



## 12th European Quality Assurance Forum

### Responsible QA – committing to impact

Hosted by the University of Latvia  
Riga, Latvia

23-25 November 2017

### Paper proposal form

**Deadline 24 July 2017**

*Please note that all fields are obligatory. For a detailed description of the submission requirements and Frequently Asked Questions please consult the Call for Contributions.*

#### Author(s)

**Name:** Ana Torres

**Position:** Researcher at Statistics and Prospective Unit

**Organisation:** Instituto Superior Técnico

**Country:** Portugal

**E-mail address:** ana.torres@tecnico.ulisboa.pt

**Short bio (150 words max):** Ana Torres works at the Statistics and Prospective Unit at the Instituto Superior Técnico (University of Lisbon). She received her master degree in Biostatistics from Faculdade de Ciências (University of Lisbon). Her recent publications include “Engaging students in the teaching quality assurance: a driver of higher education excellence” (paper presented at the EAIR 37th Annual Forum in Krems, 2015) and “Working in partnership with students: the importance of experience and engagement” (paper presented at the EAIR 38th Annual Forum in Birmingham, 2016). Her research interests include quality assurance and management in higher education.

**Name:** Filipa David

**Position:** Researcher at Statistics and Prospective Unit

**Organisation:** Instituto Superior Técnico

**Country:** Portugal

**E-mail address:** filipasdavid@tecnico.ulisboa.pt

**Short bio (150 words max):** Filipa David collaborates with the Statistics and Prospective Unit at Instituto Superior Técnico (University of Lisbon) since 2012. She has a master degree in Statistics and Operational Research from Faculdade de Ciências (University of Lisbon). Since 2015, her main work has been developing case studies (e.g. characterization of the new students' profile), and giving support to the institution's teaching and decision-making processes such as the curricular units quality assurance system.



**Name:** Marta Graça

**Position:** Coordinator at Statistics and Prospective Unit

**Organisation:** Instituto Superior Técnico

**Country:** Portugal

**E-mail address:** marta.graca@tecnico.ulisboa.pt

**Short bio (150 words max):** *Marta Graça is the coordinator of the Statistics and Prospective Unit at the Instituto Superior Técnico (University of Lisbon). She received her bachelor degree in Statistics and Operational Research from Faculdade de Ciências (University of Lisbon). Her recent publications include "Degree programme internal benchmarking" (paper presented at the EAIR 35th Annual Forum in Rotterdam, 2013) and "Working in partnership with students: the importance of experience and engagement" (paper presented at the EAIR 38th Annual Forum in Birmingham, 2016).*

## Proposal

**Title:** Quality assurance of teaching and learning: validity and usefulness of student ratings

**Abstract (150 words max):**

The Quality Assurance System for Course Units at Instituto Superior Técnico monitors the course units taught at the institution, by involving the entire academic community in the teaching and learning assessment process.

Although this mechanism is well established, there are factors which require further study, namely: Are these student ratings stable? What is the relation between student ratings, student grades and subjective assessment of learning?

We find that students' ratings of instruction are stable and much more a function of the perceived quality of teaching than of the student grades, supporting their validity as a quality assurance instrument. Implications of these findings for practical use of student ratings are discussed.

**The paper is based on:** research

**Has this paper previously been published/presented elsewhere?** No.

**Text of paper (3000 words max):**

### Introduction

Students' evaluations of teaching effectiveness have been a topic of great interest in universities all over the world: heated debate concerning the merits and the shortcomings of these evaluations continues to flourish, despite intensive ongoing research and international growth in their use as one indicator of teaching quality (Feldman 2007; Watkins, 1994; Marsh & Roche, 1997).

According to an often cited Marsh's literature review (1987) student ratings are: a) multidimensional; b) reliable and stable; c) primarily a function of the instructor who teaches the course; d) relatively valid against a variety of indicators of effective teaching; e) relatively unaffected by a variety of variables hypothesized as potential biases; and f) seen to be useful by the academic community.

