ĐẢM BẢO TÍNH GIÁ TRỊ NỘI DUNG CỦA TRẠM THI OSCE HỎI BỆNH SỬ CHO SINH VIÊN Y ĐA KHOA NĂM THỨ HAI

Lê Quốc Bảo[1], Nguyễn Thị Mỹ Hạnh[1], Hồ Hội[2]

[1]Đại học Y Dược TP HCM, Việt Nam

[2]Trung tâm Khoa học Sức khỏe Đại học Công nghệ Texas El Paso, Hoa Kỳ

TÓM TẮT

Giới thiệu: Đánh giá lâm sàng có cấu trúc khách quan (OSCE) là một phương pháp hiệu quả để lượng giá năng lực lâm sàng của người học trong một môi trường học tập an toàn và thuận lợi. Là một bộ phận không thể thiếu của Đại học Y Dược TP. HCM (UMP HCMC), Trung tâm Huấn luyện Nâng cao Mô phỏng Lâm sàng (ATCS) cam kết chuẩn hóa quy trình lượng giá kỹ năng lâm sàng và nâng cao chất lượng giáo dục, đặc biệt là đảm bảo tính giá trị nội dung của các kỳ thi OSCE trong cả giai đoạn tiền lâm sàng và lâm sàng. Mục tiêu của nghiên cứu nhằm đo lường tính giá trị nội dung của bảng kiểm trạm hỏi bệnh sử trong kỳ thi OSCE cuối kỳ cho sinh viên y đa khoa năm thứ hai.

Phương pháp: Chúng tôi sử dụng quy trình đánh giá tính giá trị nội dung gồm 6 bước và áp dụng thang đo 4 mức độ phù hợp (không phù hợp, khá phù hợp, phù hợp, rất phù hợp) cho biểu mẫu đánh giá. Biểu mẫu này được gửi cho hội đồng đánh giá bao gồm 6 giảng viên là những chuyên gia về mô phỏng lâm sàng và lượng giá. Những giảng viên này đã xem xét cẩn thận, góp ý điều chỉnh và cho điểm một cách độc lập các khía cạnh năng lực và chi tiết trong bảng kiểm hỏi bệnh sử dựa trên mức độ phù hợp. Chúng tôi đã tính toán bốn thông số của chỉ số giá trị nội dung (CVI): CVI tiểu mục (ICVI), trung bình CVI bảng kiểm (S-CVI/Ave), đồng thuận chung CVI bảng kiểm (S-CVI/UA), tỷ lệ giá trị nội dung (CVR), và so sánh với ngưỡng đề xuất của CVI là ≥ 0.83) và CVR là ≥ 0.8.

Kết quả: Trong số 37 tiểu mục của bảng kiểm hỏi bệnh sử ban đầu, 14 tiểu mục (37,84%) có I-CVI dưới 0,83; hầu hết trong số đó (10/14) thuộc về kỹ năng khai thác bệnh sử. Bảng kiểm này có S-CVI/Ave là 0,80, S-CVI/UA là 0,24 và CVR là 0,60. Sau khi điều chỉnh dựa trên góp ý của hội đồng đánh giá, bảng kiểm được rút ngắn chỉ còn 17 tiểu mục và các chỉ số giá trị nội dung gia tăng đáng kể (S-CVI/Ave = 0,98, S-CVI/UA = 0,88 và CVR = 0,96).

Kết luận: Chất lượng và tính giá trị của trạm thi OSCE kỹ năng hỏi bệnh sử đã được cải thiện đáng kể sau khi áp dụng quy trình đánh giá tính giá trị nội dung của bảng kiểm OSCE với sự tham gia của các chuyên gia về mô phỏng lâm sàng và lượng giá. Dựa trên những kết quả tích cực này, chúng tôi sẽ áp dụng rộng rãi quy trình đánh giá tính giá trị nội dung cho tất cả các kỳ thi OSCE để không ngừng nâng cao chất lượng và giá trị các công cụ lượng giá của trung tâm ATCS.

Từ khóa: tính giá trị, quy trình đánh giá tính giá trị, chỉ số giá trị nội dung, tỷ lệ giá trị nội dung, I-CVI, S-CVI/Ave, S-CVI/UA, CVR, ATCS, OSCE.

