9th European Quality Assurance Forum

13 – 15 November 2014

University of Barcelona, Spain

Changing education – QA and the shift from teaching to learning

Author(s)

Name: Elisa MAZZUCATO

Position: Decision Support Cell

Organisation: UNIVERSITY OF LUXEMBOURG

Country: LUXEMBOURG

E-mail address: elisa.mazzucato@uni.lu

Short bio (150 words max):

Elisa Mazzucato joined the young University of Luxembourg in 2010 with the aim of developing and providing strategically focused business information and performance data on the University. She is also responsible for working with others across the University to enhance the quality, scope and accessibility of University management information and of a number of important statutory statistical returns to government and international organization.

Elisa Mazzucato completed her master studies in statistics applied to social sciences and demography at University of Padova, Italy. She has an extensive experience in statistics, information management, research methodology and evaluation. She cooperated largely with International institutions such as ILO, Eurostat, OECD, Luxembourg Ministries and conducted research on social research methodology at KU Leuven, Belgium.

(If there are several authors, please copy and fill in the fields for each author.)

After the Forum, the full text of all papers presented at the Forum will be published on the Forum website. If you do not wish your paper to be published, please indicate so here. This has no consequences on the selection of the papers.

Proposal

Title: STUDENTS' COURSE EVALUATION. A SHIFT OF PARADIGM AT UNIVERSITY OF LUXEMBOURG

Abstract (150 words max): University of Luxembourg has recognized the difficulty of studentsᶜcourse evaluation: it has replaced a formal evaluation system with a more open and flexible one, taking into account the dynamic and nonstandard process of teaching and the complexity in quality definition. The paradigm shift has been to provide a system that would allow ongoing improvement of a given course through a solid and continuous cooperation among students, lectures and administration; and not just be the final and definite judgment of the quality of teaching. At the same time it gives managers sufficient information to steer the organization toward its goals by means of a simple set of indicators having essentially a diagnostic purpose. The new evaluation system has shown the way to an innovative paradigm shift on how to consider evaluation results validity: where improvement is the main goal, consistency of meanings across interpreters may become much less important.

Text of paper (3000 words max):

1. Courses evaluation: what has been our experience?

University of Luxembourg has used a traditional teaching evaluation system essentially having a summative purpose, based on an online questionnaire submitted to students at each end of semester. The questionnaire counted 11 items to be measured on a five-level Likert scale. The items focused essentially on students' perception of teaching quality. Students were invited to express their opinion on each lecturer of a given course by means of an electronic form accessible via the Moodle platform.

The average participation was around 30% of the registered students. The global satisfaction level was around 77%. This result together with the registered reasonable participation rate - often around 30% for an online survey - appeared satisfactory. Nevertheless the system offered some room for improvement.

The number of items to be answered by each student each semester was substantial. Each student was requested to answer on average 8.78 questionnaires each semester, for a total of 150 items per students. On the other hand, data showed a very high co-graduation of responses meaning that students' effort in completing the questionnaire did not generate an appropriate amount of information. Once the effect of items co-graduation was controlled[1], the only information that could be extracted from data was a general level of satisfaction per lecturer. This measure was indeed very important but it could not be considered informative in indicating new directions towards quality improvement.

Insufficient information return that could feed quality change in poor teaching practices impacted on the ability of giving a substantial feedback to students on evaluation results. A global participation rate and percentage of satisfied/dissatisfied respondents were the only two figures that could be returned to students. As the global satisfaction was generally good and no specific teaching issues were brought out by data analysis, students demonstrated a certain disappointment. Such a positive feedback on general satisfaction did not reassure students that University uses the course evaluation results to detect and appropriately treat possible issues. On the other hands, students had the impression of not being able to reward very good teachers. The opinion that the students' effort on completing evaluation questionnaire was useless and the impression that it could not boost a real action for quality improvement was reinforced. Moreover the motivation of students in participating in the survey could be reduced through similar feedback. In particular, dissatisfied students may be less keen to express their opinion, as they can imagine that it is not worthwhile to do so: "nothing will change[2]". Students may become more involved in the evaluation exercise if they are reassured that their comments are taken into account and that they may lead to a real action for improvement[3].