By gathering evidence of teaching effectiveness among their students, the universities are able to make informed and objective decisions about retention, promotion, teaching awards or funding decisions (Bennett & Sid Nair 2010; Chen & Hoshower 2003), has measures available for accreditation requirements (Scoles, Bilgutay, & Good 2000), besides providing information to students for the selection of their courses and teachers (Marsh & Roche 1993). Yet probably the most important advantage of student evaluations is the feedback the forms provide directly to instructors, so that they can refine their courses (content, format and structure) and their teaching practices, in order to offer students better learning experiences (Center for Teaching and Learning - Stanford University 1997).



However, even when the ratings are technically rigorous, one of the major problems is the daily practice: student ratings are often misinterpreted, ignored, or not accompanied by other information that allows users to make conscious decisions. As a result, there is a great deal of suspicion, and hostility towards ratings (Theall 2002). Among the myths that have crept into the academic community as arguments against the use of student ratings is the following: "Students are too immature, capricious, and inexperienced to give reliable feedback on teaching" (Lawall 1998).

But, is this true? In fact, research on the subject indicates the opposite, as Guthrie (1954), Hativa (1996), Albanese (1991), Palchik (1988), Costin, Greenough & Menges (1971), and Hogan (1973) found substantial correlations between student ratings of the quality of their teachers from one year to the next. Additionally, there are consistently high correlations between student ratings of teachers/course units, and student ratings of achievement in a course ('amount learned'), which provides strong support for the validity of student ratings as measures of teaching effectiveness (Cohen 1981, Cohen 1986, Baird 1987). In the words of Hoffman (1979), "Students learn more from better teachers".

Another controversy is centred around the relationship between student ratings and their course grades (Aleamoni 1999). The general feeling is that teachers can 'buy' good student ratings by giving good grades. However, this aspect of student evaluations has been studied extensively and the majority of the studies have reached the same conclusion: there is no consistent correlation between the grades a faculty member gives and the ratings he or she receives (Aleamoni & Hexner 1980; Baird 1987; Gigliotti & Buchtel 1990). Nevertheless, even in those studies that do see a significant effect there is the lack of proof of 'lenient grading' as opposed to good teaching, which might stimulate students to perform well in a course and as a consequence lead them to give the course high ratings (Lawall 1998; Theall et al. 2001; Howard & Maxwell 1980).

Lastly, we have to keep in mind that the students are the only direct observers of a teacher's classroom teaching performance and therefore, are certainly qualified to express their satisfaction or dissatisfaction with their experience (Seldin 1989; Theall 2002). As it follows, students should actively engage in the institutional quality assurance mechanisms. Aware of this essential aspect, and in accordance with the European and Portuguese guidelines for quality assurance, Instituto Superior Técnico (IST) has invested in the development of an internal system (Quality Assurance System for Course Units - QUC) to evaluate the quality of teaching and learning at the curricular unit level, involving all the academic community.

In this mechanism, students are 'main actors' and besides giving information about their learning experience on the courses, they also work in collaboration with the academic staff, on identified problems regarding the teaching process. In this way, IST promotes a culture of ongoing dialogue between the academic community, in cooperation for fostering quality.

However, as student ratings are used as the primary measure of teaching effectiveness at IST, active participation from students, and assuring the validity of the quality assurance mechanism are critical factors. The challenges of QUC include maintaining the academic community engaged in the process by assuring the validity of the ratings, and especially by ensuring its outcomes have visible effects in the enhancement of the learning processes over time.

The factors and challenges outlined above raised a number of questions, among which are the following: Are student ratings stable? What is the relation between student ratings, student grades and subjective assessment of learning?

This paper aims to: present the QUC system and discuss its challenges; give an overview on its results and validity of the ratings; and finally, discuss the implications of these findings for practical use of student ratings in the enhancement of teaching excellence.

