Danish University Colleges

HOW TO INCREASE TEACHERS' USE OF THEIR QUANTITATIVE EVALUATION DATA
Case: The Technical Faculty of Lund University
Andersen, Line Palle

Publication date:
2018

Document Version
Post-print: The final version of the article, which has been accepted, amended and reviewed by the publisher, but without the publisher's layout.

Link to publication

Citation for published version (APA):
Abstract .................................................................................................................................................. 7
1. Introduction ....................................................................................................................................... 8
2. Methodology – Empirical Data ........................................................................................................... 11
   2.1. Participatory Design .................................................................................................................... 11
   2.2. Focus Group Interviews ............................................................................................................... 13
3. Analysis – Change Laboratory .......................................................................................................... 14
4. Charting the situation – The CEQ-system at LTH ........................................................................... 15
   4.1. The structure of the CEQ at LTH ............................................................................................... 15
   4.2. CEQ and university teaching ...................................................................................................... 17
   4.3. The evaluation process at LTH ................................................................................................... 18
   4.4. Recognising the need for change ............................................................................................... 19
5. Analysing the situation .................................................................................................................... 20
   5.1. Historicity ................................................................................................................................... 20
      5.1.1. From questionnaire through Work Report to End Report ..................................................... 21
      5.1.2. Central student roles in the development and execution of the evaluation system .......... 23
      5.1.3. Continuous Research-based studies of the CEQ-system .................................................... 24
      5.1.4. From paper to web-based evaluations ................................................................................ 25
      5.1.5. Evaluations as part of the psycho-social working environment ....................................... 25
      5.1.6. Sub-conclusion – the historicity of the CEQ-system at LTH ............................................. 26
   5.2. Empirical analysis – What are the present troubles and contradictions? ............................... 27
      5.2.1. Negative evaluations – who cares? ..................................................................................... 27
      5.2.2. New teachers are particularly vulnerable .......................................................................... 29
      5.2.3. Student participation – a double-edged sword .................................................................. 30
      5.2.4. When teachers go solo – Improvements or alternatives to the CEQ. ............................... 33
      5.2.5. Appropriate Workload – when is my evaluation positive? ................................................ 35
      5.2.6. Subconclusion – troubles and contradictions ..................................................................... 36
6. Engeström’s third generation of Activity Theory ............................................................................. 37
   6.1. Who are the subjects of learning? ................................................................................................. 38
   6.2. Why do they learn? – What makes them make the effort? ......................................................... 39
      6.2.1. Teachers’ Object: Evaluations will be presented to the management ............................... 40
6.2.2. Teachers’ Object: Quality enhancement because of vanity. The devil’s favourite sin
6.2.3. Management Object: Evaluations contributes to ensuring the reputation of LTH
6.2.4. Shared object: It makes sense to everyone that students learn as much as possible
6.2.5. Studying evaluations are closely related to looking for progression. The question is which one?

6.3. What do they learn? – What are the contents and outcomes of learning?
6.3.1. Evaluations as recurrent reminders of guidelines for learning and teaching
6.3.2. Teachers learn to survive through their informal networks

6.4. How do they learn - What are the key actions or processes of learning?
6.4.1. The Management is learning through formal communities
6.4.2. Teachers learning about student development primarily happens through informal communities
6.4.3. Writing comments for the End Report forces teachers to relate to their data

7. Creating a New Model
7.1. Adjustment of questionnaire and associated reports
7.2. Systematic AD supervision – especially for new teachers
7.3. Video tutorials aimed at teachers and/or students
7.4. Upgrading of Student representatives / Student assistants
7.5. Interface – teachers work with their own data
7.5.1. Introductory questions
7.5.2. Visualisation of quantitative data
7.6. The Dashboard-unit from explorance as a supplement to the evaluation system Blue

8. Conclusion

Bibliography
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 Work 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 is unlikely that it is a lack of understanding that prevents teachers from using a comprehensive dataset in order to improve their teaching. Neither is it a lack of pedagogical interest. The empirical data in this paper is based on focus group interviews with teachers from LTH and they have revealed that they disregard the Work Reports as soon as they can. They feel uncomfortable about the evaluations and especially about the inconsiderate comments that they often receive from students. This paper examines how to stimulate a culture in the organisation that will make teachers increase their use of the quantitative data and pave the way for expansive learning in terms of more openness and knowledge sharing regarding evaluations. 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 questionnaire, a revision of the Work 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.
1. Introduction

*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 –*  
Teacher at LTH

This thesis is about evaluations. My empirical data is collected at the Technical Faculty at Lund University (Lunds Tekniska Högskola, henceforth LTH) and consists of focus group interviews where I have spoken to teachers about evaluations and their ways of dealing with a comprehensive dataset that is orchestrated (among other things) with the intention to help teachers improve their teaching. All informants had an interest in academic development and wanted their students to learn as much as possible, but internal reviews at LTH have revealed that teachers make very little use of their evaluation data to improve their teaching and especially the quantitative feedback is ignored (Roxå & Mårtensson, 2011, p. 11).

