HOW TO INCREASE TEACHERS’ USE OF THEIR QUANTITATIVE EVALUATION DATA
CASE: THE TECHNICAL FACULTY AT LUND UNIVERSITY

Andersen, Line Palle

Publication date: 2018

Citation for published version (APA):
HOW TO INCREASE TEACHERS’ USE OF THEIR QUANTITATIVE EVALUATION DATA AT THE TECHNICAL FACULTY AT LUND UNIVERSITY

L. Palle Andersen¹, L. Birch Andreasen¹, T. Roxå²
¹Aalborg University (DENMARK)
²Lund University (SWEDEN)

Abstract

Student evaluations at institutions for Higher Education became mandatory in Sweden in 2003 and the Technical Faculty at Lund University (LTH) decided to implement an adapted version of the research-based Course Experience Questionnaire (CEQ) that focuses upon in-depth learning. The questionnaire at LTH consists of 26 quantitative and two free-text questions. Teachers receive the summarised data in a report, and studies at the Academic Development Institution, Genombrottet, at LTH have revealed that teachers primarily focus on the qualitative replies and neglect the quantitative data. The visualisation of the quantitative data is clear, so it seems not to be a question of lack of understanding. The empirical data in this paper is based on focus group interviews and they have revealed that teachers feel uncomfortable about the evaluations, especially regarding the inconsiderate comments that they often receive from students, and thus teachers tend to disregard the evaluation reports as soon as they can. This paper focuses on how to stimulate a culture in the organisation that will make teachers increase their use of the comprehensive dataset contained in the quantitative evaluations with the aim of including this data in the development of their teaching. Engeström’s Activity Theory has been used to analyse the organisational structure and solutions in terms of new tools including short video tutorials aimed at teachers and students respectively. An adjustment of the free-text questions, a revision of the working report and end report and finally an interface that makes it possible for teachers to work with their own quantitative data by comparing them to their previous ones or to selected averages of equivalent data at LTH.

Keywords: CEQ evaluations, Activity Theory, Quantitative data analysis, participatory design

1 INTRODUCTION

This paper is based on an MA-thesis in ICT & Learning written at Aalborg University in Denmark. The empirical data is collected at the Technical Faculty at Lund University (Lunds Tekniska Högskola, LTH). The research question is: How can quantitative evaluation data be made more useful to teachers in order to increase the use of evaluations to develop their teaching?

The strength of evaluations is that they offer the possibility of measuring progress, which is a key aspect of higher education [1]. Evaluations provide intelligence on how to improve the effectiveness of materials and teaching methods and to remove the elements that are not [2].

However, there are many reasons to have reservations regarding evaluation data:

1. Course evaluations are anonymous and in general have a very low response frequency (usually below 40%). Thus we cannot be sure that the students answering are representative of the whole class.
2. Evaluation is teaching- and teacher-focused. It is the teacher being rated even when the question concerns the learning activity. The teacher’s attitude influences the evaluation and a feel-good factor is rewarded [2].
3. Student judgment is affected by class size and study level [3]
4. LTHs in-house studies of the first five years of using the CEQ-system (from 2003-2008) reveal that the education programme a student belongs to affects the outcome of the course evaluation and there is a significant difference between how the different programmes experience the same course [4]
5. Evaluations have a tendency to be gender biased. A study at LTH based on the CEQ revealed that teachers tended to receive higher ratings in subjects that were less typical for their gender than in subjects that were more typical for their gender [5].
6. Good evaluations do not necessarily reflect how much the students learned – remember the popular teacher that gave you detailed lecture handouts, which made it easy to pass the course without reading the course book [2].

Bearing all these reservations in mind one can question whether the result of the course-level quality processes measure up to the effort put in by teachers, students and administrators. Moreover, for many teachers the compulsory evaluations constitute a stress load. Some suffer from being exposed to inconsiderate student comments even to the extent that it can be a workplace health issue [2].