### **Quality Assurance System for Course Units at IST: historical overview**

IST was founded in 1911, and aims to contribute to the development of society, promoting and sharing excellence in higher education in the fields of Architecture, Engineering, Science and Technology. IST offers Bachelor, Master and PhD programmes, lifelong training and develops Research, Development and Innovation (RD&I) activities, which are essential to provide an education based on the top



international standards. The faculty has approximately 11.458 students enrolled and 853 teachers and researchers.

In 1993, IST began the development of a teaching activity evaluation mechanism for their undergraduate programmes through the 'Subject evaluation system', always assuming that it would have a positive rebound on the outcomes of their courses and on the performance of their teachers. The process consisted in applying a student paper survey to each functioning course unit, at the end of each semester. This procedure was time consuming, had an underlying relative high number of errors, a low response rate (between 30% and 40%) and the results were not disseminated through the student community.

In 2005, the questionnaires transitioned to a digital format, on an attempt to minimize the amount of errors and logistics involved. However, this transition was reflected on a significant drop of the students' response rate (20% in the first application and later to values even lower), which made it difficult to validate the results. Nonetheless, given the benefits of online surveying mechanism over the paper forms, IST decided to investigate the motives responsible for this deficit.

A preliminary research, on the related literature, revealed some probable reasons why students failed to respond to online evaluations: apathy, technical problems, perceived lack of anonymity, lack of importance, inconvenience, inaccessibility, and time for completion (Avery et al. 2006; Ballantyne 2003; Dommeyer, Baum & Hanna 2002; Sorenson & Reiner 2003). On the other hand, Chen and Hoshower (2003) found that students' motivation to participate in a rating system was based on the following survey outcomes (in order of decreasing importance): 1) improvements in teaching; 2) improvements in course content and format; 3) faculty personnel decisions (promotion, tenure, salary increase).

Keeping these motives in mind, IST identified some factors that needed an overall review to ensure the enquires effectiveness: the involvement of all stakeholders in the process in a clear and effective way, developing retroactive actions towards the results, and disseminating the outcomes to students in a timely fashion manner. With this perspective, in 2007, a set of directives were developed and the actual Quality Assurance System for the IST Curricular Units (QUC) was created, with the purpose of giving 'voice' to students, and improving their learning experience.

### **Design of the QUC system**

QUC system provides for a half-yearly evaluation of each course of the programmes taught at IST, involving different stakeholders in the process, in a clear and responsible manner. This mechanism is not a static model, restricted merely to data collection: it is a dynamic and continuous quality improvement mechanism, with the adjustment of the teaching and learning process, monitoring of its progress, and a cyclical review of the results.

The main sources of information are a Student Survey, a Students' Representative Report, a Lecturer-in-Charge Report, a Teaching Report, and a Programme Coordinator Report. All this information is collected online, and the specific forms are available for the different actors, at the end of each semester.

The QUC mechanism includes four phases: 1) Assessment; 2) Analysis; 3) Improvement and 4) Supervision. The first phase (Assessment) begins with the evaluation of the course when students answer an online survey at the end of each semester, organized in two sections. In the first part, students are questioned about the perceived workload in each course they were enrolled during the semester, and if they wish to answer a second section of the survey in order to evaluate teaching effectiveness (optional).

The second section consists of 23 items answered using a 9-point Likert-type scale, ranging from 'strongly agree' (9) to 'strongly disagree' (1), and is divided into five groups, as shown in Table 1: (i) 'Workload', (ii) 'Organization', (iii) 'Evaluation', (iv) 'Perceived Learning' and (v) 'Teaching Staff' (the latest group is divided into the sub-dimensions (v.1) 'Advantages from in-class learning', (v.2) 'Pedagogical ability' and (v.3) 'Interaction with students').

