

Improving and assessing the quality and effectiveness of teaching by innovative peer review approach

**Recent efforts at the Budapest University of Technology and Economics for the
modernisation and quality improvement of education**

Anikó Kálmán¹
Associate Professor
E-mail: drkalmananiko@gmail.com

Zsuzsanna Eszter Tóth
Vice-Dean for Education Quality
E-mail: toth@mvt.bme.hu

György Andor
Dean
E-mail: andor@finance.bme.hu

Budapest University of Technology and Economics
Faculty of Economics and Social Sciences
Budapest, Hungary

Conference Key Areas: Lifelong Learning, University Staff Development, Engineering
Education Research

Keywords: peer-review, staff development, innovation assessment, complex learning
environment

INTRODUCTION

The Hungarian higher education system has been undergoing significant transformations in the last 15 years. Several new institutions have been established and higher education became a mass-market service characterized by increasing number of students and increasing number and diversity of institutions. Universities meanwhile experienced dramatic changes in funding and in student numbers. These changes necessitated addressing the issue of quality with increasing attention.

At the Budapest University of Technology and Economics (BME), Faculty of Economics and Social Sciences, the branch of studies in the field of management and business administration lost 95% of state financed places in 2012, therefore quality issues have become extremely important. The competitive pressures, the growing number of institutions and the increasing costs together with demographic shifts force institutions to put greater emphasis on student satisfaction [1]. Students are now generally recognized as the principal stakeholders of higher education.

¹ Corresponding Author: Anikó Kálmán, drkalmananiko@gmail.com

1. LITERATURE REVIEW

1.1. Peer review of teaching

The issue of peer vs. student observation has been discussed in the literature. Ward Griffin and Brown [2], highlighted that peers are a more reliable source of information on the lecturer's knowledge of content. Leamon and Fields [3] emphasized that peers are better able to evaluate a lecturer's knowledge of content, selection of instructional materials that optimally supports the overall course objectives, or a lecturer's pedagogical development year by year.

While it is broadly acknowledged that peer observation of teaching is of high importance, nonetheless it can be provoking as peers are both colleagues and competitors [4]. Costello et al. [5] and Branningen and Burson [6] emphasize the subjectivity of peer reviewing. Lecturers perceive peer review positively when good relations exist between reviewer and reviewee. Quinlan [7] reports that lecturers who see themselves as good teachers routinely compare their peers' teaching strategies to their own practice or experience.

Most results support the contention that "colleagues do enter a peer observation with preconceived notions of what constitutes good teaching and such notions are generally constructed around the observer's own teaching approach" and when faculty perceive a colleague's approach to teaching is different from their own, there is a trend towards lower ratings of effectiveness.

In higher education context, Washer [8] concludes that "at its best, the process can foster and disseminate best practice and lead to closer academic links and more general team-building within and without academic departments."

With the transformation of teaching and learning methods by using more and more information and communication technology tools and integrating the approach of lifelong learning at the universities, further challenges emerge. As Kálmán [9] highlights, "the teachers' role and the institutional challenges in self directed learning raise situational, institutional and dispositional obstacles which have to be treated in the process of quality management of learning."

Perlman and McCann [10] present a particular 'handbook' for the peer review of teaching. Bingham and Ottewill [11] report a pilot project at Sheffield Hallam University where they try to create a balanced picture between student feedbacks and professional judgments of academic staff.

A prior study by Hodgkinson [12] finds out that there is value in promoting peer observation of teaching models in higher education institutions. "It must also, to be successful, be supported by senior management within the institution."

Examining the relevant literature a number of quality criteria are mentioned when designing a peer review process [13][14][15]:

- authenticity: assessment tasks need to include the relevant knowledge, skills, and attitudes.
- transparency: assessment needs to be clear and comprehensible to all participants involved
- fairness: this implies the absence of bias against certain groups and the exclusion of irrelevant variance in the assessment process.
- generalizability: it includes three specific components, namely comparability, reproducibility and transferability.
- educational consequences: it can be identified as an overall criterion equated with the consequential validity of an assessment.