Evaluations at institutions for Higher Education became mandatory in Sweden in 2003, but the educational institutions were free to decide, which evaluation system they wanted to use. From the very beginning, it was communicated by the management at LTH that evaluation results, positive as well as negative, would not influence a teacher’s salary (a promise that all informants agree has been kept). Asked directly, none of the informants feared losing their job because of poor evaluations. Focus at LTH has thus been on using evaluation data for teaching development purposes so that students may learn more.

Why do teachers then practically ignore the quantitative part of the evaluations even when the data presentation is very easy to apprehend and when they are interested in developing their teaching both in terms of content and pedagogy?

1.1 Background: The Evaluation Process at LTH

LTH has a strong tradition for Academic Development represented by the department Genombrottet. The Course Experience Questionnaire (CEQ) was selected based on their recommendations. The CEQ was originally developed at Lancaster University in the 1980s and subsequent research has confirmed the validity and usefulness of the CEQ as a performance indicator of university teaching quality [6]. LTH selected five categories each containing 3-5 questions; Good Teaching, Appropriate Workload, Appropriate Assessment, Clear Goals and General Skills [7]. The original CEQ also contains the category Independence, which LTH deselected. LTH has added two questions to the CEQ, one about the relevance (Q17 The course seems important to my education) and one about satisfaction (Q26 Overall, I am satisfied with this course). The evaluation process at LTH runs in 7 steps:

1. Teachers are asked if they would like to add 1-4 supplementary free-text questions to the questionnaire
2. Students fill in the questionnaire (paper or web-based) and add comments in free-text
3. (student assistants go through the free text answers and are supposed to remove unpleasant comments that do not contribute to course development)
4. The computer system transforms the data into a work report including a data visualisation of the separate quantitative questions and a histogram summarising the categories
5. The data in the working report is discussed at a mandatory meeting between the course coordinator (CC), student representatives (SR) and a director of studies (PD) who is responsible for the whole programme, which the course is part of
6. The end or final report is created containing the statistically processed data (now including the number of pass/fail) and the three concluding remarks from CC, SR and PD respectively
7. The End report is published on the faculty intranet and sent via e-mail to all students who participated in the course [3].

When the CEQ-system was implemented at LTH it was decided by the management that students should play a considerable role in the evaluation process. Students are encouraged to elect two representatives in each class and the chosen ones will conduct a mid-term evaluation including feedback to the course teachers about the status of the course. The student representatives (SR) are supposed to encourage their fellow course students to complete the CEQ, study the questionnaires, participate in the CEQ meeting and write a concluding remark for the End report. Student assistants are paid to go through the free-text answers and remove inconsiderate comments. The student voice is respected as a part of a democracy.
2 METHODOLOGY – PARTICIPATORY DESIGN AND FOCUS GROUP INTERVIEWS

2.1 Participatory Design

Participatory Design (PD) came into existence in Scandinavia in the 70s when computer based technology was used as a way to automate and improve efficiency primarily in the production industry. It started as a strongly political project with the aim to develop workers’ competence and knowledge about computer based technology with the long-term goal to strengthen democracy in the workplace [8]. During the 1980s PD-research changed from the explicitly political projects to focus on “design for the skilled worker”. The context also changed from the production industry to office-environments and the service industry. Designers realised that they had to change the methodology and use experiments in order to gain insight into the users’ “hidden” knowledge about their work. In the experiments, through reciprocal learning, designers and users could show each other the technological possibilities and the core of the working processes that technology could support [8].

In 1991 Joan Greenbaum and Morten Kyng published Design at Work, which has contributions from the majority of the dominating forces/talents in Scandinavian PD. The experience from the PD-projects are gathered and the authors present a number of design ideals that they consider central in the design of computer systems to support working situations. These are:

- Full participation of the users, which implies that users participate in the whole design-process and not just act as informants for the designers
- Computer systems must strengthen the professional competences at the work place – not undermine them
- Computer systems are tools that must be designed for the people who in the end have to use them
- Design of computer systems should not just aim at higher productivity, but also at quality
- Conflicts are part of the design process and must not be ignored, but must be dealt with and solved.
- Design of computer systems must focus on the actual user situation [8]

The solutions or tools are developed in this project by means of PD following those six principles to make teachers utilise their quantitative evaluation data.