Although course evaluation was generally accepted, the proposed methodology still gave rise to some controversies within academia. Some of the resistance and criticisms coming from lecturers are very common in teaching evaluation literature:

- evaluations are biased,
- students are not competent evaluators,
- ratings are impacted by student grade expectations or students' cognitive and non-cognitive abilities
- and the use of online questionnaires may be responsible for a poorer response rate and lower accuracy than traditional paper and pencil methods.

Furthermore, evaluation systems were perceived as a tool used by administrators to support personnel decision, rather than being a tool that empowers students to share problems and solutions with the academic community. So even if faculty believes evaluations to be useful in assessing teaching, lecturers and program directors rarely used the results of their own evaluation in course or professional development decisions.

---

[1] by means of a principal component analysis for ordinal data
[2] Research into complaint behavior reveals that only a fraction of dissatisfied consumers complains and, thereby, gives the service or product provider an opportunity to correct the problem. There is evidence that some consumers do not complain because they are skeptical about business's willingness or ability to resolve disputes fairly. Consumers simply withdraw their patronage and criticize the company or the product to others.
[3] Business experiences show that consumers who complain about products and services continue to frequent the businesses and buy the products they complain about if they believe the complaint was resolved fairly.

Even though the evaluation was carried out as an online survey through the Moodle platform, the complexity of the process and the need for further manual intervention in preparing questionnaires limited the possibility of efficiency gain, normally enhanced by the use of an online survey. For instance, each semester three dedicated resources had to be made available to manually input questionnaires for each course and each lecturer. The use of online support must bring the advantage of saving time and resources over the traditional paper and pencil scan sheet method, without losing flexibility and quality in data production.

2.   What do we propose?

The criticisms expressed by students, lecturers and administrative have been evaluated carefully, and the consequence was to replace a traditional evaluation system with a more open and flexible evaluation tool, that takes into account the variability and intrinsic differences in the learning process and in the definition of its quality[4]. The paradigm shift was to provide a system that would allow ongoing improvement of a given course through a solid and continuous cooperation among students, lectures and administration; and not just be the final and definite judgment of the quality of teaching. Thus the new tool had to:

- assess quality in its heterogeneity, considering a certain degree of autonomy in determine quality of teaching, empowering students as partners in quality improvement[5], and focusing on administrative efficiency.
- Empower faculty managers, teachers and study directors to constantly improve the quality of given courses by acting promptly in solving emerged issues, reinforcing the existing programs or spreading identified positive procedures to other courses.
- Maintain the possibility of use evaluation results to feed personnel decisions and course selection.

Firstly, the change has been to focus evaluation on a course considered as a whole, instead of making students evaluate each lecturer. A course is organized on the base of a series of learning outcomes. Learning outcomes are achieved by means of quality teaching but also through a good and smooth course organization and the availability of learning resources. In this context, a quality course should be seen as the product of the cooperation among lecturers, program directors and faculty more than as a sole expression of a quality lecturer. It is believed that evaluating each lecturer for themselves is not a good incentive to cooperation and interaction towards quality course improvement.

Secondly, a formative evaluation process comes to integrate a more traditional summative one. A formative feedback is described as information gathered for the purpose of improving and developing teaching. This information is meant to inform change. The feedback itself must be specific and concrete enough to suggest actions for improvement (Rando and Lenze (1994)). Indeed, Harper and Kuh (2007) noted that qualitative means of assessment can often bring to light issues that cannot emerge through conventional quantitative means. Thus a set of open-ended question has been tested for this purpose. The summative feedback has been reduced to a short, economical form. This approach is supported by the Cashin and Downey (1992) study on data from 17,183 courses representing 105 institutions. It found out that short summative forms are able to gather much of the information needed for summative purposes. Detailed information are then apprehended by the formative evaluations.

3.   The questionnaire

A study of other Universities'experiences has been conducted while developing an evaluation questionnaire (fig.1), in order to profit from their expertise and competence on that topic. The following criteria[6] have guided our choice:

- Reduce length.

---

[4] Recent attempts to define teacher quality have sought ways to broadly represent the views of the field and to benefit teacher development and assessment.

