

February – 2023

The Design and Psychometric Properties of a Peer Observation Tool for Use in LMS-Based Classrooms in Medical Sciences

Zohresadat Mirmoghtadaie¹, Mohsen Keshavarz², Mojgan Mohammadimehr³, and Davood Rasouli⁴

¹ Assistant Professor, Department of e-Learning, Virtual School of Medical Education and Management, Shahid Beheshti University of Medical Sciences (SBMU), Tehran, Iran; ² Department of E-Learning in Medical Sciences, School of Paramedical Sciences, Torbat Heydariyeh University of Medical Sciences, Torbat Heydariyeh, Iran; ³ Department of Laboratory Sciences, Faculty of Paramedicine, AJA University of Medical Sciences, Tehran, Iran; ⁴ Assistant Professor, Center for Educational Research in Medical Sciences (CERMS), Department of Medical Education, School of Medicine, Iran University of Medical Sciences, Tehran, Iran; corresponding author

Abstract

In peer observation of teaching, an experienced colleague in the educational environment of a faculty member observes the educational performance of that faculty member and provides appropriate feedback. The use of peer review as an alternative source of evidence of teaching effectiveness is increasing. However, no research has been done in the field of tool design and development to peer review in classrooms that use a learning management system (LMS). This study used mixed methods. In the qualitative stage, after studying sources and interviewing professors active in virtual education, a question bank was prepared and a 26-item initial questionnaire created. In the quantitative stage, the psychometric properties of the developed instruments, such as the face, content, and structural validity, were examined, and reliability tests were performed. IBM SPSS Statistics (Version 20) was used for analysis. Five categories, including content preparation, content presentation, effective interactions, motivation management, and support services, and 26 subcategories were determined to be effective indicators in peer observation in LMS-based classes in medical sciences. During content analysis, 9 items were removed due to lack of necessary criteria. Then, using principal component analysis and varimax rotation in the present mode (Watkins, 2018), 5 components with eigenvalues higher than 1 were extracted, which explained a total of 70.55% of the total variance. The inter-cluster correlation coefficient (ICC) was 0.88. Thus, the peer observation measurement tool, designed with 17 expressions using the answer method “yes/no”, showed good validity and reliability. The research results demonstrate that the evaluation of virtual classes of professors by their peers is effective and that the results can be used in e-learning promotion plans.

Keywords: blended learning, virtual education, psychometrics, validity, reliability

Introduction

Online learning refers to teaching and learning processes that are provided through the Internet. It includes a wide range of applications to access educational materials, as well as to facilitate teacher-student interaction (Keshavarz, Mirmoghataei, & Nayyeri, 2022). In recent years, e-learning systems have been increasingly influencing both classroom and campus-based teaching, but more primarily, such systems are leading to new models or designs for teaching and learning (Bates, 2022). In March 2020, with the emergence of the coronavirus, most schools, colleges, and universities across the world were forced to close to protect students and staff from infection (OECD, 2021). Gradually, instructors adopted blended/hybrid learning methods and asynchronous learning in online teaching. During the pandemic, lectures were often recorded and made available to download and replay at any time on online platforms (Bates, 2022). As blended learning systems developed, components and interactions became more complicated, and as a result, the expectations of students and other stakeholders from this educational environment have increased (Andone & Sireteanu, 2009). It should be noted however that certain limitations of e-learning, such as the lack of face-to-face communication and human and emotional interaction, have been largely eliminated (Kintu, Zhu, & Kagambe, 2017; Pinto-Llorente, Sánchez-Gómez, García-Peñalvo, & Casillas-Martín, 2017).

The purpose of blended education is to provide opportunities for students to use both real and virtual spaces to better benefit from learning (Henrie, Bodily, Manwaring, & Graham, 2015). This method optimizes learning outcomes and cost-effectiveness (Donnelly, 2017). Training in the medical field, part of higher education, should provide a wide range of knowledge, attitudes, and skills to students to gain job qualifications (Wood, 2003). Improving the health of the community depends on the presence of efficient and high quality manpower, trained using these new educational methods (Twomey, 2004).

Today, in the digital age, one of the basic requirements of learners is that they have the skills to learn in new digital environments. For this reason, instructors must possess digital-age teaching skills and be familiar with ways to manage and lead online classes using new learning platforms (Keshavarz & Ghoneim, 2021). Since blended learning can provide the benefits of both traditional and virtual methods, it is a good way to achieve teaching-learning goals in medical education. A review of research institutes and universities all around the world looking at the mechanisms of blended learning in medicine shows that, in recent years, blended learning is being used more often than traditional methods such as face-to-face and class lectures. Blended learning is not only capable of a more efficient transfer of concepts and skills, but is also a more effective method of educating and training self-employed and creative graduates (Benner, 2012; Missildine, Fountain, Summers, & Gosselin, 2013).