In the assessment process the following phases can be identified [16]: (1) Determining the assessment goals, (2) Selecting assessment tasks (composition of assessment assignments covering content to be mastered), (3) Establishing scoring rules, (4) Administrating or implementing the assessment (execution of the actual assessment), (5) Appraising (the actual scoring), (6) Interpreting and valuing scores, (7) Giving feedback

2. METHODOLOGY - PEER REVIEW OF TEACHING AT FACULTY OF ECONOMICS AND SOCIAL SCIENCES

Within BME, student feedback has been a well-established feature of course evaluation and review. The “peer support review” is an essential process for reviewing teaching processes, ideas and “catching mistakes” and so improve the quality of the teaching service. The peer review program is completed by both student evaluations and lecturer self-evaluations.

The experiment reported here involved members from every department of the Faculty. At first we have launched the peer review program at the compulsory courses of our BA and BSc educational programs with the highest number of students enrolled. From 28 that kind of courses we chose 10 in the case of which the peer review program has been started. In the fall semester 2015 16 lecturers of 11 courses was observed by 35 observers.

Table 1. Parameters of courses involved in the peer review program

Title of the course	Credit -point	Number of students	Observed lecturers	Number of observers	Number of observed lectures	Number of evaluations	Number of student evaluations
Business Economics	5	652	1	7	10	22	166
Applied Statistics	5	468	2	7+4	9+9	19+13	206
Marketing	5	713	1	8	16	21	248
History of economy	3	272	1	8	13	29	101
Macro-economic finance	3	130	2	5+4	5+5	14+16	57
Accounting	5	450	2	6+5	8+10	20+13	135
Environmental economics	3	427	1	7	10	27	104
Economics	5	463	2	6+6	13+11	21+23	51
Psychology	3	197	1	6	7	16	59
Social communication	2	57	1	8	11	21	28
Mathematics	6	534	2	4+5	7+8	10+19	71
Total	-	4363	16	-	152	304	1226

The peer review program launched in the fall semester 2015 involved 4363 students, 152 observed lectures, 304 peer review evaluations based on observing lectures and 1226 student evaluations regarding the evaluation of student performance.

Prior to the implementation of peer support review, the issues to be resolved were: the focus of the review processes, the preparing participants, the timing and scheduling of the review process, the conduct of review meetings and the nature of outcomes.

The peer review process in helicopter view included: a *Planning stage* – preparing the questionnaires, planning the steps of peer review process, selecting the courses and teachers to be observed, heads of departments delegate colleagues for being observers, pairing observed teachers with observers, training and informing of participants, documentation issues, an *Observation stage* – pre-meeting of participants before the semester starts, observing classroom performance and additional elements of teaching performance (communication with students, consultations, midterm tests, exams etc. above during the semester). An *Initial feedback stage* – applying the prepared forms on the generic issues for the different observations, when the observed persons receive written feedback that they can discuss at the closing meeting at the end of the semester. The *Final feedback stage* is based on the evaluations given during the semester and on the self-evaluation outcomes relevant issues are discussed, strengths and improvement opportunities are identified both for the lecturer and for the course itself. In the *Dissemination stage* the vice dean responsible for the program informs various committees. The peer review results are triangulated with student feedbacks. The outcomes of the peer review process are used as a basis for the dissemination of effective practice across faculties and within faculties. Finally the *Action* – when the committees take the necessary actions, including initiate training or awareness training as necessary.

3. IMPROVING THE REVIEW PROCESS

We have found areas in the process which may need to be refined with a wider participation level. *Cultural aspects* are of greater consequence than procedural. The lack of well-established self-reflective practice stems from cultural aspects. e.g. the benefits that can origin from cross-department collaborations. We should establish a culture where lecturers' willingness to engage in an evaluative exercise on a voluntary basis is high.

The primary aim is to bring about changes in teaching practice and introduce new teaching methods. It can be reassuring for the lecturers that they are doing a reasonable job and can identify areas for improvement. It is also important to ensure the compliance to generally accepted standards and to facilitate management decision making in such areas as establishing and improving teaching standards at faculty level, and can be utilized as inputs for promotion, awards etc.

From a *Procedural* point of view, the success of peer support review largely depends on the professional background and foundation. The reviewers must prepare for the review meetings by reading the necessary documentations, participating fully and contributing to the set of recommendations. Written guidelines are also needed.

In the beginning this program concentrates on the observations of performance in the lecture theatre which is a different activity compared to seminar groups. Later on seminar observations are to be included and the differences should be reflected in a different set of criteria. A lecturer may perform well in a seminar setting but need considerable support in lecturing to large number of students.