[5] According to OECD guidelines suggesting the acquisition of a more significant role of students in quality of teaching process.
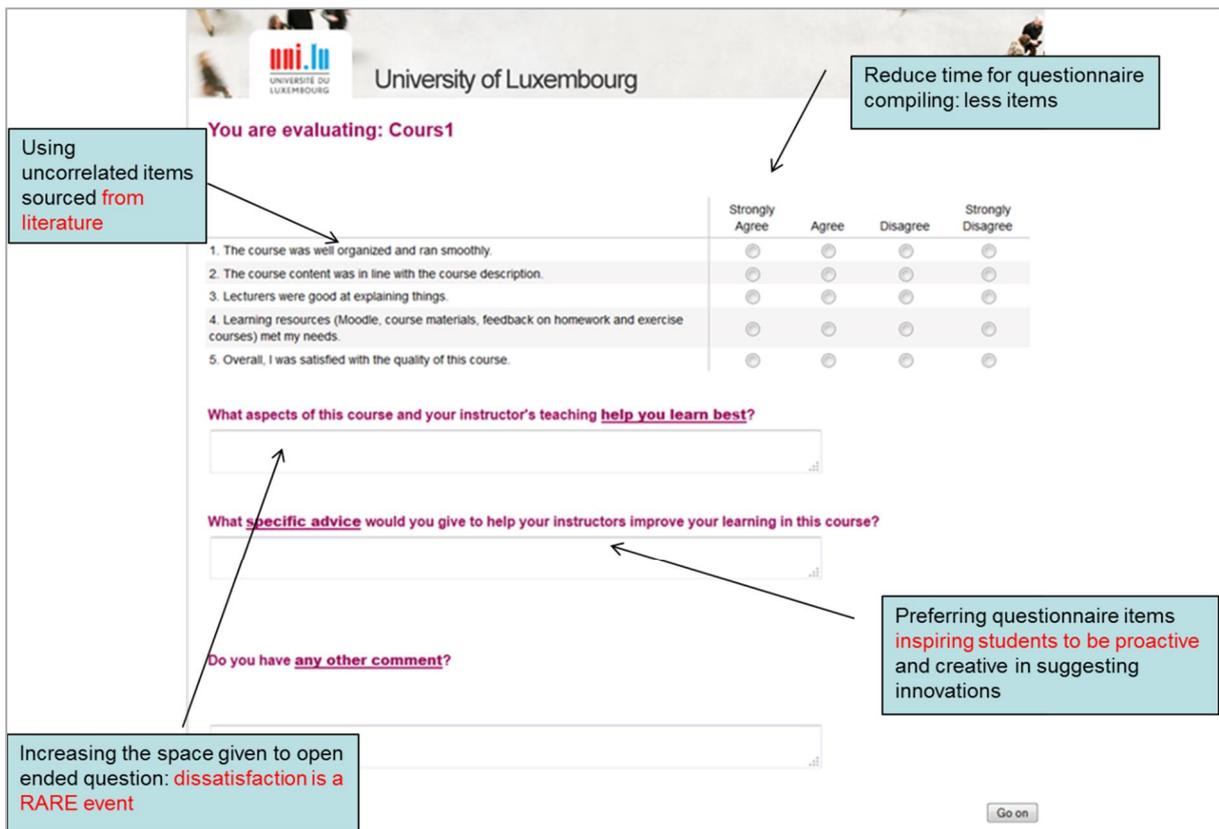
[6] A new process of revision of the instruments is foreseen after a substantial number of information and feedback are collected.

- Flexibility.
- Maximal independence of items.
- Providing both summative and formative results.

In order to reduce time for completing, the length of the questionnaire has been considerably reduced[7]. Many researchers have in fact suggested that the survey length can affect both participation rate and validity (Scriven (1995)). The number of questions has been decreased to five closed items, and three open-ended questions. The five close-ended questions come from the core set of questions in the *"Course Evaluation Questionnaire"* used at Sussex University. The Sussex questionnaire presents the advantage of being thought up as a modular system. A core set of items is defined at central level, and some other items are suggested to be used as program-level and module-level questions.

This choice is in line with the general idea of empowering program directors and lecturers to tailor their forms according to their own needs and practices - adding questions of interest as appropriate. This flexibility is the result of the maximal exploitation of the advantages offered by online surveys. Each course can be associated in advance directly at course database level to three compulsory and centrally decided questions, and to two others to be chosen by study directors. The online survey tool then automatically adapts each questionnaire to each course.

The three open ended questions were inspired by mid-term studentsqevaluation questionnaire of Princeton University. Their main characteristic is to force students to assume an active, rather than passive role in evaluation, encouraging them to be proactive and creative in suggesting innovations.