One of the tasks of medical universities is to empower faculty members to play their role as teachers, and one of the successful and effective ways of achieving this is to use the capacities and experiences of faculty members themselves. Experienced and successful instructors in teaching can contribute to the professional growth and development of their colleagues (Speer, 2010). Nowadays, peer observation of teaching is one of the new components of empowerment programs or evaluation of faculty members in different universities around the world (Johnston, Baik, & Chester, 2020).

Various terms such as *peer review* and *peer evaluation* are used synonymously in the literature, but the most common term in this field is *peer review* or *peer observation of teaching* (POT; Speer, 2010). POT is the presence of an experienced colleague in the educational environment of a faculty member observing that faculty member's educational performance and providing appropriate feedback (Cunningham, Johnson, & Lynch, 2017). The goals of POT include generating awareness of strengths and weaknesses of teaching from the perspective of colleagues, motivating faculty members in order to improve the overall teaching process, improving the teaching ability of individual faculty members, and creating an opportunity to use the experiences of other faculty members in teaching and assessment methods (Fletcher, 2018).

POT provides formative and constructive summative feedback to faculty members for the growth and development of their teaching abilities (Fernandez & Yu, 2007). This facilitates the formation of reflection and thought in teaching processes, and greatly influences the attitude and approach of faculty members towards teaching (Bernstein, Burnett, Goodburn, & Savory, 2006).

According to various studies, the use of peer review as one of the alternative sources of evidence of teaching effectiveness is increasing (Fernandez & Yu, 2007). Peer review in teaching includes two main activities: observing peers' performance in the classroom; and, reviewing written documents used in a course (Gehring, Chinn, Pérez-Quñones, & Ardis, 2005). Research has reported many different POT methods, but all are based on peer review/observation. One model is based on four phases: preparation, peer visit, peer reporting, and promotion (Speer, 2010).

In the case of formative evaluation, it is necessary to hold symposiums and provide feedback. Fernandez and Yu (2007) identified four steps in peer review:

1. Review of the educational materials-syllabus-course guide, and a sample presentation (e.g., PowerPoint slides)
2. Observer interaction, teaching observation, counselling, and post-teaching feedback
3. Written evaluation and presentation to the relevant teacher
4. Monitoring the peer review process.

If evaluation is not done according to a predetermined framework, evaluator subjectivity and biases will occur due to factors such as camaraderie, cooperation, and negative feelings. Quality teaching is also important in e-learning (Dill, 2007; Ruiz, Candler, & Teasdale, 2007).

A learning management system (LMS) is software used to implement and evaluate a learning process. A LMS provides an instructor with a way to create and deliver content and monitor student performance. A LMS may also provide students with the ability to use interactive features such as video conferencing and discussion forums. Canvas, Blackboard, and Moodle are examples of LMSs in which teachers and students are able to log in and work within an online learning environment (Bates, 2022).

Using this software, instructors and students can enter the online learning environment at designated time intervals. Course materials are often presented as PowerPoint slides or as audio podcasts or videos. Instructors take charge of teaching and introducing course materials to students. Classes with a large number of students can be divided into small groups. Students have the opportunity to discuss the course online with both the teacher and other students, and at the end of the class, the professor evaluates the learning activities. The LMS is primarily asynchronous in that students can access the learning process at any time and any place with an Internet connection (Bates, 2022).

Despite the extensive research that has been done, we found that there has been no research in the field of tool design and development related to peer review in LMS-based classrooms. Therefore, this study aimed to identify and prioritize the effective issues in peer observation in the LMS-based class in medical sciences.

Methodology

The present study was carried out using a mixed-method approach. It was conducted at the Tehran University of Medical Sciences in 2020. The mean age of the professors participating was 44.36 years, with a standard deviation of 6.47 years. Just over half (54.4%) of participants were male, and the rest were female. They came from three universities: 37.9% were faculty members of the Tehran University of Medical Sciences, 31.9% were from the Iran University of Medical Sciences, and the rest were from Shahid Beheshti University of Medical Sciences.

Qualitative Stage

Semi-structured interviews were used to collect data at this stage. Following a systematic review of related texts and articles, the questions were developed. Preliminary questions were as follows: “What do you think about peer observation in LMS-based education?” “What do you think are the challenges of peer observation?” and “What is the viable solution for improving e-learning using peer review?”

The semi-structured interviews were conducted with expert professors who were selected by purposive sampling. Inclusion criteria were having experience in virtual teaching and willingness to participate in the study. Each interview was conducted at a time and place convenient to the interviewee. The interviews were conducted individually, and the duration of each was 30–45 minutes. All interviews were recorded and then transcribed. Content analysis was performed after each interview.