Regular review of the scheme by all participators is welcomed. The experience deriving from the application in the first two weeks were immediately fed back and according to the first comments both the reviewing process and its criteria system were refined. The review process and the criteria system applied in the second semester were improved based on the feedbacks of the first semester.

4. RESULTS

The following figures highlight some of the results derived from the program in fall semester 2015.

In *Fig.1.* the diagram on the left illustrates the average evaluation scores (1-5, 1: worst and 5: best) given for each lecturer (L1-L16) (based on the average scores given by all reviewers in 19 evaluation dimensions: *D1 Communicating learning objectives, D2 Communicating course requirements, D3 Positioning the current subject to the curriculum, D4 Quality of the introduction part, D5 Volume, intonation, D6 Grammar, intelligibility, speech rate, D7 Learning, explaining the terminology, D8 Explanatory capability, D9 Maintaining students' attention, D10 Choosing the proper presentation techniques, D11 Quality of the slide show, D12 Consonance of the slide show with verbal communication, D13 Political correctness, D14 Consonance of the current subject with the course objectives, D15 Consonance of the lecture structure and the applied course materials with the course objectives, D16 Structuring course materials and the slideshow, D17 Consonance of course materials with requirements of the course, D18 Logical structure of the presentation, D19 To what extent the current lecture supports the student preparation for evaluation*). It can be seen that for most of the lecturers the average evaluations do not differ more than a half unit, and some lecturers were more divisive.

The diagram on the right presents the differences in average evaluations in each dimension by taking all lecturers and reviewers into consideration. It highlights those dimensions where general problems can be addressed as the means are lower (e.g. D9, D10, D4), or where there are bigger differences between lecturers' performance (e.g. D9, D16, D6). These dimensions shed a light on general problems.

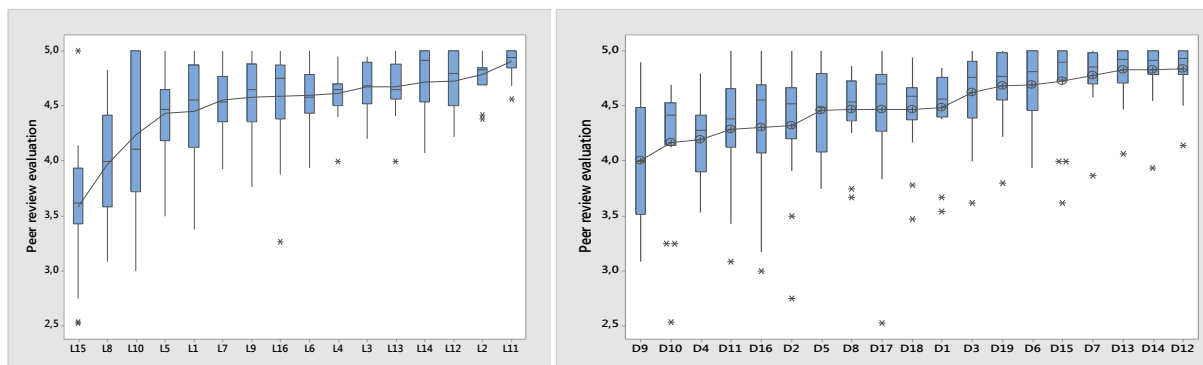


Fig.1. Bloxplot diagrams of peer review evaluations by lecturers (left), and by evaluation dimensions (right)

Fig.2. maps the stochastic relationship between the average value of peer review evaluation and the aggregate students' evaluation results for the lecturers. This shows strong correlation ($R^2=86.6\%$).