FIG.1 Course evaluation questionnaire, English version

---

[7] Each students attend on average eight to nine courses each semester. Each semester students should answer a global number of items that is equal to the number of questionnaire items times the number of courses

EUA
European University Association

ENQA
EUROPEAN ASSOCIATION
FOR QUALITY ASSURANCE
IN HIGHER EDUCATION

EURASHE

eSU
european students' union

4. Integration with courses database

The adoption of online questionnaire can considerably decrease the administrative workload of course evaluation. Efficiency can be boosted by building a system that integrates, as much as possible, the courses/students databases, containing all information necessary to conduct evaluation (programs, registered students, courses, lecturer names, types of question, evaluation dates, etc..) and the online survey tool.

University decided to shift to Qualtrics[8], a well-known online survey software. Qualtrics allows a better integration of databases and the creation of semester panels containing all information needed to routinely manage the evaluation process (fig.2). The questionnaire is automatically tailored to each student and each course, according to the uploaded semester panel, and it does not need to be modified manually each semester. This represents an enormous gain of time and resources. We have been able to reduce personnel cost from 15 man-day to 3 man-day. A further reduction of man-day is ongoing by increasing the quality of students/courses databases. Part of the current manual workload concerns students/courses data quality checks as some errors are still present when inputting data or some coursesqstructure change during a given semester.

The reduction of manual intervention in the process has also the advantage of minimizing error. Furthermore, an indirect effect of this integration is the improvement of the students/courses database, as the quality check needed to have a correct lists to be uploaded into Qualtrics, makes it necessary to correct the data source each semester.
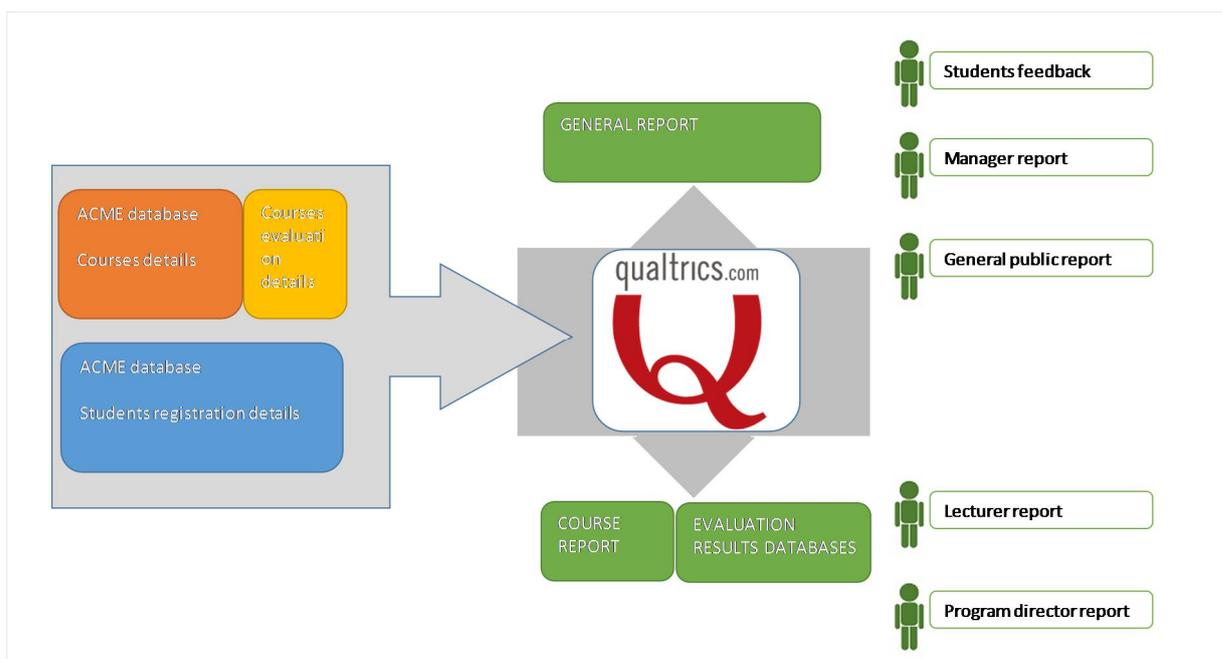

FIG.2 Information flows

5. Data interpretation and reporting

Summative aggregated results are used as a first diagnostic instrument that is able to detect global issues or extremely good courses. Formative feedback then clarifies scores obtained in summative evaluation[9]. Detailed reports are produced for each course for lecturers and program directors, and global reports for managers, students and the general public.