During the interviews it appeared to be a common pattern among the informants that when they received the Work Report, which contains a summary of their evaluation data, they would only skim the summarising quantitative data on the front page and then turn to the qualitative data in terms of free-text answers. It was a returning issue spoken of more or less openly that those evaluations were often written in a tone that was so disagreeable that teachers had to stop reading them. All informants had suffered from negative responses and even though they knew that they had acted correctly, they could remember the worst comments years afterwards. Research at the Royal Institute of Technology in Stockholm confirms the observations at LTH and documents that 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 (Edström, 2008, p. 97).

There are plenty of reasons for having reservations regarding evaluation data:

- 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.
- Evaluations are teaching and teacher-focused and consequently the teacher’s attitude influences the evaluation and a feel-good factor is rewarded (Edström, 2008, p. 97).
- Student judgment is affected by class size and study level (Roxå & Mårtensson, 2011, p. 13).
• LTH’s in-house studies of the first five years of using the evaluation system (2003-2008) reveal that the education programme that a student belongs to affects the outcome of the course evaluation and there are significant differences between how students from the different programmes experience the same course (Björnsson, Dahlblom, Modig, & Sjöberg, 2009, s. 1).

• Evaluations have a tendency to be gender biased. A study at LTH revealed that teachers tended to receive higher ratings when they were teaching subjects that were considered less typical for their sex (i.e. a woman teaching hard-core physics or a man teaching environmental studies) than in subjects that were not (Price, Svensson, Borell, & Richardsson, 2017, p. 281).

• Good evaluations do not necessarily reflect how much the students learned. A teacher that makes it easy for students to pass a course without a heavy workload may be popular, but the lack of challenge made the students learn less (Edström, 2008, p. 99). The question is also if students are capable of assessing what they have learned?

Despite the fact that answers given by students are likely to depend on a wide range of different factors out of which only a fraction lies under the control of the teacher or the course coordinator (Alveteg & Svensson, 2010, s. 1), evaluations have been mandatory in Sweden since 2001 (and implemented at LTH in 2003). In other words, evaluations must be executed and as a worldly-wise saying goes: Life may not be the party that we hoped for, but while we are here we might as well dance!¹. How do we turn evaluations to our benefit, bearing in mind that the underlying intention is to improve the teaching so students learn more? The strength of evaluations is that they offer the possibility of measuring progress, which is a key aspect of higher education (Alveteg & Svensson, 2010, s. 1), and evaluations provide intelligence on how to improve the effectiveness of materials and teaching methods and to remove the elements that are not (Edström, 2008, p. 105).

The process of evaluations involves two dimensions; gathering data and then using the data for assessment and decision making with respect to agreed-on standards (Talukdar, Aspland, & Datta, 2013, s. 27). When evaluations became mandatory for all higher education in Sweden the institutions were free to decide, which system they wanted to use. At LTH they focused upon development rather than control and they chose the research-based Course Experience Questionnaire (henceforth CEQ) that measures in-depth learning. The management also communicated that neither positive nor negative results would influence a

¹ Several persons take credit for this quotation, so I have not attributed it to anyone
teacher’s salary. During the interviews, teachers confirmed that the management had kept their promise and none of them feared losing their job because of poor evaluations. They also corroborated that the official processes, where the evaluation data were being discussed, focused upon analysis and discussion rather than control and registration of the data. How come teachers practically ignore the evaluation data and especially the quantitative part, when apparently there are so few (if any) job related risks?

When I began looking for reasons for this, my pre-understanding was that it was a question of visualising the data in different and more intuitive ways. However, a closer study of the present data-visualisation revealed that it would be difficult to present the data in a clearer and more readable way. In the questionnaire students can answer a question through a five-step scale that ranges from fully disagree to fully agree. In the data-system the steps are given points from $-100$ (fully disagree) to $+100$ (fully agree).

In the Work Report the answer to a question is presented in this way:

**Clear Goals and Standards (+37)**

<table>
<thead>
<tr>
<th>Question</th>
<th>CEQ-score</th>
<th>Histogram</th>
</tr>
</thead>
<tbody>
<tr>
<td>It was easy to know the standard of work expected</td>
<td>+46</td>
<td></td>
</tr>
</tbody>
</table>

It is apparent what the question is, the number of answers to each step on the scale and the average score (+46) and I concluded that it was not the lack of understanding, which made teachers dismiss their quantitative results.