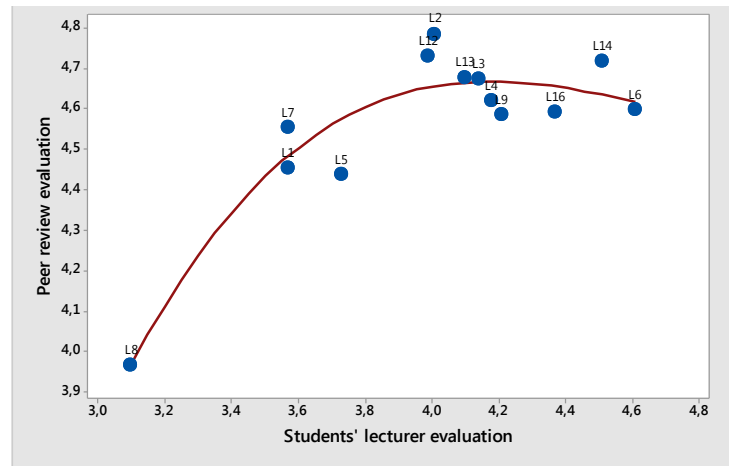


Fig.2. Relationship between peer review and student evaluations

The first investigations focused on measuring correlation between evaluation aspects and student course evaluation results. The following figures (Fig. 3) highlight those dimensions where at least moderate correlations were found.

Table 2. Correlation between specific evaluation aspects (D19, D10, D11, D12) and students' lecturer evaluation (OHV-OMI)

Variables	Pearson correlation	p-value
D19 – OHV-OMI	0,856	0,000S
D10 – OHV-OMI	0,814	0,001
D11 – OHV-OMI	0,609	0,027
D12 – OHV-OMI	0,585	0,036

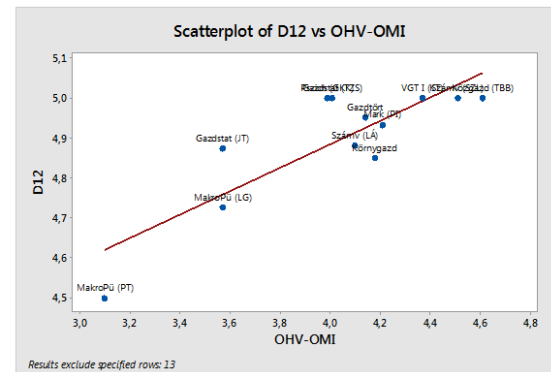
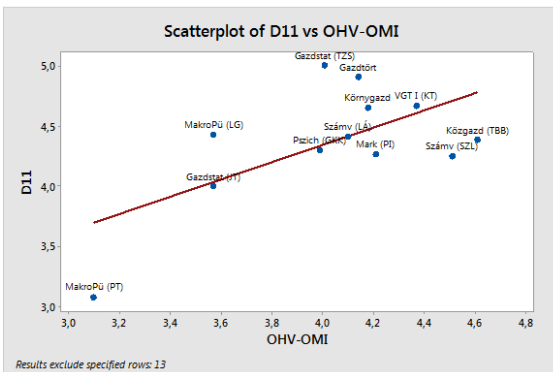
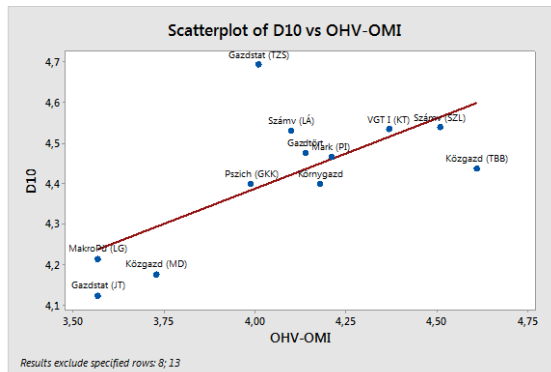
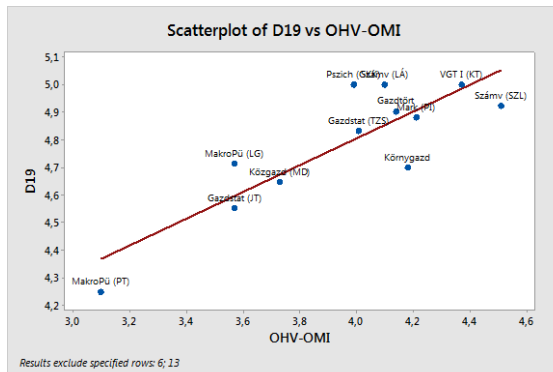


Fig.3. D19, D10, D11 and D12 evaluation aspects – Students' lecturer evaluation

Students gave immediate feedbacks after student performance evaluations by expressing their judgement on 1-5 scale in 8 aspects (*HSZ1 availability of instructional materials, HSZ2 midterm test, exam circumstances, HSZ3 – review the course of tests, exams, HSZ4 clarity of exam questions, HSZ5 consonance of exam questions with requirements, HSZ6 clarity of result calculation, HSZ7 standard of consultation opportunities, HSZ8 standard of midterm test/ exam viewing opportunities*). The following figure (Fig. 4.) shows the correlation between average student ratings and the aggregate students' course evaluations.

