếp theo nữa là nhóm ý kiến “tham gia các hoạt động KT-ĐG ở cấp phòng giáo dục” đứng thứ ba về tỉ lệ, tuy nhiên đây là 1 tỉ lệ khá nhỏ so với hai nhóm tham gia cấp trường và không tham gia ở trên.

Để phân tích sâu hơn về đánh giá của GV về kinh nghiệm tham gia những hoạt động kiểm tra đánh giá, các số liệu thống kê mô tả đối với từng câu hỏi nhỏ (item) được sử dụng gồm giá trị trung bình ( $\bar{X}$ ), độ lệch chuẩn ( $SD$ ), trung tuyến (Median), mode, độ tin cậy ( $\alpha$ ) và tương quan ( $r$ ) giữa các item với câu hỏi

chính. Các số liệu mô tả về câu hỏi 6 (Q6) được thể hiện qua Bảng 3.16 dưới đây:

### 3. Kết Luận

Kết quả khảo sát cho thấy hầu hết GV tiếng Anh khẳng định rằng KT-ĐG là 1 trong những thành tố quan trọng nhất trong dạy học ngoại ngữ. Qua các điểm số đánh giá có thể nhận thấy các GV không cho rằng vai trò của KT-ĐG chỉ dừng lại ở việc xét tốt nghiệp. Nghĩa là họ đều nhận thấy tầm quan trọng của KT-ĐG trong việc nâng cao chất lượng dạy học và điều chỉnh quá trình dạy học ngoại ngữ. Tuy nhiên việc kiểm tra đánh giá ở các cấp học chưa hoàn toàn phù hợp với nội dung chương trình giáo trình, đặc biệt là bài thi tốt nghiệp THPT Quốc gia. Kết quả đánh giá về vai trò của KT-ĐG giữa 3 nhóm GV tiểu học, THCS và THPT thì giáo viên THPT đánh giá cao nhất về vai trò của KT-ĐG, sau đó là GV THCS, và GV tiểu học là nhóm đánh giá thấp nhất về vai trò của của KT-ĐG; và có sự khác biệt có ý nghĩa về chênh lệch giữa GV tiểu học và THPT.

GV tiếng Anh thường xuyên sử dụng một số hoạt động KT-ĐG trong khi dạy học. Tuy nhiên mức độ thực hiện thường xuyên hoạt động nào đó có thể không hoàn toàn tương quan thuận với khả năng thực hiện thành thạo hoạt động trên. Hầu hết GV tự nhận là biết thiết kế các bài KT-ĐG để sử dụng trên lớp. Tuy nhiên đa số GV thừa nhận có một số các hoạt động GV ít sử dụng trên lớp, điều này có liên quan với khả năng không thực hiện thành thạo hoạt động đó. Qua phỏng vấn, gần như tất cả GV thừa nhận rằng họ không thể đo độ khó, độ tin cậy, độ phân biệt, độ xác trị của các bài kiểm tra do họ thiết kế, họ cũng không thể thiết kế và đảm bảo các ma trận kiến thức kĩ năng để đảm bảo ra đề kiểm tra bám sát nội dung chương trình học.

### Tài liệu tham khảo

1. Alderson, C. J., & Wall, D. (1993). Does washback exist? *Applied Linguistics*, 14(2), 115–129. [https://doi.org/10.1093/a\\_pplin/14.2.115](https://doi.org/10.1093/a_pplin/14.2.115).
2. Brown, D. H. (2004). *Language assessment principles and classroom practices*. Pearson Education.
3. Barnes, M. (2016). The Washback of the TOEFL iBT in Vietnam. *Australian Journal of Teacher Education*, 41(7), 158–174. <https://doi.org/10.14221/ajte.2016v41n7.10>.
4. Chính phủ nước CH XHCN Việt Nam (2008). Quyết định số 1400/QĐ-TTg. Quyết định của Thủ tướng Chính phủ ban hành Đề án “Dạy và học tiếng Anh trong hệ thống giáo dục quốc dân, giai đoạn 2008-2020.