Quantitative Stage

In the quantitative section, the psychometric properties of the developed instruments such as face validity, content validity, construct validity, and reliability were examined. The questionnaire was developed based on information obtained during the qualitative stage. The sample consisted of faculty members of the Tehran, Iran, and Shahid Beheshti universities of medical sciences who were selected by available sampling method. Inclusion criteria in this stage were having at least two years' experience in virtual teaching and being interested in participating.

Face Validity

To evaluate face validity, two approaches were used, one qualitative and the other quantitative. In the study of qualitative face validity, items were corrected with a qualitative approach. The impact score index was used to determine the quantitative face validity (Mohammadbeigi, Mohammadsalehi, & Aligol, 2015; Neuendorf, 2017). To do this, a checklist tool with a 5-point Likert scale (1 = *not important at all* to 5 = *absolutely important*) was provided to 15 professors. After calculating the score of each question, questions with a score above 1.5 were deemed acceptable and saved for next steps. A score of 1.5 was considered the minimum acceptable score for an item (Lacasse, Godbout, & Series, 2002; Neuendorf, 2017).

Content Validity

To evaluate content validity, two approaches were used, one qualitative and the other quantitative. In the qualitative approach, a checklist was provided to 10 professors active in the field of virtual education to help them review and comment on issues such as observing Persian grammar, using the right words, placing the items in the right order, and the appropriateness of the items. Then, using their comments, we examined content validity using the content validity index (CVI) and content validity ratio (CVR) to quantify our findings. CVI was reviewed by 10 expert professors based on the formula proposed by Waltz and Bausell (1981). The total number of agreeable scores, i.e., “which is relevant but needs to be reviewed” and “fully relevant,” was divided by the total number of specialist professors, and the index scores with a content validity of less than 0.7 were removed. Scores between 0.7 and 0.79 were revised (modified based on the recommendations of the panel members and the research team), and scores above 0.79 remained unchanged on the checklist (Polit, Beck, & Owen, 2007). To determine the CVR, experts were asked to review each item based on a three-part range of “essential,” “useful but not essential,” and “not essential.” Then, answers were calculated according to the following formula, where N_e represents the number of panelists indicating “essential,” and N is the total number of panelists.

$$CVR = \frac{N_e - \frac{N}{2}}{\frac{N}{2}}$$

Based on the number of experts who evaluated the questions, the minimum acceptable CVR value in this study was determined to be 0.49, which is, in turn, based on the Lawshe table for 15 participating specialists. Questions for which the CVR value was less than the minimum were excluded from the test (Lawshe, 1975).

Construct Validity

Construct validity using exploratory factor analysis (EFA) after examining Kaiser-Meyer-Olkin (KMO) sampling adequacy indices and Bartlett’s test of sphericity, and after ensuring the ability to perform exploratory analysis with the participation of 182 faculty members of Tehran, Iran, and Shahid Beheshti universities of medical sciences was evaluated using principal component analysis and varimax rotation. In other studies, different ratios for the sample size required for EFA have been expressed. In this regard, a minimum ratio of subjects to variables has been reported as 1 to 3, 1 to 10, 1 to 15, as well as 1 to 20 (Stevens, 2012; Westen & Rosenthal, 2003).

Reliability

Considering that the final tool was a checklist with two options, yes and no, we gave the checklist to five faculty members to evaluate. The degree of their agreement was calculated based on the intraclass correlation coefficient (two-way mixed and consistency).

Data Analysis

Analysis of interview data at the qualitative stage of the study was performed through content analysis. We used Colaizzi's 7-step method which includes: (a) reading important findings to get a grasp on participants' understanding of the topic, (b) extracting important sentences related to the subject under study, (c) giving specific concepts to the extracted sentences, (d) classifying the concepts and clusters obtained, (e) referring to the main and comparative contents of the data, (f) describing the studied phenomenon, and finally, (g) returning the description of the phenomena to the participants to check reliability. After these steps were taken, the main categories and subcategories were coded and extracted (Drisko & Maschi, 2016). Data analysis was performed using MAXQDA software (Version 12). Further quantitative analyses were performed using IBM SPSS Statistics (Version 20).

Trustworthiness of Qualitative Data

Numerous frameworks have been developed to evaluate the rigor or assess the trustworthiness of qualitative data (Patton, 1983), and various strategies for determining credibility, transferability, dependability, and confirmability have been established. In this study, the credibility of the qualitative findings was ensured by using member check and immersion techniques, as well as our ongoing engagement with the data and participation in similar congresses. Then, to complete the data and examine the transferability of our findings, we asked peers who had experience conducting qualitative research to review the initial interviews, coding, and categories. We focused on the research topic and also controlled and checked the findings to increase the reliability of the data.

Ethical Considerations

All ethical considerations were observed in conducting this research. Professors participating gave their informed consent after being told of the objectives of the research, its voluntary nature, our commitment to confidentiality of information, and of their right to withdraw at any time. The university's code of ethics number assigned to this research is IR.SBMUS.REC.1400.1214.

