

DEVELOPING A PROGRAM FOR EVALUATING THE ACTIVITIES OF PROFESSORS AND TEACHERS IN CLASS LESSONS IN HIGHER EDUCATION INSTITUTIONS: METHODOLOGICAL APPROACH AND CULTURAL ADAPTATION ISSUES

 <https://doi.org/10.70728/tech.v3.i04.009>

Faxriddinov Muhammad Faxriddin o'g'li
Samarkand State Pedagogical Institute,
PhD Candidate.
fakhriddinvmukhammad0@gmail.com

Abstract. This article analyzes the methodological foundations for developing a program to evaluate the activities of professors and teachers during class lessons in higher education institutions. In recent years, serious discussions have been ongoing internationally regarding the reliability, validity, and cultural appropriateness of Student Evaluations of Teaching (SET) systems. Research has identified shortcomings of traditional assessment tools, such as biases related to demographic factors, excessive focus on student satisfaction, and failure to fully reflect teaching effectiveness. Synthesizing advanced research conducted in Thailand, Ecuador, the USA, Indonesia, and other countries, the article elucidates a seven-stage process of cultural adaptation of assessment tools, psychometric testing methods, and practical application mechanisms. Furthermore, through a comparative analysis of assessment approaches in Eastern and Western universities, practical recommendations have been developed for creating an adapted assessment program for the higher education system of Uzbekistan.

Keywords: teacher performance evaluation, student evaluation of teaching, cultural adaptation, validation, SET instruments, higher education, psychometric properties, IMRAD, evaluation program, S-IMPACT, SEEQ.

Introduction. Systems for evaluating teacher performance and course usefulness for students in higher education institutions have become an integral part of the academic environment in recent decades. The results of Student Evaluations of Teaching (SET) are used not only to improve teachers' pedagogical skills but also as one of the main criteria in making important personnel decisions such as promotion, tenure, and financial incentives [5]. However, recent research is revealing serious shortcomings of traditional SET systems.

Research conducted by University of Nebraska-Omaha scientists Christopher Moore and Tracie Reding within a project funded by the National Science Foundation (NSF) shows that existing assessment models have biases related to demographic, disciplinary, and personal characteristics [2]. These systems often focus on student satisfaction levels rather than actual learning outcomes and suffer from reliability issues [2][5]. A study by Chinese researchers Ding Yan, Wang Ying, and Chen Kan, covering 24 world-renowned

universities, identified another important aspect – cultural and regional differences in assessment approaches. According to it, European universities tend towards a "development model" of assessment, while Asia-Pacific universities predominantly use an "attainment model" [4].

These problems necessitate the development of new, scientifically grounded, and culturally adapted programs for evaluating teacher performance and course usefulness in higher education institutions. Particularly for the Uzbek higher education system, applying international experience adapted to the local context is a pressing scientific and practical task.

The purpose of this article is to systematize international experience in developing programs for evaluating teacher performance and course usefulness for students in higher education, analyze methodological approaches, and develop practical recommendations applicable in the context of Uzbekistan.

Methodology. In preparing the article, methods of systematic literature analysis, comparative analysis, and generalization were used. International scientific articles, conference materials, and institutional reports published between 2019 and 2025 were selected as the research object. When selecting sources, attention was paid to their indexing in international scientific databases such as Scopus and Web of Science, their methodological robustness, and practical significance.

Geographically, the analyzed sources cover the regions of East Asia (China, Japan, Korea), Southeast Asia (Indonesia, Thailand, Malaysia), the Middle East (Turkey, Jordan), Europe (Spain, the Netherlands), and the Americas (USA, Mexico, Canada). This enabled a comparative analysis of different cultural approaches to assessment systems.