The global report includes aggregated summative results (examples of report in fig.3-4). Each course is classified on the base of a global indicator depending on the overall studentsqsatisfaction and the number of

---

[8] http://www.qualtrics.com/
[9] Several studies in fact have shown that there is a correlation between the qualitative and quantitative ratings (Cashin, 1995) supporting the diagnostic power of the summative ratings.

respondents per course. The course is then defined as excellent, good or it receives a warning. The number of excellent/good and warning courses are then communicated to central administration and students by mean of feedback sessions. Students and administrators can then appreciate the evolution of course quality over time[10].
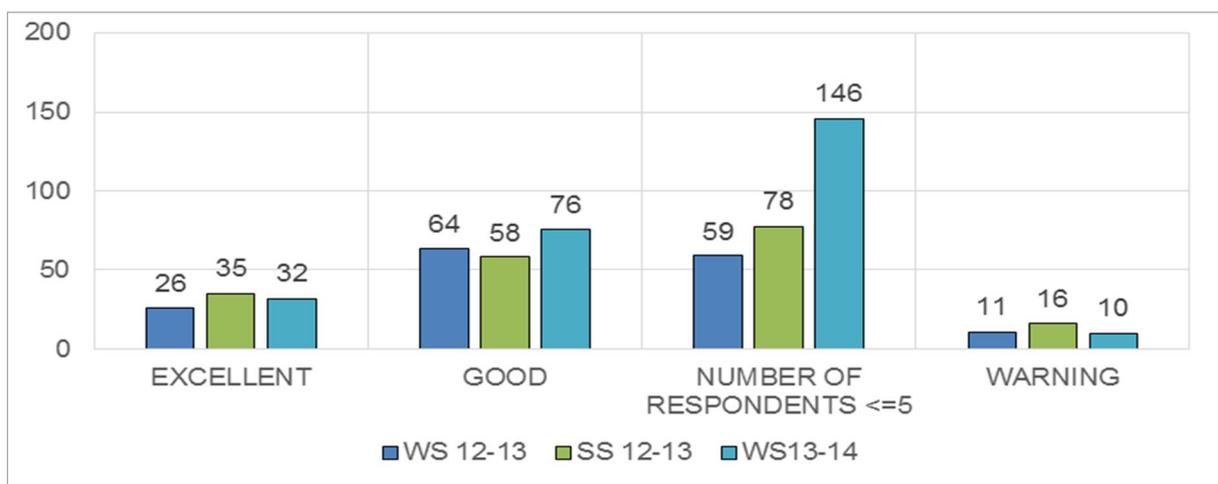
	Target	Respond ent	Question naire	Respons e rate	Question naire per responde nt	Overall satisfacti on	In line with descripti on	Lectures	Resource s	Organisa tion	# courses Excellent	Warning
BACHELOR1	397	109	340	27%	3,1	77%	82%	78%	73%	81%	9	4
BACHELOR2	353	95	371	27%	3,9	78%	81%	80%	73%	81%	1	2
BACHELOR3	447	76	279	17%	3,7	77%	80%	79%	74%	77%	4	3
MASTER1	33	16	52	48%	3,3	85%	89%	88%	84%	88%	-	-
MASTER2	96	37	124	39%	3,4	74%	88%	77%	72%	79%	3	1
MASTER3	20	7	35	35%	5,0	89%	100%	91%	91%	91%	-	-
MASTER4	171	74	386	43%	5,2	92%	90%	92%	91%	90%	15	-
Total	**1517**	**414**	**1587**	**27%**	**3,8**							

FIG.3 Summative reporting example per program

FIG.4 Summative reporting example: courses classification over time

The course report contains the frequency distribution of closed questions and all answers to open-ended questions (see Annex). The latter do not receive any treatment and they are integrally repeated in the report. Results are stored in a shared database, and lecturers and program directors can access their own results any time, appreciating quality improvements, if any, over time.

Moreover each lecturer or program director is invited to classify studentsʹ feedback into complaints/praises/general note/suggestion or improper remark. For each remark an implemented action should be recorded, if any. The action can be of different types: clarification, modification of procedure, and investigation

for more details, etc.