ABSTRACT

PILOT STUDY TO ENHANCE THE CONTENT VALIDITY
OF HISTORY TAKING OSCE FOR THE SECOND-YEAR MEDICAL STUDENT

Q.B. Le[1], T.M.H. Nguyen[1], H. Ho[2]
[1]University of Medicine and Pharmacy at Ho Chi Minh City (VIETNAM)
[2]Texas Tech University Health Sciences Center El Paso (UNITED STATES)

Introduction: Objective Structured Clinical Examination (OSCE) is a well-established platform to assess the learners' clinical ability in a safe and conducive learning environment. As an integral component of the University of Medicine and Pharmacy –Ho Chi Minh City– Vietnam (UMP-HCMC), the Center for Advanced Training in Clinical Simulation (ATCS) is committed to standardize the clinical skills assessment process and to enhance the quality and especially the validity of the OSCE platform in assessing the learners' competency in both pre-clinical and clinical phases. We decided to initiate a pilot study to measure the content validity of the History Taking checklist of the OSCE station for second-year medical students.

Methods: We used the 6-step content validation process that is commonly reported in the literature, applied the 4-level rating scale of relevance (not relevant, somewhat relevant, relevant, highly relevant) to prepare the content validation form, and sent it to the review panel that consisted of 6 faculty assessors who are experts in standard setting and clinical simulation. These assessors critically reviewed, refined, and independently scored the domain and its related checklist items according to their relevance. We calculated four parameters of the Content Validity Index (CVI): Item-CVI (ICVI), Scale-CVI/ Average (S-CVI/Ave), Scale-CVI/ Universal Agreement (S-CVI/UA), and Content Validity Ratio (CVR), and interpreted the results according to the recommended cut-off score of CVI (=> 0.83) and CVR (=>0.8).

Results: Of the original 37-item checklist, 14 (37.84%) had I-CVI under 0.83; most of them (10/14) belong to the competency of taking present illness. The S-CVI/Ave was 0.80, S-CVI/UA was 0.24, and CVR was 0.60. The revised checklist was reduced to only 17 items, and S-CVIs increased dramatically (S-CVI/Ave = 0.98, S-CVI/UA = 0.88, and CVR = 0.96).

Conclusion: The validation process for content validity of the OSCE's checklists conducted by faculty assessor experts in standard setting and clinical simulation has greatly improved the quality and validity of the history taking OSCE station. The positive results of this pilot study have prompted us to continue this validation process for all current and future OSCE checklists to improve the quality and validity of our assessment tools.

Keywords: Validity, validation process, content validity index, content validity ratio, I-CVI, S-CVI/Ave, S-CVI/UA, CVR, ATCS, OSCE.

INTRODUCTION

Objective Structured Clinical Examination (OSCE) is a well-established platform to assess the learners' clinical ability in a safe and conducive learning environment. As an integral component of the University of Medicine and Pharmacy –Ho Chi Minh City– Vietnam (UMP-HCMC), the Center for Advanced Training in Clinical Simulation (ATCS) is committed to standardize the clinical skills assessment process and to enhance the quality and especially the validity of the OSCE platform in assessing the learners' competency in both pre-clinical and

clinical phases. We decided to initiate a pilot study to measure the content validity of the History Taking checklist of the OSCE station for second-year medical students.

Our study aimed to (1) establish a panel of experts to review checklist of the summative OSCE on history taking skills station for MS-2; (2) ensure the relevance and representativeness of the checklist; (3) calculate CVIs of the original and revised checklist; and (4) interpret results of the study.

METHODOLOGY

We used the Yusoff's six steps approach to measure the content validity and revise the checklist in order to enhance the quality and validity of the summative OSCE for assessing history taking skills of MS2.

Step 1: Preparing content validation form.

We prepared the content validation form to assist the review panel about expectations of the learner's performance. The 4-level rating scale of relevance will be used for scoring individual items (Box 1) that are representative of tasks stated in the included door sign (Box 2). The rating of checklist relevance is based on the scenarios and scripted information provided by the standardized patients (SP)

Dear Experts and Faculties,

This OSCE checklist used for the **HISTORY TAKING** station contains **04 domains** (A, B, C, D), and **37 items** which are scored on a scale of 0 (not performed) – 1 (completely implemented). The objectives of this checklist are to assess the communication skills (domains A, C), history-taking skills (domain B), and professionalism (domain D) when students communicate with a SP.

We need your expert judgment on the degree of relevance of each item to the measured domains. Your review should be based on the definition and relevant terminologies that are provided to you. Please be as objective and constructive as possible in your review and use the following rating scale.

Degree of relevance:

1 = The item is **NOT** relevant to the measured domain

2 = The item is **SOMEWHAT** relevant to the measured domain

3 = The item is **QUITE** relevant to the measured domain

4 = The item is **HIGHLY** relevant to the measured domain

Box 1. *The Content Validation Form with Domains, their Definitions, and Individual Items of the History Taking Checklist*

Information: A 47-year-old male/ female teacher has complained of "cough a lot" during the last 4 days.

Tasks: Please communicate and take the history of the patient pertaining to her/his chief complaint.