Results. Methodology for Cultural Adaptation of Assessment Tools

The analyses show that when transferring assessment tools from one cultural context to another, it is necessary to re-examine and adapt their psychometric properties. The seven-stage cultural adaptation process developed by Thai scientists Chanisra S., Sanguantrakul T.R., and Puranitee P. from Mahidol University in Thailand is recognized as the most comprehensive methodological guide [3]. This process includes the following stages:

1. **Literature Analysis** – identifying 44 teacher performance characteristics from 29 scientific articles.
2. **Focus Groups** – conducting 6 focus groups with 45 students, gathering 252 descriptions.
3. **Delphi Method** – selecting 23 important characteristics through a two-round survey involving 35 experts.
4. **Content Validity Index (CVI)** – ensuring a CVI value >0.8 for each item.
5. **Cognitive Interviews** – verifying that students correctly understand the questions.
6. **Pilot Testing** – initial trial on a small sample.
7. **Exploratory Factor Analysis (EFA) and Reliability Testing.**

The Thai clinical teacher evaluation form developed based on this methodology has a two-factor structure (Clinical Facilitator – 14 items and Supporting Professional Identity – 6 items), achieving a high reliability indicator like Cronbach's $\alpha = 0.976$ [3].

Research on the cultural adaptation of tools measuring the "flipped classroom" experience, conducted by the Monash University research center in Malaysia, substantiates why adapting tools to a new context is important. The diversity of student responses from different countries and cultures makes the direct translation of assessment tools impossible [6].

Analysis of Main Assessment Tools Used Internationally

Based on the analyzed sources, the most widely used and psychometrically robust assessment tools in higher education are the following:

SEEQ (Student Evaluation of Educational Quality) – Developed by the University of Maryland, this tool consists of 9 dimensions and 34 questions [5]. Its dimensions include: Learning, Enthusiasm, Organization, Group Interaction, Individual Rapport, Breadth of Coverage, Examinations, Assignments, and Overall Rating. A 5-point Likert scale is used.

S-IMPACT and F-IMPACT Tools – Developed at the University of Nebraska-Omaha, these tools are based on the Concerns-Based Adoption Model (CBAM) change theory [2]. S-IMPACT is intended for student evaluation, while F-IMPACT is for teacher self-evaluation. Both tools have been validated against classroom observations (COPUS – Classroom Observation Protocol for Undergraduate STEM). Tests conducted across 14 courses (N=691) confirmed the tools' reliability and validity.

SPOT (Student Perceptions of Teaching) – Developed at West Virginia University, this tool places special emphasis on open-ended questions. Open-ended questions, used in conjunction with a student consent protocol, are based on the concept of "useful for my learning" [5].

EEDocente (Teacher Performance Evaluation Scale) – Developed at the Autonomous University of Baja California in Mexico, this tool was tested with the participation of 1,422 students. The two-factor model explains 68% of the variance, and measurement invariance across different groups (age, gender) has been proven for the tool [4].

Comparative Analysis of Eastern and Western Approaches

Research conducted by Ding Yan and colleagues identified fundamental differences in assessment approaches between Eastern and Western universities. In European universities, the "developmental assessment" model prevails, meaning assessment results primarily serve to improve the teaching process. In East Asian universities, the "attainment assessment" model is widespread, where assessment is more focused on recording and reporting the final outcomes of teacher performance [4].

Research by André et al. (2020) conducted across 6 countries (Spain, Turkey, South Africa, South Korea, Indonesia, the Netherlands) proved that cultural and contextual differences play a significant role in how teaching practice is evaluated by students [4].

Measurement invariance analysis showed that students from different countries approach the same questions differently.

Methods for Testing the Psychometric Properties of Assessment Tools.

In modern research, the following methods are used to test the psychometric properties of assessment tools:

Content Validity Index (CVI) – Expert evaluation of how well each item's content matches the construct being measured. The Thai study required a CVI value >0.8 [3].

Exploratory Factor Analysis (EFA) – Used to determine the factor structure of the tool's items. In research conducted by Aniel M.K. in the Philippines, EFA with principal component analysis and varimax rotation was used to develop a scale reduced from 37 items to 12 items [9].

Confirmatory Factor Analysis (CFA) – Tests whether the identified factor structure fits the theoretical model. In the study by Boukari et al. (2025) adapting the Engaged Teachers Scale (A-ETS) for Arabic-speaking physical education teachers, CFA in a second-order model recorded excellent fit indicators: $\chi^2=90.91$, $df=101$, $CFI=0.98$, $RMSEA=0.045$ [6].

Reliability Analysis – Assessing internal consistency through Cronbach's alpha (α) coefficient. In the psychometric evaluation of the Arabic version of the TEACH (Teachers' Evaluation of Aural/Oral Performance of Children) questionnaire, high indicators were obtained: Cronbach's $\alpha = 0.96$ (overall), 0.94 (auditory function), and 0.95 (listening ease) [3].