Clearly the challenge is the perceived increase in time and effort needed to assess and classify written comments[11]. Measuring and reporting faculty membersʹ engagement by using studentsʹ feedback can encourage

---

[10] At the moment an aggregate summative indicator is studied to publish evaluation results via a program on the web.
[11] Several studies (Beran, Violato & Kline, 2007; Beran et al., 2005; Wagenaar, 1995) of the use of evaluations by instructors and administrators indicate that these groups rarely review written comments, preferring instead to use only what they perceive to be the more time-efficient global ratings

lecturersꞌclassification effort. The lecturersꞌeffort can be summarized by a global indicators counting number of centered/meaningful complaints over number of improvement actions[12]. Lecturers should then be endowed accordingly. In more general terms a high level of student participation and lecturer activities in acting promptly can be detected by such an indicator and can give managers a measure of quality maturity level of the teaching process.

It appears as evident that the system is clearly not based on a sanctioning approach, and it tends to give a certain autonomy and responsibility to lecturers and program directors in interpreting and evaluating studentsꞌ feedback. It is in fact believed that lecturers must play a central role in integrating formative information into the more traditional summative one and have a certain degree of freedom in interpreting their own results and describing their own reality, acting accordingly. At the management level, this approach allows to focus on the most committed lecturers and best practices, and to quickly detect possible issues. Particularly good lectures can then be rewarded and may lead university as a whole towards course quality improvement, as well as more information can be requested when courses are judged as poor. The tool is the expression of a total quality logic: the more the students participate through their remarks, either negative or positive, the more the organization is pushed into the direction of quality and excellence.

6.   Building a common quality culture

Some authors suggest that training should be offered to students, instructors and administrators about the value of written comments and on techniques for, respectively, writing and analyzing these comments effectively. In general, it is believed that a better use of this tool is possible only if a common quality culture lies underneath. To build a shared concept of quality we proceed by giving feedback on evaluation results to students, lecturers, faculty and university managers, and by creating a student quality circle. This quality circle is involved in collaborative endeavors that encourage continuous university development. It is composed of a voluntary group of students, who meet regularly and are encouraged to identify, analyses and solve university-related problems and present their solutions to management and where possible implement the solutions themselves in order to improve the performance of the organization, and motivate and enrich the general university processes and output. On the lecturersꞌside, faculties put in place according to their own quality strategy, various activities aimed at the development of teaching quality and spreading and evaluation culture.

7.   Is response rate a real central issue?

Despite a significant majority of lecturers considering student evaluations to be a useful measure of the instructional behaviors that contribute to teaching efficiency and regardless of the advantages of the evaluation system here described, the instructorsꞌresistance to evaluation is still present. The most important concern is about evaluation validity, caused by lower return rates, often associated with the use of online questionnaire.

Both literature[13] and experiences at University of Luxembourg showed that paper and pencil in class evaluation give better results in terms of participation rate. This being so we decided to use online questionnaire to simulate a paper and pencil survey in class, profiting from free Wi-Fi accessible within the university perimeter. Lectures having access to Qualtrics system can send the link for accessing evaluation to studentsꞌportable devices during classes. This approach maintains the advantages of the online platform and integrates those of the more traditional paper and pencil methods. The first pilot experience will be conducted during the winter semester 2014/15.

Nevertheless, despite the effort in increasing studentsꞌparticipation, validity concerns remain as

- most of the courses are attended by a small number of students,
- poor courses are still a statistically rare event,
- variability of course issues is very large[14],
- and non-sampling errors may affect data even with a 100% participation rate.

---

[12] Other indicators can then be computed according to various reporting goals as shown in Fig.2.

[13] It worth to be noted that Donovan, Mader, and Shinsky (2006) have found that by using online formats open-ended comments will not only be quantitatively greater in number and length, but they will contain more qualitative detail than is likely to be found in traditional evaluations.

[14] Validity can also be questioned as how quality can be measured or evaluated if there is no general agreement on what is effective teaching, as no "universal set of characteristics of effective teachers and courses that should be used as a targetō appears to exist+(Ory and Ryan (2001).

It is then worth asking whether or not we should strive at maximum in increasing statistical significance of summative estimates rather than using results as an efficient first diagnose of the level of wellbeing of the course, and then analyze and treat peculiar issues using formative assessments. On the other hand, where improvement is the main goal, consistency of meanings across interpreters is much less important. If two different assessors interpret a particular teaching performance differently, and suggest different professional development activities that would be equally successful in moving the teacher's learning forward, then, they would be equally valid in terms of their impact on teacher learning[15].