Results

Results of the Qualitative Section

Based on the analysis of interview data and open coding results in the qualitative part of the research, five main categories and 26 subcategories of indicators related to peer observation of teaching in a LMS environment were identified. These are shown in Table 1.

Table 1

Main and Sub Categories Affecting Peer Observation in LMS-Based Classes in Medical Sciences

Category	Subcategory
Content preparation	Provide up-to-date scientific content
	Proportion content to meet the needs of learners
	The fit of content with the course goals
	Observe professional principles for educational design
	Proportion of content volume based on the course unit
	Use of new technologies
	Provide content appropriate to learning styles
	Quality course content (technically)
Content delivery	Provide content at the right time
	Time management
	Course management
	Management of contradictions and conflicts in electronic debates
Effective interactions	Provide appropriate feedback
	Provide timely feedback
	Control and supervision of learners
	Encourage interaction between learners
	Create attractive discussions
	Observe the appropriate period for completing homework
Motivation management	Follow up on the reason for students' non-participation
	Encourage and encouragement to provide creative assignments
	Comparison of students' assignments and introduction to the top student
	Create an environment for the free expression of opinions
	Guidance and encouragement for group work
Supportive services	In-person appointment guide

Follow-up development of students
Debugging classes

The conceptual model of the qualitative part of this study is shown in Figure 1. Five general categories affect the main focus of the research, namely, effective indicators in observing peers: content preparation, content delivery, effective interaction, motivation management, and supportive services.

Figure 1

Conceptual Model of the Qualitative Part of the Research



Results of the Quantitative Section

Face and Content Validity

The results of qualitative face validity measurement showed that five items needed to be corrected and applied to the checklist. Quantitative face validity measurement on the 26 subcategories showed that all items had a score above 1.5 and were suitable for content validity testing.

In the qualitative part of content validity, the checklist was revised and modified based on the opinions of 10 professors participating in this part of the study. Based on quantitative content validity results, according to 15 participating experts, nine items were deleted due to not receiving an appropriate content validity index, and finally, 17 items remained (Table 2).

Table 2

Initial CVR and CVI Values of Peer Review Checklist Questions

Row	Item	Content validity	
		CVR	CVI
1	The scientific content is up to date.	1	1
2	The content presented is relevant to the objectives of the training.	1	1
3	Professional principles of educational design are observed.	0.73	1
4	The volume of content fits the course unit.	0.86	1
5	New technologies are used to deliver content.	0.73	1
6	Feedback is given appropriately.	0.73	1
7	Learners are monitored during the training process.	0.73	1
8	By creating forums, the interaction between learners is created.	0.6	1
9	Appropriate discussions have been organized.	0.6	1
10	Content is provided at the right time	0.6	1
11	The appropriate period for completing homework is observed.	0.6	1
12	The assignments presented are tailored to the needs of learners.	1	1
13	Meeting time and consultation are provided.	0.73	0.86
14	Class time is well managed.	0.6	1
15	The course is well managed.	0.6	1
16	Contradictions and conflicts in online discussions are well managed.	0.6	1
17	Feedback is given at the appropriate time.	0.86	1
Total		0.74	0.99

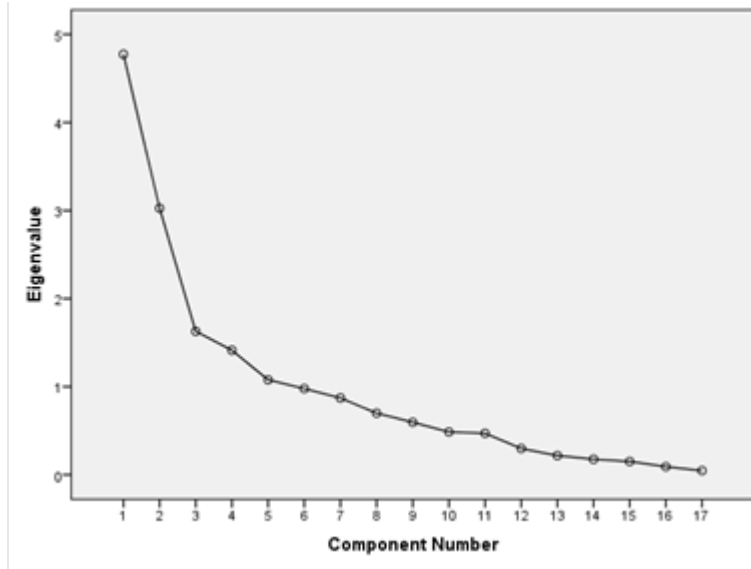
Construct Validity

The possibility of factor analysis on the research sample was investigated using the Bartlett test and the KMO sampling adequacy index where $KMO = 0.61$ and the approximate chi-square = $187/32$, $p = 0.000$, and $df = 136$.