Duration: 06 minutes (DO NOT PERFORM PHYSICAL EXAMINATION)

<u>**NOTE:**</u> Students will write down this patient's medical history at the next station.

Box 2. *The Door Sign of the history-taking station*

Step 2: Selecting a review panel of experts.

The selected panel includes individuals who have a minimum of 2 years of experience in standard setting, performance assessment and clinical simulation. All of them completed at least one comprehensive course of faculty development.

Table 1. *The Review Panel of Experts*

No.	Full Name	Qualification	Expertise (years)
1	Doan Thi Thu Hoa, MD. MHPE.	CEP*, AFDP†	21
2	Nguyen Thi My Hanh, MD. MHPE.	CEP*, AFDP†	21
3	Le Quoc Bao, MD. MSc.	CEP*, AFDP†	5
4	Ho Ngoc Loi, MD. MSc.	CEP*, AFDP†	5
5	Doan Truc Quynh, MD. MSc.	In-house FDC‡	3
6	Tang My Ngan, MD. MSc.	In-house FDC‡	2

CEP Collaborative Educational Program; AFDP† Advanced Faculty Development Program; FDC‡ Faculty Development Courses*

Step 3: Conducting content validation.

The ATCS Chief of Education Division conducted the session on content validation for the expert panel who reviewed and rated the scenarios and its checklist.

Figure 1. *The content validation meeting*

Step 4: Reviewing domain and items.

The experts critically reviewed the scenario, scripted SP information, the domain and its items before providing a score on each item. The experts were encouraged to provide verbal comments or written comments to improve the relevance of items to the targeted domain. (Please refer to Appendix 2 for details)

Step 5: Providing a score on each item.

The experts assigned a score on each checklist item based on the provided relevance scale.

Step 6: Calculating CVI

We calculated item CVI (I-CVI),scale CVI (S-CVI), average scale CVI (S-CVI/Ave) and universal agreement S-CVI (S-CVI/UA) [3-6].

In steps 5 and 6, the data were recorded in Microsoft form and analysed by using MS Excel, and SPSS 20. If I-CVI, S-CVI/Ave, and S-CVI/UA >= 0.83, they meet a satisfactory level.

RESULTS

1. The Content Validity of the original 37-item checklist

Table 2. *The calculated CVIs of the previous 37-item checklist*

Item	Expert 1	Expert 2	Expert 3	Expert 4	Expert 5	Expert 6	I-CVI	UA
1	1	1	0	1	1	1	0.83	0
2	1	1	0	1	1	1	0.83	0
3	0	1	1	1	1	1	0.83	0
4	1	1	1	1	0	1	0.83	0
5	1	0	1	1	1	1	0.83	0
6	1	0	1	1	1	1	0.83	0
7	0	0	1	1	1	1	0.67	0
8	0	1	1	1	1	0	0.67	0
9	1	1	1	0	1	0	0.67	0
10	1	0	1	0	1	1	0.67	0
11	1	1	1	1	1	1	1.00	1
12	1	1	1	1	1	1	1.00	1
13	1	1	1	0	1	0	0.67	0
14	1	1	1	0	1	1	0.83	0
15	0	1	1	1	1	1	0.83	0
16	0	1	1	1	1	1	0.83	0
17	1	1	1	1	1	1	1.00	1
18	1	1	1	0	1	1	0.83	0
19	1	1	1	1	1	1	1.00	1
20	1	1	1	1	1	1	1.00	1
21	0	1	0	1	1	1	0.67	0
22	0	1	0	0	1	1	0.50	0
23	1	1	1	1	0	1	0.83	0
24	1	1	1	1	0	1	0.83	0
25	1	1	1	1	1	1	1.00	1
26	1	1	1	1	1	1	1.00	1
27	1	1	1	1	1	1	1.00	1
28	0	0	1	1	1	1	0.67	0
29	1	1	0	0	1	1	0.67	0
30	1	1	1	0	0	1	0.67	0
31	1	1	1	1	0	0	0.67	0
32	1	1	0	1	1	0	0.67	0
33	1	1	1	1	1	0	0.83	0
34	1	0	0	1	1	1	0.67	0
35	1	1	0	1	1	1	0.83	0
36	1	1	1	1	1	1	1.00	1
37	1	1	0	0	1	0	0.50	0
S-CVI/Ave				0.80				
S-CVI/UA							0.24	

As shown in Table 2, the S-CVI of the total checklist is 0.80 and CVR was 0.60. There are 14 items out of 37 items (37.84%) that have the I-CVI under 0.83. Among these low valid items, most of them (10/14 = 71.43%) belong to the competencies of taking history in domain B.