*Table 1: Items used to assess teaching effectiveness on QUC survey.*

| <b>i. Workload</b>  |
|---|
| Previous knowledge was enough to follow up this course unit   |
| Characterization of the level of importance attached to the study resources, if used in this course unit: |
| Attending theoretical/seminar classes   |
| Attending problem-solving classes   |
| Attending laboratory classes  |
| Suggested bibliography  |
| Notes and other teacher paperwork   |
| Notes and other student paperwork   |
| Further information that is publicly accessible   |
| In which range of values do you include your evaluation in this course unit                               |
| <b>ii. Organization</b>   |
| The programme was taught as scheduled   |
| The course unit was well structured   |
| The bibliography suggested was appropriate  |
| The supporting materials were appropriate   |
| <b>iii. Evaluation</b>  |
| The evaluation method was appropriate to the course unit contents   |
| The evaluation process was fair/equitable   |
| <b>iv. Perceived Learning</b>   |
| The course unit has contributed to the acquisition and/or development of skills:                          |
| To develop knowledge and understanding of the theme   |
| To increase the ability to apply acquired knowledge on the theme  |
| To develop the critical judgment and the ability to discuss the theme                                     |
| To promote the ability to cooperate and communicate   |
| To increase the ability to learn autonomously   |
| To increase the ability to analyze the implications of the theme on the social context                    |
| <b>v. Teaching Staff</b>  |
| <u>v.1. Advantages from in-class learning</u>   |
| The teacher has regularly and timely carried out the academic activities                                  |
| The content and pace of the classes were appropriate  |
| <u>v.2. Pedagogical ability</u>   |
| The teacher showed strong commitment  |
| The teacher has presented the contents in a attractive manner   |
| The teacher has presented the contents with clarity   |
| The teacher has shown confidence during his/her classes   |
| <u>v.3. Interaction with students</u>   |
| The teacher has stimulated participation and discussion   |
| The teacher has been willing to clarify doubts, in class or otherwise                                     |

The information is statistically processed, and each one of the groups is classified according to its median. Based on the classification obtained for the sub-dimensions v.1, v.2 and v.3 a Global Rating for the teacher (an overall, summative judgement ranging from 'Inadequate' to 'Excellent') is also computed, according to the conditions summarized in Table 2.

*Table 2: Conditions for computing the QUC Global Rating.*

|  |
|--|
| <i>Inadequate results</i> : If the average of the median classifications in v.1, v.2 and v.3 is $\leq 3$ (1-9 scale) |
| <i>To improve results</i> : If the average of the median classifications in v.1, v.2 and v.3 is ]3,5[ (1-9 scale)    |
| <i>Regular results</i> : If the average of the median classifications in v.1, v.2 and v.3 is [5,8[ (1-9 scale)       |
| <i>Good results</i> : If the average of the median classifications in v.1, v.2 and v.3 is [8,9[ (1-9 scale)          |
| <i>Excellent results</i> : If the average of the median classifications in v.1, v.2 and v.3 is = 9 (1-9 scale)       |



A compilation of these results is displayed on the QUC website (<http://quc.tecnico.ulisboa.pt/en/>) available for the IST academic community and, in order to avoid presenting statistics that give little guidance for action, the ratings have colours associated (from red to blue). For example, in Figure 1 is represented an evaluation sample for the group 'Teaching Staff'.

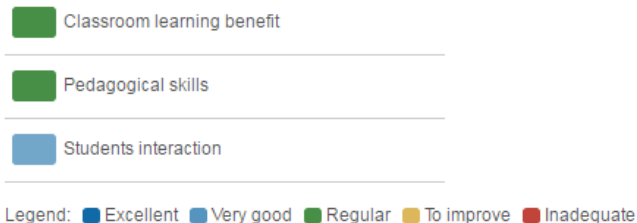


Figure 1: Example of QUC 'Teaching Staff' results.

Whenever a teacher/course is identified with low results the students' representative is required to make a comment in which the problem is itemized, and triggers the second phase of this process (Analysis). In this stage, the teacher of the course together with lecturer in charge analyse and comment the students' representative remarks.