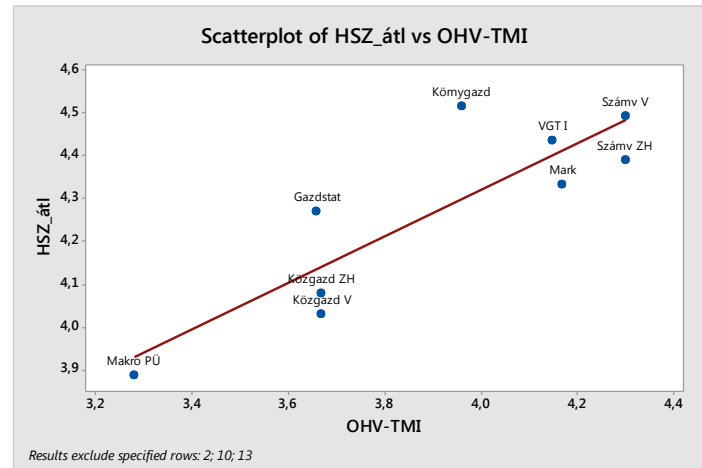


Fig.4. Aggregate students' course evaluations – average student ratings concerning test and exams ($R^2=0,864$, $p=0,003$)

Observers were also required to give feedbacks regarding student performance evaluations by expressing their judgement on 1-5 scale in 4 aspects (*OSZ1 – review the course of tests, exams OSZ2 midterm test, exam circumstances, OSZ3efforts minimizing the use of cheat sheets, OSZ4 structuring the worksheet*). The results of a paired t-test are presented in Table 3. including the comparison of average students' and observers' ratings in the same aspect.

These tests show that there is no difference between average students' and observers' ratings.

Table 3. Paired t-tests

Paired T for OSZ1 - HSZ3					Paired T for OSZ2 - HSZ				
N	Mean	StDev	SE	Mean	N	Mean	StDev	SE	Mean
OSZ1	21	4,5319	0,3116	0,0680	OSZ2	21	4,6043	0,4293	0,0937
HSZ3_1	21	4,5895	0,1715	0,0374	HSZ2_1	21	4,5758	0,1745	0,0381
Difference	21	-0,0576	0,3780	0,0825	Difference	21	0,028	0,494	0,108
95% CI for mean difference: (-0,2297; 0,1144)					95% CI for mean difference: (-0,196; 0,253)				
T-Test of mean difference = 0 (vs \neq 0): T-Value = -0,70					T-Test of mean difference = 0 (vs \neq 0): T-Value = 0,26				
P-Value = 0,493					P-Value = 0,795				

5. CONCLUSIONS

The primary aim of the present peer review was to stimulate debate and experimentation and revitalize the contribution of reviewers to course evaluation and review and to provide clear public evidence of any improvements in practice.

A “culture of peer reviewing” is a critical factor in order to enhance quality improvements in teaching, for searching best practices, ideas, tips and identifying common mistakes, bringing about changes in teaching practice and introducing new teaching methods both on individual and faculty level.

An important conclusion may be that most identified mistakes are not connected with classroom teaching activities but with such other supplementary elements like structuring course outlines, course materials, or the organization of midterm tests and exams.

The program got people to talk about what they are doing and how they can achieve that. The program contributed significantly to a professional community building and to motivate more interactions between the different institutions of the faculty.

The outcome of this program would suggest that there is value in promoting such a model in institutions of higher education. Several (at least 5) observers attending lectures during the semester ensures that objectivity is not endangered. A well thought out training program needs to be devised and provided for all who participate in the program.

A thorough revision of the criteria system based on the application in the first two semesters is ahead.