In the study of item commonality, it was found that all items had more than 0.5 subscriptions. Factors in the test were extracted by principal component analysis and varimax rotation. In the present model, five components with eigenvalues higher than 1 and scree plot diagrams were obtained (Figure 2).

Figure 2

Pebble Test (Scree Test) on Peer Evaluation Checklist Factors



The five extracted factors with eigenvalues higher than 1 in total explained 70.55% of the total variance of the test variables. The eigenvalues values of the 5 factors extracted after rotation were 3.92, 3.04, 1.64, 1.56, and 1.11, respectively, each of which was 24.51%, 19.04%, 10.27%, 9.76%, and 6.95% of the variance explained respectively (Table 3).

Table 3

Primary and Extractive Exploration of Exploratory Agent Analysis of the Peer Review Checklist

	Factors extracted by rotation			Extracted agents without rotation			Special value		
	Compression variance %	Explanation variance %	Total	Compression variance %	Explanation variance %	Total	Compression variance %	Explanation variance %	Total
1	24.51	24.51	3.92	26.32	26.32	4.21	26.32	26.32	4.21
2	43.56	19.04	3.04	45.13	18.81	3.01	45.13	18.81	3.01
3	53.83	10.27	1.64	55.07	9.93	1.58	55.07	9.93	1.58
4	63.6	9.76	1.56	63.82	8.74	1.40	63.82	8.74	1.40
5	70.55	6.95	1.11	70.55	6.73	1.07	70.55	6.73	1.07

Based on factor analysis with varimax rotation, all questions with a factor load of at least 0.5 were examined (Yong & Pearce, 2013), and finally, a 17-item checklist was extracted in the form of five factors. The factors were: (a) content management (five items), (b) classroom management (five items), (c) conflict management (two items), (d) assignment management (2 items), and (e) feedback management (3 items). These are shown in Table 4 along with the results of factor analysis.

Table 4

Rotated Factor Matrix by Principal Component Analysis and Varimax Rotation After Exploratory Factor Analysis

Item	Factor bar				
	1	2	3	4	5
Factor 1: Content management					
The scientific content is up to date.	0.910				
The content presented is relevant to the objectives of the training.	0.860				
Professional principles of educational design are observed.	0.909				
The volume of content fits the course unit.	0.769				
New technologies are used to deliver content.	0.834				
Factor 2: Classroom management					
Feedback is given appropriately.		0.558			
Learners are monitored during the training process.		0.870			
By creating forums, the interaction between learners is created.		0.911			
Appropriate discussions have been organized.		0.639			
Content is provided at the right time.		0.532			
Factor 3: Conflict management					
The appropriate period for completing homework is observed.			0.835		
The assignments presented are tailored to the needs of the learners.			0.783		
Factor 4: Assignment management					
Meeting time and consultation are provided.				0.827	
Class time is well managed.				0.845	
Factor 5: Feedback management					
The course is well managed.					0.837
Contradictions and conflicts in online discussions are well managed.					0.861
Feedback is given at the appropriate time.					0.791

Reliability

The result of the intraclass correlation coefficient (ICC) for the checklist was 0.88 which shows acceptable reliability.

Discussion and Conclusion

In this study, a Peer Observation Tool (POT) to be used in LMS-based classrooms was designed, comprised of a list of items related to the homogeneous observation of five main categories: content preparation, content presentation, effective interactions, motivation management, and support services. Furthermore, the results of face validity, content validity, and design tool reliability show that the tool has appropriate validity and reliability for peer observation.

Continuous evaluation of teaching plays an important role in improving the quality of teachers. How the evaluation is performed and the criteria measured are very important. According to Keig (2000), teaching should be seen as a process and follow a path similar to what a research manuscript goes through before being published in a reputable scientific journal, which includes a review and strict judgments by peers.

Peer review, according to Min (2006), is still unknown in e-learning. With the new technological developments in the field of education over the last two decades, these components should be reviewed. Assessing quality in an e-learning system requires attention to the criteria of teaching in general and the field of e-learning in particular. On the other hand, many criteria of the face-to-face classroom must be transferable to the virtual learning space to be examined. The results of this study show that, from the perspective of peers, the items “electronic content enrichment,” “interaction promotion,” “appropriate timing of course delivery,” “content assurance,” “face-to-face interaction,” and “process maturity teaching” are of great importance.

The work done in the development and distribution of multimedia content has raised the hope that students will have access to a wider range of content (Garrison, 2016). New technologies have provided many possibilities for professors to produce attractive and rich content (Collis & Moonen, 2012). As content moves from static and inactive to multimedia, the volume of cognitive processing of memory is reduced and learning is facilitated (Garrison, 2016).