2.2 Focus group interviews

As methodology for gathering the empirical data I have chosen focus group interviews, because the source of the data is the social interaction. The knowledge about the complexity of creating meanings and social practices will be produced during the group processes when the participants compare their experiences and understandings. This knowledge will be more difficult to obtain in an individual interview where the group is a means to obtain more complex data [9].

In comparison with field notes and participatory observation, it is considered a weakness for the focus group interview that there is a lot of interesting things that you can only access by being present in people’s existing social context [9]. However, particularly concerning evaluations, they are generally kept very private and rarely discussed in public. A small group of teachers that has accepted an invitation to participate in a focus group interview about evaluations is more likely to share experiences, understandings and actions they would otherwise have kept to themselves.

In order to be able to generalise analytically based on empirical patterns you normally have to make sure that important characteristics regarding the problem area are represented in the selection [9]. In collaboration with Academic Developer Dr. Torgny Roxå we decided to focus on teachers who had shown an interest in the evaluation process and shared a general interest in pedagogy.

The participatory design process was structured in the following way:

1. Focus group interview with three teachers about the evaluation system at LTH. Two of the participants knew each other in advance. (Due to fully booked calendars this interview ended up being split in two with one interview with two teachers (who knew each other) and one with the third teacher). The informants had 14 years or more of working experience at LTH.
2. Focus group interview with the same three teachers discussing suggested prototypes to increase and improve the use of quantitative data.
3. Focus group interview with another three teachers (two knew each other in advance) discussing revised and new prototypes based on the response in the second interview. Among the informants were two younger teachers with less than 7 years of experience (none of them were Swedish) and one who had been teaching for 18 years. A set of final prototypes was adjusted and developed based on the third interview. Those are the results/suggestions in this paper that should help teachers make more use of their quantitative evaluation data.

2.2.1 The Academic Developer as gatekeeper

An Academic Developer (AD) at LTH selected and invited the participants based on his acquaintance with the teachers at LTH’s pedagogical institution, Genombrottet, where they participated in university pedagogical courses and some taught at them. The AD functioned as what would be defined as a gatekeeper in anthropological studies. I had informants that I would not have been able to interview had I simply advertised asking for teachers that wanted to share their experiences about evaluations. I had the credit of the AD, which meant engaged, very competent teachers regarding CEQs, evaluations in general and pedagogy. Moreover, I was introduced to the technical developer of the CEQ-system who provided me with all the relevant and requested documents.

2.3 Analysis using Engeström’s Activity Theory

I use Engeström’s Activity Theory in order to illustrate and conceptualise areas that affect the learning of a group at a working place, which in this case is LTH. The lines that form the triangles represent double arrows and indicate that the subject in the activity system (the learners) is affected by many different factors, which is the strength of the model [10].

![Activity Theory Diagram](image)

*Figure 1. Two interacting Activity Systems [10]*

The activity systems above illustrate how the two subjects, the teachers and the management, share a common objective in terms of increased learning for the students. In the index below I illustrate the difference between the various parts of the figure belonging to the different subjects:

<table>
<thead>
<tr>
<th>Subject</th>
<th>Teachers at LTH</th>
<th>Management at LTH</th>
</tr>
</thead>
<tbody>
<tr>
<td>Object (1)</td>
<td>Quality enhancement of the teaching</td>
<td>Quality assurance of the teaching</td>
</tr>
<tr>
<td>Object (2)</td>
<td>Better teaching</td>
<td>Better reputation/consolidate the good reputation of LTH</td>
</tr>
<tr>
<td>Object (3)</td>
<td>Students learn more</td>
<td>Students learn more</td>
</tr>
</tbody>
</table>
One of the suggested solutions in this paper is to increase teachers’ use of their quantitative data by adjusting the tools and signs/the mediating artefacts. The intention is to change the community among teachers from informal to more formal networks and thus from private to organisational learning. Presently, evaluations at LTH are shared and discussed through informal chats during coffee breaks with either a close colleague or an academic developer like Mattias Alveteg. By speaking openly about his own evaluations (especially the negative ones) he has tried to establish an environment among teachers at LTH and as a result several teachers have confided in him when they felt a need to discuss their evaluations [11]