In the most serious cases, an auditing process is activated for the detailed analysis of the problems, aiming for their consequent rectification, corresponding to the third phase of the QUC mechanism (Improvement). The audit team (constituted by two members of the Pedagogical Board, one student and one teacher) starts this process by interviewing those who can make an impact in the enhancement of the teaching practices. Afterwards, these feedback together with the inputs from the students' representative and the teachers of the audited course, allow the audit team to propose a solution along with a number of recommendations for the future.

At this point, one question arises: what is the adequate solution for improving an audited teacher's performance? There is not a formula that fits all the cases. Nevertheless, at IST the strategy adopted was the following: in order to identify what could be improved and help the audit team, a psychologist specialized in pedagogical practices observes the teacher's performance in class, and then recommends some workshops (e.g. conflict resolution, teamwork or public speaking) that could benefit the subject. This leads to the last phase of the process (Supervision), which is continuous in time. The progresses in the teaching practises of the audited teachers are monitored by the Pedagogical Board and if the problem reappears on the following years, stronger measures will be taken.

Rather than focusing only in the negative aspects, positive emphasis is given to excellent results, by publishing a list of excellent teachers in each semester, by disseminating identified good teaching practices on short online videos, and by giving a money award to the best teacher in each cycle of studies. However, despite many teachers applaud this use of ratings, not everyone is so enthusiastic. Some teachers view ratings as popularity contests with meaningless quantification, and warn against teachers who 'buy' their ratings by giving good grades, or students who use this opportunity to get even at teachers.

As meaningful academic community engagement is crucial for the process, one problem emerges: how can one prove the validity of QUC's student ratings? For purposes of the present study, the previous question was resolved into two more specific ones: How stable are student ratings of the same teacher giving the same course in two different years? What is the relation between student ratings, student grades and subjective assessment of learning?

Data for these study came from the latest QUC student surveys available. With the intention of comparing student ratings between two consecutive years, we used the Global Ratings of 2014/15 and 2015/16, selecting the instances in which the same teacher taught the same course in both years, and performed a paired t-test. To assess the relation between student ratings, student grades and subjective assessment of learning we selected the following information from the 1<sup>st</sup> semester of 2016/17: Global Ratings, Student Grades and Perceived Learning computed as the mean classification of group (iv) from



the QUC survey (Table 1). A Pearson product-moment correlation coefficient was computed to assess the relationship between the three variables.

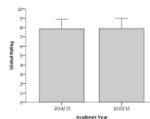
## Results

Since its implementation, QUC has an average of 600 curricular units monitored, 10000 students engaged, 80 students' representatives involved and 700 teachers evaluated per semester. The students' survey response rate has been stable through the years, and on average 70% of the students choose to complete the QUC survey, on each semester.

If the ratings students give to a teacher are stable (as found in the studies described before) then student ratings of the same teachers giving the same course in two consecutive years should not differ significantly. Conducting a paired t-test, using as data the Global Ratings from the years of 2014/15 and 2015/16 ( $n = 1366$  pairs), we observed that there is no significant difference between the generations ( $7.87 \pm 1.02$  vs  $7.90 \pm 1.06$ ,  $p = 0.44$  - Table 3, Figure 2).

*Table 3: Global Ratings of the same teachers giving the same courses in 2014/15 and 2015/16 (paired t-test).*

|               | 2014/15         | 2015/16         | p    |
|---------------|-----------------|-----------------|------|
| n             | 1366            | 1366            | -    |
| Global Rating | $7,87 \pm 1,02$ | $7,90 \pm 1,06$ | 0,44 |



*Figure 2: Mean Global Ratings ( $\pm 1SD$ ) of the same teachers giving the same course units in 2014/15 and 2015/16.  $n = 1366$ . Paired t-test,  $p = 0.44$ .*

The correlation coefficients between the variables Global Ratings, Student Grades and Perceived Learning ( $n = 1155$ ) are shown in Table 4. Two scatterplots summarize the results (Figure 3, Figure 4).