2. *The Content Validity of the revised 17-item checklist*

After the panel discussion, the revised checklist contained only17 items). This remarkable reduction is due to either deletion or combination of items. The new checklist has a wider scale range of 0, 1, and 2 for the combined items (items 1, 5, 6, 7, 14, 15, and 17 in Appendix 2). Traditionally the binary rating of checklists (0 and 1) appeared to have ability to provide an objective assessment and lead to greater inter-rater reliability; however, a growing body of evidence showing that objectivity does not necessarily translate into greater reliability [8]. This is particularly applicable if expert examiners are used in an OSCE [9].

Table 3. *The CVIs of the revised 17-item checklist*

Item	Expert 1	Expert 2	Expert 3	Expert 4	Expert 5	Expert 6	I-CVI	UA
1	1	1	1	1	1	1	1.00	1
2	1	1	1	1	1	1	1.00	1
3	1	1	1	1	1	1	1.00	1
4	1	1	1	1	1	1	1.00	1
5	1	1	1	1	1	1	1.00	1
6	1	1	1	1	1	1	1.00	1
7	1	1	1	1	1	1	1.00	1
8	1	1	1	1	1	1	1.00	1
9	1	1	1	1	1	1	1.00	1
10	1	1	1	1	1	1	1.00	1
11	1	1	1	1	1	1	1.00	1
12	1	1	1	1	1	1	1.00	1
13	1	1	1	1	1	1	1.00	1
14	1	1	1	1	1	1	1.00	1
15	0	1	1	1	1	1	0.83	0
16	0	1	1	1	1	1	0.83	0
17	1	1	1	1	1	1	1.00	1
S-CVI/Ave				0.98				
S-CVI/UA							0.88	

3. *Comparison of the original and revised checklist*

Table 4 shows significant differences between the two checklist versions. Domain B (history-taking skills) accounts for the vast majority of the checklist, 72.97% and 66.67% respectively. All markers for content validity increased dramatically in the revised checklist (S-CVI/Ave = 0.98, S-CVI/UA = 0.88, and CVR = 0.96).

Table 4. *Psychometric comparison between two checklist versions*

	37-item checklist	17-item checklist	P value
Domain A Initiation	8.11%	12.50%	< 0.01
Domain B History taking	72.97%	66.67%	< 0.01
Domain C Communication	13.51%	12.50%	< 0.01
Domain D Professionalism	5.41%	8.33%	< 0.01
S-CVI/Ave	0.80	0.98	< 0.01
S-CVI/UA	0.24	0.88	< 0.01
CVR	0.60	0.96	< 0.01

CONCLUSIONS

The validation process for content validity of the OSCE's checklists conducted by faculty assessor experts in standard setting and clinical simulation has greatly improved the quality and validity of the history taking OSCE station. By using this process, the checklist renewal is shorter, valid, and more friendly for assessors. The positive results of this pilot study have prompted us to continue this validation process for all current and future OSCE checklists to improve the quality and validity of our assessment tools.

REFERENCES

[1]   M.S.B. Yusoff, "ABC of Content Validation and Content Validity Index Calculation," *Education in Medicine Journal*, 2019.

[2]   M.S.B. Yusoff, "A systematic review on validity evidence of Medical Student Stressor Questionnaire," vol. 9, pp. 1-16, 2017.

[3]   D.A. Cook, and T. J. Beckman, "Current concepts in validity and reliability for psychometric instruments: theory and application," *Am J Med*, vol. 119, no. 2, pp. 166.e7-16, Feb 2006.

[4]   S.N. Haynes, D.C.S. Richard, and E.S. Kubany, "Content validity in psychological assessment: A functional approach to concepts and methods," *Psychological Assessment*, vol. 7, pp. 238-247, 1995.

[5]   M.R. Lynn, "Determination and quantification of content validity," *Nurs Res*, vol. 35, no. 6, pp. 382-5, Nov-Dec 1986.

[6]   D.F. Polit, C.T. Beck, and S.V. Owen, "Is the CVI an acceptable indicator of content validity? Appraisal and recommendations," *Res Nurs Health*, vol. 30, no. 4, pp. 459-67, Aug 2007.

[7]   R.M. Harden, M. Stevenson, W.W. Downie, et al., "Assessment of clinical competence using objective structured examination," *Br Med J*, vol. 1, no. 5955, pp. 447-51, Feb 22, 1975.

[8]   T.J. Wilkinson, C.M. Frampton, M. Thompson-Fawcett et al., "Objectivity in objective structured clinical examinations: checklists are no substitute for examiner commitment," *Acad Med*, vol. 78, no. 2, pp. 219-23, Feb 2003.

[9]   B. Hodges, and J.H. McIlroy, "Analytic global OSCE ratings are sensitive to level of training," *Med Educ*, vol. 37, no. 11, pp. 1012-6, Nov 2003.