Another important point that was obtained in the research is “promoting teacher-student interaction.” Roslin et al. also showed that for interaction to occur at a high level, effective teaching must be participatory and emphasize teamwork (Amira & Jelas, 2010). Many educators are not aware of the importance and effective methods of live or virtual interactions with learners, and teachers need to be trained to design and implement appropriate interactions (Ibrahimzadeh, Zandi, Alipour, Zare, & Yazdani, 2010).

Another important feature important in evaluating an e-learning system is whether lessons and assignments are uploaded by the instructor following an appropriate schedule. One of the main concerns in this area is the production and management of educational content (Snyder, 2009). A study titled *Academic Quality Assessment* showed two important criteria of a good professor: ability in scientific reasoning and knowledge of how to teach to convey understanding of concepts (Clipa, 2011). Other research

has shown that the specialty of a professor is another factor in student satisfaction. Educational equipment and facilities, as well as intimate teacher-student interaction and cooperation, are equally important factors in evaluation (Butt & Ur Rehman, 2010). The results of another research study have shown that emotional factors and having a correct and appropriate social relationship play an important role in education (Opre, Calbaza-Ormenisan, & Opre, 2011).

Lee has stated that although the goal of the e-learning method is self-learning, feedback plays a major role (2009). On the other hand, for producing quality electronic content, one of the important points to pay attention to is learning styles, and the cognitive and emotional preferences of learners (Kay & Knaack, 2008). E-learning, with all its benefits, is defective due to a lack of direct social interaction and face-to-face contact and the absence of non-verbal cues (Al-Qahtani & Higgins, 2013).

Research has shown that e-learning can be very useful when combined with face-to-face training. In blended learning, the learner benefits from the combination of e-learning and face-to-face learning (Akkoyunlu & Soyulu, 2006). In the present study, the emphasis on creating a face-to-face block during the term was confirmation of these past findings.

The criteria described in this research, in addition to being useful in the evaluation of professors in the field of e-learning, may also empower professors in this field. The empowerment of faculty members, especially in virtual education, will help achieve the mission and goals of higher education institutions and improve performance in this field.

Limitations

In this study, researchers have tried to simultaneously design a valid tool for peer observation in virtual classes as well as evaluate the validity and reliability of that tool, so that the reader will be aware of and able to themselves evaluate the quality of the designed tool. The design of this tool was based on a psychometric process, using the opinions of the target group and specialists and experts. This, being the first time such a research path was taken, is one of the positive points of this tool. However, since the validity and reliability of this tool have only been performed by medical professors, such tests would need to be undertaken in different populations.

In this study, researchers provided complete and accurate information on how to determine the validity and reliability of the designed tool, which has contributed to the clarity of the issues in this field. On the other hand, it should be noted that in medical science education, e-learning is blended learning and not just virtual and LMS-based education. This may affect results in other disciplines. Evaluation of virtual classes by professors' peers can clarify the status quo, and the results can be used in e-learning promotion plans. To confirm or reject the components obtained in this study, it is suggested that these indicators be tested in future research both before and after the empowerment of professors in virtual learning.

Acknowledgment

We would like to thank the esteemed Vice-Chancellor of the virtual college and all the professors who, with their compassionate support, helped us hold peer review sessions.

Conflict of Interest

There was no conflict of interest.

Ethical Affirmation

The present study is part of an educational process, and the meetings were conducted with the coordination and approval of relevant authorities. The participation of faculty members in this research was voluntary.

Funds

No financial support has been received for this study.

References

- Akkoyunlu, B., & Soyulu, M. Y. (2006). A study on students' views on blended learning environment. *Turkish Online Journal of Distance Education*, 7(3), 43–56.
<https://dergipark.org.tr/en/pub/tojde/issue/16925/176657>
- Al-Qahtani, A. A. Y., & Higgins, S. E. (2013). Effects of traditional, blended, and e-learning on students' achievement in higher education. *Journal of Computer Assisted Learning*, 29(3), 220–234.
<http://dx.doi.org/10.1111/j.1365-2729.2012.00490.x>
- Amira, R., & Jelas, Z. M. (2010). Teaching and learning styles in higher education institutions: Do they match? *Procedia-Social and Behavioral Sciences*, 7, 680–684.
<https://doi.org/10.1016/j.sbspro.2010.10.092>
- Andone, I., & Sireteanu, N.- A. (2009). Strategies for technology-based learning in higher education. *The FedUni Journal of Higher Education*, 4(1), 31–42. <https://ssrn.com/abstract=1337962>
- Bates, A. W. (2022). *Teaching in a digital age: Guidelines for designing teaching and learning* (3rd ed.). Tony Bates Associates Ltd. <https://opentextbc.ca/teachinginadigitalage/>
- Benner, P. (2012). Educating nurses: A call for radical transformation— How far have we come? *Journal of Nursing Education*, 51(4), 183–184. <https://doi.org/10.3928/01484834-20120402-01>
- Bernstein, D., Burnett, A. N., Goodburn, A. M., & Savory, P. (2006). *Making teaching and learning visible: Course portfolios and the peer review of teaching* (1st ed.). Jossey-Bass.
- Butt, B. Z., & Ur Rehman, K. (2010). A study examining the students' satisfaction in higher education. *Procedia-Social and Behavioral Sciences*, 2(2), 5446–5450.
<https://doi.org/10.1016/j.sbspro.2010.03.888>
- Clipa, O. (2011). The profile of the academic assessor. *Procedia-Social and Behavioral Sciences*, 12, 200–204. <https://doi.org/10.1016/j.sbspro.2011.02.027>
- Collis, B., & Moonen, J. (2012). *Flexible learning in a digital world: Experiences and expectations* (2nd ed.). Routledge.
- Cunningham, I., Johnson, I., & Lynch, C. (2017). Implementing peer review of teaching: A guide for dental educators. *British Dental Journal*, 222(7), 535–540.
<https://doi.org/10.1038/sj.bdj.2017.316>
- Dill, D. (2007). Quality assurance in higher education: Practices and issues. In P. P. Peterson, E. Baker, & B. McGaw (Eds.), *International Encyclopedia of Education* (3rd ed., pp. 377–383). Elsevier.