3 RESULTS

Mattias Alveteg, who is a former AD at Genombrottet and who has focused on statistical analysis of the CEQ and LTH from the start in 2003 says with a reference to CEQ data, that the answer given by a student likely depends on a wide range of different factors out of which only a fraction lies under the control of e.g. the course coordinator [1]. This reflects a central problem about evaluations: that teachers are exposed to (anonymous) criticism that also their colleagues have access to, and often the comments are not related to the content of their teaching, their pedagogical abilities or how much the students learned.

Teachers only skim the working report. During the focus group interviews, it appeared to be a common pattern among the informants that they skimmed the summarising data on the front page of the working report and then went directly to the free-text answers. It was a returning issue spoken of more or less openly, that evaluations were often written in a tone that was so disagreeable that teachers had to stop reading them. It requires strength for a teacher to do anything that might attract negative comments [2]. All informants had suffered from negative responses and even though they knew that they had acted correctly, they remembered the worst ones years afterwards [12].
The informants expressed contradictory attitudes to student participation. Students have a central role in the evaluation process and whereas the two very experienced male teachers in the first focus group were very satisfied with their cooperation with SRs, the two experienced women teachers (one in each focus group) were more hesitant and especially critical about the number of inconsiderate comments that were not removed by the assistants. The group of teachers with only a few years of experience were not course coordinators and thus did not cooperate with students through the official evaluation system with the meeting between Course Coordinator, Student Representatives and Program Director. They relied on informal chats during breaks and mid-term evaluations arranged by the students, which had the form of a dialogue. Their focus was primarily on the general unpleasantness of reading the evaluations. *It is horrible every time I open the mail and I don’t have much to be afraid of, but I’m still really scared*, as one of the informants expressed it [13].

**Evaluation procedures have been changed in order to improve the psycho-social working environment.** Today, working reports are never distributed during a holiday, as it turned out that many teachers had their break ruined when reading the comments in free-text answers [13]. In addition, the summarising graph at the front page of the working report used to be a pie chart, and if the students were dissatisfied, the teachers would be facing a black circle on the front page. The pie has been replaced with a histogram where the evaluations are divided into the five main categories and LTH’s own questions about relevance and satisfaction. This gives a more balanced visualisation [11].

Some teachers choose to do their own evaluations. The CEQ in its standardised form is not considered informative to them and if there are several teachers in a course, they can’t differentiate between neither the quantitative nor the qualitative responses. One informant developed a questionnaire where she evaluated the specific learning objectives one by one [14]. Another made a questionnaire with four quantitative questions and 2 qualitative ones [13].

Students write in the free-text answers as if they are reporting to a higher authority. *I can see from the language and the form that they are writing, they don’t really expect that it is me who reads it…how they write about me is written for someone else* [13]. This is also why Informal meetings with students are preferred to the formal CEQ evaluation. Especially the informants with few years of teaching experience preferred the informal mid-term meetings arranged by the students. It was more comfortable to speak to the students in person and they could ask specific questions about the aspects of their teaching that they wanted to know about. It was also more informative to them when students during those meetings had the possibility to address them directly and thus give them a personal response [13]. Students are also more likely to use a moderate tone when physically facing the teacher in comparison to when they write comments having the freedom of anonymity.

The deselection of the CEQ might also be due to lack of understanding of the system, what it measures and that it is based on considerable research. For the informants (except Alveteg, who had analysed the statistical data comprehensively) it was the standard form that was distributed by the end of a course and where they were in doubt about what was a satisfactory result in the category *“appropriate workload”*. That is – except Alveteg, who could reveal that an average between +20 and +40 was acceptable! [12].