*Table 4: Pearson Correlations between Global Ratings, Student Grades and Perceived Learning.*

| Variable              | a | b    | c    |
|-----------------------|---|------|------|
| a. Global Ratings     | - | 0.10 | 0.43 |
| b. Student Grades     | - | -    | 0.27 |
| c. Perceived Learning | - | -    | -    |

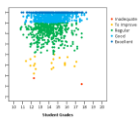


Figure 3: Global Ratings versus Student Grades.

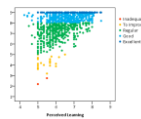


Figure 4: Global Ratings versus Perceived Learning.

The relation between Global Ratings and Student Grades, as well as the relation between Student Grades and Perceived Learning were low ( $r = 0.1$  and  $r = 0.27$ ). However, a more positive relation emerged between Global Ratings and Perceived Learning ( $r = 0.43$ ), supporting the superiority of subjective learning as a predictor of student ratings.

### Conclusions

These findings add clarification to the ratings' validity issues, and we hope they will enhance the acceptability and usefulness of student data in the teaching performance evaluations. We find that students do not rate their teachers impulsively, or according to their grades, but rather according to how much they believe they have learned. The practical implication is that students' ratings of instruction are stable and much more a function of the perceived quality of teaching than of the received grades.

These results suggest that students are qualified to rate their teachers and the quality of teaching they receive: they can report the teacher behaviours (including his availability, the fit of his pedagogical approach, and the clarity of his explanations), the amount of work required, how much they feel they have learned, and the difficulty of the course. Consequently, the weakness of ratings has more to do with their use, than with the validity of ratings themselves: when policy is unclear, or when results are disseminated unconsciously, ineffective use of ratings can easily come.

It is vital to build up trust between faculty members and students, by establishing internal forums, with the aim to share and discuss the teaching and learning quality assurance procedures, and thus, ensuring a broadly shared understanding of quality. The QUC system is widely known among IST community, not only for the clarity of the information provided, or the efforts that were put into planning and implementing a comprehensive faculty-wide communication strategy, but especially because of the existence of practical results. The results of the evaluations have allowed the faculty to take action on problems such as lack of fit between the students' expectations and the contents of the courses, unfair evaluation processes, or teachers with insufficient pedagogical ability.

Student evaluations of teaching performance should aid the institutions to evaluate their teaching and learning effectiveness, provide information for administrative decisions, and be a source of useful information for current and potential students in the selection of programmes and courses. Above all, student ratings should enable enhancement and change: in the students themselves, in the student experience, and in the institution as a whole.

### References:

Albanese, M. A., Schuldt, S. S., Case, D. E., & Brown, D. (1991). The validity of lecturer ratings by students and trained observers. *Academic Medicine*, **66**(1), 26-8.