- Donnelly, R. (2017). Blended problem-based learning in higher education: The intersection of social learning and technology. *Psychosociological Issues in Human Resource Management*, 5(2), 25–50. <https://doi.org/10.22381/PIHRM5220172>
- Drisko, J. W., & Maschi, T. (2016). *Content analysis: Pocket guides to social work research methods*. Oxford University Press.
- Fernandez, C. E., & Yu, J. (2007). Peer review of teaching. *The Journal of Chiropractic Education*, 21(2), 154–161. <https://doi.org/10.7899/1042-5055-21.2.154>
- Fletcher, J. A. (2018). Peer observation of teaching: A practical tool in higher education. *The Journal of Faculty Development*, 32(1), 51–64. <http://doi.10.13140/RG.2.2.19455.82084>
- Garrison, D. R. (2016). *E-learning in the 21st century: A community of inquiry framework for research and practice* (3rd ed.). Routledge.
- Gehring, E. F., Chinn, D. D., Pérez-Quiñones, M. A., & Ardis, M. A. (2005). Using peer review in teaching computing. *ACM SIGCSE Bulletin*, 37(1), 321–322. <https://dl.acm.org/doi/10.1145/1047124.1047455>
- Henrie, C. R., Bodily, R., Manwaring, K. C., & Graham, C. R. (2015). Exploring intensive longitudinal measures of student engagement in blended learning. *The International Review of Research in Open and Distributed Learning*, 16(3), 131–155. <https://doi.org/10.19173/irrodl.v16i3.2015>
- Ibrahimzadeh, I., Zandi, B., Alipour, A., Zare, H., & Yazdani, F. (2010). The kinds of e-learning and different forms of interaction on it. *Interdisciplinary Journal of Virtual Learning in Medical Sciences*, 1(1), 11–22. https://ijvlms.sums.ac.ir/article_46026.html
- Johnston, A. L., Baik, C., & Chester, A. (2020). Peer review of teaching in Australian higher education: A systematic review. *Higher Education Research & Development*, 41(2), 390–404. <https://doi.org/10.1080/07294360.2020.1845124>
- Kay, R. H., & Knaack, L. (2008). An examination of the impact of learning objects in secondary school. *Journal of Computer Assisted Learning*, 24(6), 447–461. <https://doi.org/10.1111/j.1365-2729.2008.00278.x>
- Keig, L. (2000). Formative peer review of teaching: Attitudes of faculty at liberal arts colleges toward colleague assessment. *Journal of Personnel Evaluation in Education*, 14(1), 67–87. <https://link.springer.com/article/10.1023/A:1008194230542#citeas>
- Keshavarz, M., & Ghoneim, A. (2021). Preparing educators to teach in a digital age. *The International Review of Research in Open and Distributed Learning*, 22(1), 221–242. <https://doi.org/10.19173/irrodl.v22i1.4910>