BIBLIOGRAPHY

Abedin, N.F.Z., Taib, J.M., Jamil, H.M.T. “Comparative Study on Course Evaluation Process: Students’ and Lecturers’ Perceptions”, *Procedia - Social and Behavioral Sciences* 123 (2014) pp. 380–388

Bingham, R. and Ottewill, R. “Whatever happened to peer review? Revitalising the contribution of tutors to course evaluation”, *Quality Assurance in Education*, Vol. 9 Iss 1 (2001) pp. 32–39

Courneya, C.A., Pratt, D.D., Collins, J. “Through what perspective do we judge the teaching of peers?”, *Teaching and Teacher Education* 24(2008) pp. 69–79

Hodgkinson, M. “Peer Observation of Teaching Performance by Action Enquiry”, *Quality Assurance in Education* Vol.2Iss2 (1994) pp. 26–31

Marlowe, K.F., Wargo, K.A., Kelley, K.W. “Peer review of the curriculum as a continuous process of improvement”, *Currents in Pharmacy Teaching and Learning* 4 (2012) pp157–164

Samson, S., McCrea, D.E. “Using peer review to foster good teaching”, *Reference Services Review*, Vol.36 Iss1 (2008) pp. 61 – 70

Kálmán, A. „Formation of Competences Establishing Lifelong – Learning in the Dimension of Non-Formal and Informal Learning” - In: *Modern Management Research Conference: Insights into the Sustainable Growth of Business*, Vilnius, Lithuania, November 19-21. (CD) (2009)

Kálmán, A. „Learning - in the New Lifelong and Lifewide Perspectives Tampere” - Tampere University of Applied Sciences, 2016. 139 p. (ISBN:978-952-5903-80-5)

REFERENCES

- [1] Kotler, P. and Fox, K. (1995), "Strategic Marketing for Educational Institutions", (2nd ed.), Prentice Hall, New Jersey
- [2] Ward Griffin, C., Brown, B. (1992): Evaluation of teaching: A review of the literature. *Journal of Advanced Nursing*, 17, 1408-1414.
- [3] Leamon, M., Fields, L. (2005): Measuring teaching effectiveness on a pre-clinical multi instructor course. *Teaching and Learning in Medicine*, 17(2): 119-129.
- [4] Ciesla, J., Lovejoy, N. (1997): Peer review: A method for developing faculty leaders. *Nurse Educator*, 22(6): 41-47.
- [5] Costello, J., Pateman, B., Pusey, H., Longshaw, K. (2001): Peer review of classroom teaching: An interim report. *Nurse Education Today*, 21, 444-454.

- [6] Branningen, C. N., Burson, J. (1983): Revamping the peer review process. *Journal of Nursing Education*, 22 (7): 287-289.
- [7] Quinlan, K. M. (2002): Inside the peer review process: How academics review a colleague's teaching portfolio. *Teaching and Teacher Education*, 18, 1035-1049.
- [8] Washer, P. "Designing a system for observation of teaching", *Quality Assurance in Education* (2006) Vol.14 Iss 3. pp. 243–250
- [9] Kálmán, A.(2016): Learning – in the New Lifelong and Lifewide Perspectives Tampere University of Applied Sciences, Tampere, (2016), pp. 130
- [10] Perlman, B., McCann, L.I. "Peer Review of Teaching", University of Wisconsin Oshkosh (1998)
- [11] Bingham, R. and Ottewill, R. "Whatever happened to peer review? Revitalising the contribution of tutors to course evaluation", *Quality Assurance in Education*, Vol. 9 Iss 1 (2001) pp. 32–39
- [12] Hodgkinson, M. "Peer Observation of Teaching Performance by Action Enquiry", *Quality Assurance in Education* Vol.2 Iss2 (1994) pp. 26–31
- [13] Baartman, L. K. J., Bastiaens, T. J., Kirschner, P. A., Van der Vleuten, C. P. M. (2006): The wheel of competency assessment: Presenting quality criteria for competency assessment programmes. *Studies in Educational Evaluation*, 32, 153-170.
- [14] Baartman, L. K. J. (2008): Assessing the assessment: development and use of quality criteria for competence assessment programmes. Doctoral dissertation, Utrecht, The Netherlands: University of Utrecht
- [15] Ryan, G. J., Marshall, L. L., Porter, K., Jia, H. (2007): Peer, professor and self-evaluation of class participation. *Active Learning in Higher Education*, 8, 49-61
- [16] Freeman, R., Lewis, R. (1998): Planning and implementing assessment. London: Kogan Page