- Aleamoni, L. M. (1999). Student rating myths versus research facts from 1924 to 1998. *Journal of personnel evaluation in education*, **13**(2), 153-166.
- Aleamoni, L. M., & Hexner, P. Z. (1980). A review of the research on student evaluation and a report on the effect of different sets of instructions on student course and instructor evaluation. *Instructional Science*, **9**(1), 67-84.
- Avery, R.J., Bryan, W.K., Mathios, A., Kang, H., & Bell, D. (2006). Electronic course evaluations: Does an online delivery system influence student evaluations?, *Journal of Economic Education*, **37**(1), 21–37.
- Baird, J. S. (1987). Perceived learning in relation to student evaluation of university instruction. *Journal of educational Psychology*, **79**(1), 90.
- Ballantyne, C. (2003). Online evaluations of teaching: An examination of current practice and considerations for the future, *New Directions for Teaching and Learning*, **96**, 103–112.
- Bennett, L., & Sid Nair, C. (2010). A recipe for effective participation rates for web-based surveys, *Assessment & Evaluation in Higher Education*, **35**(4), 357–365.
- Center for Teaching and Learning - Stanford University (1997). *Using Student Evaluations to Improve Teaching, Speaking of Teaching*, **9**(1), 1-4.
- Chen, Y., & Hoshower, L. B. (2003). Student evaluation of teaching effectiveness: An assessment of student perception and motivation, *Assessment & Evaluation in Higher Education*, **28**(1), 71–88.
- Cohen, P. A. (1981). Student ratings of instruction and student achievement: A meta-analysis of multisection validity studies. *Review of educational Research*, **51**(3), 281-309.
- Cohen, P. A. (1986). An Updated and Expanded Meta-Analysis of Multisection Student Rating Validity Studies.
- Costin, F., Greenough, W. T., & Menges, R. J. (1971). Student ratings of college teaching: Reliability, validity, and usefulness. *Review of Educational Research*, **41**(5), 511-535.
- Dommeyer, C. J., Baum, P., Hanna, R. W. & Chapman, K. S. (2004). Gathering faculty teaching evaluations by in-class and online surveys: their effects on response rates and evaluations, *Assessment & Evaluation in Higher Education*, **29**(5), 611–23.
- Feldman, K. A. (2007). Identifying Exemplary Teachers and Teaching: Evidence from Student Ratings. In *The scholarship of teaching and learning in higher education: An evidence-based perspective* (pp. 93-143). Springer Netherlands.
- Gigliotti, R. J., & Buchtel, F. S. (1990). Attributional bias and course evaluations. *Journal of Educational Psychology*, **82**(2), 341.
- Guthrie, Edwin Ray. The evaluation of teaching: A progress report. University of Washington, 1954.
- Hativa, N. (1996). University instructors' ratings profiles: Stability over time, and disciplinary differences. *Research in higher education*, **37**(3), 241-265.
- Hoffman, C. D. (1979). Students learn more from better teachers. *Teaching of Psychology*.
- Hogan, T. P. (1973). Similarity of student ratings across instructors, courses, and time. *Research in Higher Education*, **1**(2), 149-154.
- Howard, G. S., & Maxwell, S. E. (1980). Correlation between student satisfaction and grades: A case of mistaken causation?. *Journal of Educational Psychology*, **72**(6), 810.
- Marsh, H. W. (1987). Students' evaluations of university teaching: Research findings, methodological issues, and directions for future research. *International journal of educational research*, **11**(3), 253-388.
- Marsh, H. W., & Roche, L. A. (1997). Making students' evaluations of teaching effectiveness effective: The critical issues of validity, bias, and utility. *American psychologist*, **52**(11), 1187.



Marsh, H. W. & Roche, L. (1993) The use of students' evaluations and an individually structured intervention to enhance university teaching effectiveness, *American Educational Research Journal*, **30(1)**, 217–251.

Lawall, M. (1998). Students rating teaching: How student feedback can inform your teaching. University Teaching Services, Centre for Higher Education Research and Development, University of Manitoba.

Palchik, N. S., Burdi, A. R., Hess, G. E., & Dielman, T. E. (1988). Student assessment of teaching effectiveness in a multi-instructor course for multidisciplinary health professional students. *Evaluation & the Health Professions*, **11(1)**, 55-73.

Scoles, K., N. Bilgutay, and J. Good. (2000). A new course evaluation process, *IEEE Transactions on Education*, **43(2)**, 125-131.

Seldin, P. (1989). Using student feedback to improve teaching. *New Directions for Teaching and Learning*, **1989(37)**, 89-97.

Sorenson, D. L., & Reiner, C. (2003). Charting the uncharted seas of online student ratings of instruction, *New Directions for Teaching and Learning*, **96**, 1–29.

Theall, M. (2002). Student ratings: Myths vs. research evidence. Retrieved June, 26, 2008.

Theall, M., Abrami, P. C., & Mets, L. A. (2001). The student ratings debate: Are they valid? How can we best use them?.

Watkins, D. (1994). Student evaluations of university teaching: A cross-cultural perspective. *Research in Higher Education*, **35(2)**, 251-266.

**Discussion questions:** Do you feel that students are qualified to rate their teachers? Experiences of other institutions at this level: What are the difficulties encountered when using student ratings? What actions does your institution take with the output of student evaluations?