New teachers are left to deal with evaluations on their own. *Don’t take it too hard*, was the best advice the (young) informants would give to new colleagues that had just begun teaching [13]. Some of the experienced course coordinators had taken upon themselves to go through the evaluations with younger colleagues to make sure that they were not left alone when receiving their first student feedback, but it was a personal responsibility, not something arranged through the official channels [11].

Quantitative data becomes interesting when you can measure a progression. At the very end of the first interview an informant expressed that he would like a shadow column showing the result from last year to measure a development. This lead to a number of different prototypes: Shadow columns representing data from the previous 1-3 years; histograms from the previous year placed next to the present histogram showing the data as either numbers or percentages; columns in different colours and patterns and general discussions about, which data would be relevant to compare (e.g. the summarising histogram or the histogram representing a single question). However, it was very individual, which data the teachers saw a point in comparing and for how long a time span. This led to the idea of developing an interface where teachers could work with their quantitative data either comparing them to their previous results or to relevant averages at LTH.

**Supplying the end report with the numbers of pass/fail and the distribution of marks.** This was suggested by one of the informants, because she thought it would add perspective when you could compare the student evaluations to the results they had actually obtained [13].
In the introduction is listed a number of research-based reservations about evaluations. In the three interviews, the informants added supplementary ones based on their own experiences:

1. Good teaching is an indication, but does not say anything about what students have learned but how they experienced the situation. Teachers who demands an effort from their students are not necessarily popular [14].
2. The group dynamic at a course can influence evaluations. It is not a general trend, but something can happen in a course that reinforces the atmosphere, whether it is positive or negative [11].
3. Evaluations of a course are influenced by previous courses, e.g. if students have experienced either a very heavy workload or a very low pressure of work in previous courses, even when a teacher delivers the same teaching, the experience of the workload will vary depending on the experience of the workload at the previous course [11].
4. Students change their opinion about a course over time, e.g. when they get a job, they understand why the content of the course was useful to learn [13].

Evaluations cannot be considered as the truth about the quality of teaching and learning at a course. It is more an indicator that should be acted upon through dialogue, reassuring chit-chats during coffee breaks, self-reflection and recurring offerings of academic development when teachers feel ready for it. As Alveteg put it: *The whole result of a CEQ is the reason to ask questions about a course. The result is not an answer to a course, but rather to question the answers* [12]. However, this requires that evaluations are studied thoroughly and are not just skimmed and quickly put aside because the comments are too offensive. And it requires a systematic effort from LTH to stimulate the community among teachers and ADs and offer tools that encourage an increased use of evaluations and more openness about their content. Presently, evaluations seem to be used mainly as a “fire alarm”. If the results are exceptionally low, the administration (including programme directors) may investigate the causes but apparently it does not happen systematically [15].

### 3.1 Solutions

Based on empirical data I have developed a set of support questions for teachers that updates the present system and contains pedagogical support while keeping the process private.

*Table 1. The suggested solutions for teachers to make more use of their quantitative evaluation data*

<table>
<thead>
<tr>
<th>Suggestions</th>
<th>Aim</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Video tutorials (No video lasting more than 3 min)</strong></td>
<td>- Tutorials are aimed at both teachers and students. They can study them in private at a time convenient for them.</td>
</tr>
<tr>
<td>- What is a CEQ?</td>
<td>- Basic introduction to the CEQ and the evaluation process at LTH</td>
</tr>
<tr>
<td>- How does the CEQ system work at LTH?</td>
<td>- Teachers sharing negative as well as constructive comments intended both to give teachers support (i.e. they are not alone) and at the same time communicate to students that inconsiderate comments are offensive and useless, whereas constructive comments affect teaching in a positive way</td>
</tr>
<tr>
<td>- Different teachers sharing inconsiderate comments from their evaluations</td>
<td></td>
</tr>
<tr>
<td>- Same teachers telling about constructive comments from their evaluations, which helped them improve their teaching</td>
<td></td>
</tr>
<tr>
<td>- How to “filter” the free-text answers (a skills development for student assistants)</td>
<td></td>
</tr>
</tbody>
</table>

| **Revised CEQ-questionnaire** | |
| - Please, specify which teacher you are addressing in your comment and please, write to him/her directly using 2nd person (you) | Teachers cannot differentiate who comments are addressed to. Presently students write in 3rd person as if reporting to a higher authority. It might influence the evaluations if they have to address a teacher directly. |
Second free-text question changed from “What do you think is most in need of improvement?” to “What do you think is most in need of improvement – and why?”