- Keshavarz, M., Mirmoghtadaie, Z., and Nayyeri, S. (2022). Design and validation of the virtual classroom management questionnaire. *The International Review of Research in Open and Distributed Learning*, 23(2), 121–135. <https://doi.org/10.19173/irrodl.v23i2.5774>
- Kintu, M. J., Zhu, C., & Kagambe, E. (2017). Blended learning effectiveness: The relationship between student characteristics, design features, and outcomes. *International Journal of Educational Technology in Higher Education*, 14(1), 1–20. <https://doi.10.1186/s41239-017-0043-4>
- Lacasse, Y., Godbout, C., & Sériès, F. (2002). Health-related quality of life in obstructive sleep apnoea. *European Respiratory Journal*, 19(3), 499–503. <https://doi.org/10.1183/09031936.02.00216902>
- Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel Psychology*, 28(4), 563–575. <https://doi.org/10.1111/j.1744-6570.1975.Tb01393.x>
- Lee, J.-K. (2009). The effects of self-regulated learning strategies and system satisfaction regarding learner's performance in e-learning environment. *Journal of Instructional Pedagogies*, 1, 30–45. <https://aabri.com/manuscripts/08053.pdf>
- Min, H.-T. (2006). The effects of trained peer review on EFL students' revision types and writing quality. *Journal of Second Language Writing*, 15(2), 118–141. <https://doi.org/10.1016/j.jslw.2006.01.003>
- Missildine, K., Fountain, R., Summers, L., & Gosselin, K. (2013). Flipping the classroom to improve student performance and satisfaction. *Journal of Nursing Education*, 52(10), 597–599. <https://doi.org/10.3928/01484834-20130919-03>
- Mohammadbeigi, A., Mohammadsalehi, N., & Aligol, M. (2015). Validity and reliability of the instruments and types of measurements in health applied research. *Journal of Rafsanjan University of Medical Sciences*, 13(12), 1153–1170. <http://journal.rums.ac.ir/article-1-2274-en.html>
- Neuendorf, K. (2017). *The content analysis guidebook*. SAGE. <https://dx.doi.org/10.4135/9781071802878>
- OECD (2021) *The state of school education: One year into the Covid pandemic* Paris, France: OECD
- Opre, D., Calbaza-Ormenisan, M., & Opre, A. (2011). University teaching: Didactic expertise reflected by metacognitive abilities and emotional control. *Procedia-Social and Behavioral Sciences*, 29, 670–677. <https://doi.org/10.1016/j.sbspro.2011.11.291>
- Patton, M. Q. (1983). Effective evaluation: Improving the usefulness of evaluation results through responsive and naturalistic approaches [Review of the book *Effective evaluation: Improving the usefulness of evaluation results through responsive and naturalistic approaches* by E. G. Guba & Y. S. Lincoln]. *The Journal of Higher Education*, 54(3), 339–342. <https://doi.org/10.1080/00221546.1983.11778201>

- Pinto-Llorente, A. M., Sánchez-Gómez, M. C., García-Peñalvo, F. J., & Casillas-Martín, S. (2017). Students' perceptions and attitudes towards asynchronous technological tools in blended-learning training to improve grammatical competence in English as a second language. *Computers in Human Behavior*, 72, 632–643. <https://doi.org/10.1016/j.chb.2016.05.071>
- Polit, D. F., Beck, C. T., & Owen, S. V. (2007). Is the CVI an acceptable indicator of content validity? Appraisal and recommendations. *Research in Nursing & Health*, 30(4), 459–467. <https://doi.10.1002/nur.20199>
- Ruiz, J. G., Candler, C., & Teasdale, T. A. (2007). Peer reviewing e-learning: Opportunities, challenges, and solutions. *Academic Medicine*, 82(5), 503–507. <https://doi.10.1097/acm.0b013e31803ead94>
- Snyder, M. M. (2009). Instructional-design theory to guide the creation of online learning communities for adults. *TechTrends*, 53(1), 45–57. <https://doi.org/10.1007/s11528-009-0237-2>
- Speer, S. (2010). Peer evaluation and its blurred boundaries: Results from a meta-evaluation in initial vocational education and training. *Evaluation*, 16(4), 413–430. <https://doi.org/10.1177/1356389010382265>
- Stevens, J. P. (2012). *Applied multivariate statistics for the social sciences* (5th ed.). Routledge.
- Twomey, A. (2004). Web-based teaching in nursing: Lessons from the literature. *Nurse Education Today*, 24(6), 452–458. <https://doi.org/10.1016/j.nedt.2004.04.010>
- Waltz C. F., & Bausell R. B. (1981). *Nursing Research: Design statistics and computer analysis*. F. A. Davis.
- Watkins, M. W. (2018). Exploratory Factor Analysis: A Guide to Best Practice. *Journal of Black Psychology*, 44(3), 219–246. <https://doi.org/10.1177/0095798418771807>
- Westen, D., & Rosenthal, R. (2003). Quantifying construct validity: two simple measures. *Journal of Personality and Social Psychology*, 84(3), 608–618. <https://doi.org/10.1037/0022-3514.84.3.608>
- Wood, D. F. (2003). ABC of learning and teaching in medicine. Problem-based learning. *BMJ*, 326, 328–330. <https://doi.org/10.1136/bmj.326.7384.328>
- Yong, A. G., & Pearce, P. (2013). A beginner's guide to factor analysis: Focusing on exploratory factor analysis. *The Quantitative Methods for Psychology*, 9(2), 79–94. <https://doi.org/10.20982/tqmp.09.2.p079>