Test-Retest Reliability – Assessing the stability of the tool over time using the intraclass correlation coefficient (ICC). In the aforementioned study, ICC values were 0.985 , 0.971 , and 0.991 [3].

Practical Application Mechanisms and Question Structure

The following important aspects regarding the practical application of assessment tools were identified in the analyzed sources:

Open-Ended Question Protocols – In the West Virginia University experience, students are given special instructions before open-ended questions: "Please provide constructive feedback... remember that open expression in an environment of courtesy, sensitivity, inclusivity, and mutual respect is valued" [5]. Open-ended questions are formulated as follows to obtain qualitative data:

- "Please provide constructive suggestions on how this course could be changed to improve your learning."
- "Please provide examples of things in the course that were helpful for your learning and should be continued."

"Because" Formula – The 10 core questions developed by the research group at the University of Alabama in Huntsville are structured based on the "because" formula. For example: "Because the course was well organized, I understood what I needed to do to be successful" [8].

Importance-Performance Analysis (IPA) – This approach, used at the University of the Balearic Islands in Spain, allows teachers to visually see which factors are important to students and how well they are performing on those factors [8].

Discussion. Developing a program for evaluating teacher performance and course usefulness for students in higher education institutions is a complex and multifaceted task. The analysis results indicate several significant trends and problems in this area.

Firstly, the shortcomings of traditional SET systems are being recognized internationally, and new approaches aimed at addressing them are being developed. The S-IMPACT and F-IMPACT tools presented by Moore and Reding are an important step in this regard, enabling a comprehensive assessment of teacher performance not only from the students' perspective but also through teacher self-evaluation and independent observations [2]. This approach serves to increase the reliability and objectivity of evaluation.

Secondly, the issue of culturally adapting assessment tools acquires particular importance. The international adaptation guidelines proposed by Beaton et al. (2000) are widely used today in various countries [3][9]. Research conducted in Thailand, Indonesia, Malaysia, and Arab countries shows that cultural adaptation of tools requires not only linguistic translation but also conceptual and contextual alignment [3][6][9].

Thirdly, the difference between Eastern and Western approaches once again confirms the importance of considering cultural factors when developing assessment systems. As identified in the Chinese study, European universities tend towards a "developmental assessment" model, while Asian universities predominantly use an "attainment model" [4]. Determining which model is more suitable for the Uzbek higher education system requires separate research [7].

Fourthly, it is necessary to use modern statistical methods when testing the psychometric properties of assessment tools. Methods such as exploratory and confirmatory factor analyses, content validity index, Cronbach's alpha, and test-retest reliability ensure the scientific validity of the tool [3][6][9].

Fifthly, the mechanisms for using assessment results are of great importance. As researchers from Washington and Lee University emphasized, SET tools should be used not as the sole instrument for evaluating teaching effectiveness, but as part of a comprehensive system [5]. Assessment results should serve as a mechanism for teachers' professional development, not as a tool for punishment [7].

Conclusion. The development of a program for evaluating the activities of professors and teachers during class lessons in higher education institutions should be based on the following key principles:

1. **Methodological Foundation** – The seven-stage cultural adaptation process used in the Thai research can serve as a methodological basis for the Uzbek context as well. This process, which includes literature analysis, focus groups, Delphi method, content validity index, cognitive interviews, pilot testing, and factor analysis, guarantees the high quality of the instrument.

2. **Utilizing Existing Instruments** – Instead of creating a new instrument from scratch, adapting internationally recognized and psychometrically robust tools such as SEEQ, SPOT, and S-IMPACT to the local context is a more effective approach.

3. **Cultural Adaptation** – Based on Monash University's recommendations, it is necessary not only to translate the instruments linguistically but also to adapt them to the educational experiences, cultural values, and expectations of local students. Cognitive interviews play a crucial role in this process.

4. **Comprehensive Approach** – Based on the University of Nebraska's experience, it is advisable to create an integrated system combining student evaluations (S-IMPACT) with teacher self-evaluations (F-IMPACT) and classroom observations (COPUS).

5. **Integration of Quantitative and Qualitative Methods** – Based on the West Virginia University model, a system should be established that gathers qualitative data through open-ended questions alongside quantitative ratings on Likert scales.