The added “and why” is to encourage students to be constructive in their comments when they are reminded to clarify what their statements are based on.

Revised working report
- Eliminate satisfaction scales

Two pages of the working report consist of a statistical analysis where the summative answers to the five main categories are filtered through the responses to Q26, Overall, I am satisfied with this course. None of the informants could explain the meaning of it and since the statistical analysis was introduced 15 years ago to please the math-statisticians, I suggest that it is removed.

Revised End report
- Add the statistics about the distribution of grades and the numbers for pass/fail

Suggested by one of the informants because the students’ exam results are a quality parameter and an indicator of how much importance you should attach to the evaluations in question.

Interface to work with your quantitative data
- Possible to compare your own quantitative data to your previous data or to equivalent data at LTH

Informants wanted to compare their quantitative data to other data but it varied if it was their own data and for how many years. Or it could be averages at their own study programme or at LTH in general.

4 CONCLUSIONS

Evaluations are mandatory and thus must be undertaken, but we shall avoid the reconfirmation of the views of students who see themselves in a (passive) consumer role, which is the risk when evaluation focuses on teaching in itself [2].

It is worth remembering that it is not the evaluation system in itself that produces the positive effects. Studies have shown that improvement of the teaching is possible when the evaluation is combined with interventions such as support to teachers (e.g. individual consultations with an AD) or otherwise is used as part of an aligned set of strategies [2].

Surprisingly little research has been done about how teachers are affected by evaluations although it was a dominant and recurring issue during the interviews. Genombrottet at LTH is one of the few institutions that has studied it [16] and [17]. It reveals a pattern where teachers are strongly affected by student comments and in the worst cases remember them for years. Whereas experienced teachers have their informal networks to reassure them, new teachers do not have these networks to rely on and are generally left on their own to deal with it.

To answer the research question: How can quantitative evaluation data be made more useful to teachers in order to increase the use of evaluations to develop their teaching? A precondition for teachers to make more use of their evaluations is to create a comfort zone, where especially new teachers are offered support both professionally (e.g. supervision with an AD) and from more experienced colleagues. One of the main issues is to create tools to provide all teachers with a basic understanding of what the CEQ measures and to make it clear that they are not the only ones being exposed to negative criticism. Standard documents should be adjusted so the questionnaire prepares the ground for more constructive comments. The statistical analysis in the working report that no one seems to understand should be removed and the end report should include the numbers of pass/fail.
and the distribution of marks in order to put the evaluations into perspective. A simple interface where teachers can work in private with their quantitative data and compare them to either their own or equivalent averages at LTH could be another tool. Finally, an increased focus on qualifying student assistants as well as student representatives to pay more attention to inconsiderate comments – both in order to remove them and to remind their fellow students to be constructive in their criticism. Keeping in mind that the reason for evaluating is to get better teaching.

The suggested solutions are intended to pave the way from private to organisational learning regarding evaluations, but we must always remember that evaluations are guidelines for quality enhancement not quality assurance. There are simply too many uncertainties related to evaluations for them to be considered as more than indications. Evaluations represent the students’ perception of the teaching and learning and we cannot assume that this apprehension is shared by the teacher, which is why evaluations are only a starting point for questions to be asked rather than to be considered the nominal truth about the teaching and learning in a course.

ACKNOWLEDGEMENTS
Thanks to all informants for sharing their comprehensive knowledge and teaching experience
Karim Andersson (System Administrator)
Mattias Alveteg (Head of Department, AD)
Federico Gomez (Senior Lecturer)
Christin Lindholm (Program Director and Associate Professor)
Jan-Olle Malm (Professor)
Sandra Nilsson (Lecturer)
Charlotte Sjödell (Senior Lecturer and Vice Program Director),

REFERENCES