While the focus moves from quality improvement to personnel decision or course selection, validity can be reinforced by looking at summative results over time. This indication together with the increase in response rate (as it is believed that teachers' commitment and communication on courses evaluation can indeed increase the students' motivation in participating in evaluation exercise) and the number of actions implemented for quality improvements should be used to build evidence based decisions on personnel awards and course selection.

8. Conclusion

University of Luxembourg has recognized the difficulty of students' course evaluation: it has replaced a formal evaluation system with a more open and flexible one, taking into account the dynamic and nonstandard process of teaching. The tool allows summative and formative evaluation, and gives an important part to students and teacher in developing, creating and improving the teaching process. At the same time it gives managers sufficient information to steer the organization toward its goals by means of a simple set of indicators having essentially a diagnostic purpose.

In the long run, it is believed that this evaluation system as well as the valuable job provided by quality officers, represent a step forward in building a solid shared quality and evaluation culture at University of Luxembourg.

---

[15] Nevertheless, as already mentioned, teachers need to share the implicit concepts of quality of students when they judge about quality to allow assessment to function formatively.

<Name of program director>

COURSE REPORT

<COURSE NAME>

<Faculty>

<Program name>

Principal indicator[1] :	WARNING/GOOD/ECELLENT or NUMBER OF RESPONDENTS <=5
Number of submitted questionnaires by students[2] (#Q)	X

Lectures
<Name Lectures1>
<Name Lectures2>

Close ended questions, % frequencies

	++	+	-	--
Overall, I was satisfied with the quality of this course.	%	%	%	%
The course was well organized and ran smoothly.	%	%	%	%
The course content was in line with the course description.	%	%	%	%
Lecturers were good at explaining things.	%	%	%	%
Learning resources met my needs.	%	%	%	%

What specific advice would you give to help your instructor improve your learning in this course?

References:

Ory, J.C., & Ryan, K. (2001). How do student ratings measure up to a new validity framework? In M.Theall, P.C Abrami, & L.A. Mets (Eds.), The student ratings debate: Are they valid? How can we best use them? [Special issue]. New Directions for Institutional Research, 109, 27-44.

Scriven, M. (1995). Student ratings offer useful input to teacher evaluations. Practical Assessment, Research & Evaluation, 4(7).

Judy Donovan;Cynthia E. Mader;John Shinsky, Constructive student feedback: Online vs. traditional course evaluations, Journal of Interactive Online Learning 2006 Volume 5, Number 3, Winter 2006 ISSN: 1541-4914

Beran, T., Violato, C., & Kline, D. (2007). What's the ±use of student ratings of instruction for administrators? One university's experience. Canadian Journal of Higher Education, 17(1), 27-43.

Beran, T., Violato, C., Kline, D., & Frideres, J. (2005). The utility of student ratings of instruction for students, faculty, and administrators: A ‰consequential validity+study. Canadian Journal of Higher Education, 35(2), 49-70.

Wagenaar, T.C. (1995). Student evaluation of teaching: Some cautions and suggestions. Teaching Sociology, 23(1), 64-68.

Cashin, W.E. (1995). Student ratings of teaching: The research revisited (IDEA Paper No. 32).Manhattan, KS: Kansas State University Center for Faculty Evaluation and Development.

Cashin, W.E., & Downey, R.G. (1992). Using global student rating items for summative evaluation. Journal of Educational Psychology, 84(4), 563-572.

Rando, William C. & Lenze, Lisa Firing (1994), Learning from students: early term students feedback in Higher Education, University Park, PA: National Center of PostsecondaryTeaching, Learning and Assessment.

Harper, S.R., & Kuh, G. (2007). Myths and misconceptions about using qualitative methods in assessment. In S.R. Harper & S.D. Museus (Eds.), Using qualitative methods in institutional assessment [Special issue]. New Directions for Institutional Research, 136, 5-14.

Questions for discussion:

- Is response rate a real crucial concern in course evaluation?
- Online support to evaluation can increase course evaluation efficiency?
- Is course evaluation exercise sustainable without a shared quality approach among academia, students and university managers?
- Should we prefer quality in students information return, more than quantity?
- Which is the most important goal in course evaluation?.