6. **Psychometric Testing** – The developed instrument must be meticulously examined through content validity index (CVI), exploratory factor analysis (EFA), confirmatory factor analysis (CFA), and Cronbach's alpha reliability coefficient.

7. **Documenting the Process** – The process of instrument development and cultural adaptation should be documented in detail, as this constitutes the main contribution of the scientific research.

In developing an adapted assessment program for the Uzbek higher education system, it is necessary to adhere to the above principles, as well as to deeply analyze the specific characteristics of the local context (educational traditions, student-teacher relationships, assessment culture, student satisfaction, etc.). Future research should appropriately be conducted precisely in this direction.

REFERENCES

1. Asadullayeva, M. A. qizi. (2024). Mechanisms for assessing educational quality in higher education institutions. *Educational Research in Universal Sciences*, 3(3 SPECIAL), 35–38.
2. Moore, C., & Reding, T. (2025). Using instructor insights to generate scalable metrics [Conference presentation]. ASCN Transforming Institutions Conference, University of Nebraska–Omaha, USA.
3. Boukari, S., Guelmami, N., Chalghaf, N., Ceylan, H. İ., Stefanica, V., Mekni, R., Dhahbi, W., & Dergaa, I. (2025). Cross-cultural adaptation and psychometric evaluation of the Engaged Teachers Scale among Arabic-speaking physical education teachers. *Frontiers in Education*.
4. Chanisra, S., Sanguantrakul, T. R., & Puranitee, P. (2025). Development and testing of a culturally adapted clinical teacher evaluation form in Thailand. *Korean Journal of Medical Education*, 37(2), 145-162. [Scopus]

5. Woodzicka, J. A., Greer, L., Murdock, K. K., Johnson, D. R., Locy, T., & Goldsmith, A. H. (2025). One piece of the puzzle: Developing an empirically informed open-ended student evaluation of teaching. *Frontiers in Education*.
6. Sai, B., Mohamed, M., & Dergaa, I. (2025). Cross-cultural adaptation of the "Teacher Stress Questionnaire" into Arabic. *Frontiers in Education*.
7. Yuldashev, F. I. (2025). Prospects for developing an integrated system of indicators to increase the level of competitiveness in higher education. *Digital Economy*, 7(7), 440–451.
8. Vera-Leon, X. E., Rios-Lopes, T. D., & Kaizaluisa-Barros, N. F. (2025). A new method for measuring the effectiveness of teacher evaluation tools in improving pedagogical activity in higher education. *Neutrosophic Sets and Systems*, 58, 112-128. [Scopus]
9. Lutfatulatifah, Indreswari, H., Eva, N., Hotifah, Y., & De Guzman, D. D. (2025). Cross-cultural adaptation and psychometric evaluation of the Interpersonal Reactivity Index (IRI) for Indonesian pre-service early childhood teachers. *TPM – Testing, Psychometrics, Methodology in Applied Psychology*, 32(2).
10. Wagiran, W., & others. (2024). Development of a teaching process evaluation instrument in higher education. *Indonesian Journal of Educational Research and Technology*, 4(1), 23-38.
11. Aniel, M. K. (2025). Student perceptions in evaluating college teachers: Instrument development and testing. *PhilArchive*.
12. Marshik, T., Driscoll, A., Remsburg, A., Yu, A., Dutta, N., & Klein, J. (2023). New frontiers in student evaluations of teaching: University efforts to design and test a new instrument for student feedback. *Assessment & Evaluation in Higher Education*, 48(8), 1164–1177.
13. University of Maryland, Baltimore County. (2026). Student Evaluation of Educational Quality (SEEQ) [Institutional source]. UMBC Teaching Center.
14. West Virginia University. (2025). WVU SPOT: Student Perceptions of Teaching – Instructor's guide [Institutional source]. WVU Office of Assessment.
15. University of Alabama in Huntsville. (2025). Blue/Explorance Student Instruction Evaluation (SIE) implementation guide [Institutional source]. UAH Academic Affairs.
16. Monash University Malaysia. (2021). Cross-cultural adaptation and testing of instruments measuring the "flipped classroom" experience. *Currents in Pharmacy Teaching and Learning*, 13(5), 521-534.
17. Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*, 25(24), 3186–3191.