As stated in the beginning, teachers feel uncomfortable about the inappropriate comments by students that are anonymous; comments that are often not related to the content of their teaching, their
pedagogical abilities and how much the student learned. Students have the freedom to say what they want to without a matching responsibility. This feeling is intensified because evaluations are not private and teachers from the same course, student representatives and the programme director have access to them. For that very reason, said several of the teachers interviewed, looking at evaluations was simply something that they just had to get through. At the same time, the teachers made a strong effort to constantly improve their teaching, but by other means than their evaluation data. Since evaluation is mandatory and contains a richness of data, the question is how to pave the way for teachers to increase their use of especially the quantitative data, so they can also contribute to the development of the teaching? My research question is:

_How can we make the quantitative evaluation data more useful to teachers so as to increase their use of or work on evaluations?_

The suggested solutions have the overall aim of moving from private to organisational learning, but has as a starting point respect for the privacy of evaluations. Proposed solutions imply video tutorials aimed at both students and teachers. An adjustment of the questionnaire, the Work Report and the End Report and most importantly a simple interface (or dashboard) that makes it possible for teachers as well as the administration to work with the quantitative data and compare them to either their own data from previous years or equivalent averages from LTH.

2. Methodology – Empirical Data
In the research about how to make quantitative data more useful to teachers, it was clear that teachers should be an integrated part of the process. I thus took Participatory Design as a starting point using Engeström’s Change Laboratory methodology (Engeström, 2001). I gathered the data through focus group interviews that mets the principles formulated by Greenbaum and Kyng in the book Design at Work (1991).

2.1. Participatory Design
Participatory Design (henceforth: PD) came into existence in Scandinavia in the 70s when computer based technology was used for efficiency improvement and automation primarily in the production industry. Especially Scandinavian researchers criticised the traditional approach to the implementation of computer-based technology. Researchers and trade unions united and undertook projects with the declared aim to
develop workers’ competence and knowledge about computer based technology with the long-term goal to strengthen democracy in the workplace. (Christensen & Pedersen, 1999, pp. 28-29).

During the 1980s PD-research changed from the explicit political projects to focus upon “design for the skilled worker”. Focus was upon the user participation in itself and how you could give the users in an organisation influence upon the implementation of computer based technology. It was the development of the “tool-perspective”, where the technology has to be a tool for the skilled worker. It is through the highly specialised knowledge about these tools that the workers will obtain influence and democracy in the workplace. The methodology changed when designers realised that they had to use experiments in order to gain insight into the user’s “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. This change in the methodological focus was because of a change in the areas where the PD-projects were now conducted. From the production industry to office-environments and the service industry (Christensen & Pedersen, 1999, pp. 32-33).

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

1. Full participation of the users, which imply that users participate in the whole design-process and not just act as informants for the designers
2. Computer systems must strengthen the professional competences at the work place – not undermine it
3. Computer systems are tools that must be designed for the people who in the end have to use them
4. Design of computer systems cannot just aim at higher productivity, but also at quality
5. Conflicts are a part of the design process and must not be ignored, but must be dealt with and solved.
6. Design of computer systems must focus upon the actual user situation (Christensen & Pedersen, 1999, p. 34)

The aim of my research is to, through participatory design that meets the six principles, make teachers use their quantitative evaluation data to improve their teaching and thus make students learn more.
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 meaning 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 whereas the group is a means to obtain more complex data (Halkier, 2015, s. 139).

In comparison to field notes and participatory observation it is considered a weakness of the focus group interview that there are a lot of interesting things that you can only access by being present in people’s existing social context (Halkier, 2015, s. 140). 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 of informants (Halkier, 2015, s. 140). In collaboration with Academic Developer (henceforth: AD) Dr. Torgny Roxå we discussed if the selected teachers should be representative in terms of their branch of scholarship? If they should represent a negative, a neutral or a positive attitude to evaluations – or one of each? We decided to focus on teachers who had shown an interest in the evaluation process and shared a general interest in pedagogy, believing that it would reduce the number of disturbing factors, such as a general lack of willingness to use evaluations, no matter which changes were suggested.

The preliminary PD process was structured the following way:

1. Focus group interview with three teachers about the evaluation system at LTH. Due to fully booked calendars, this interview ended up being split in two with one interview with two teachers (who knew each other in advance) and one with the third teacher. The informants had a minimum of 14 years 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 discussing revised and new prototypes based on the response at the second interview. Among the informants were two younger teachers, who had
been teaching for less than 7 years and none of them were Swedish. The last informant was also Programme Director and had been teaching at LTH for 18 years.

A set of final prototypes were adjusted and developed based on the third interview.

The focus of the interview was upon the content of the discussions and not about the social interactions among the participants. During the interviews, I used the tight model for structuring and my involvement as moderator. This model can also be used explanatorily as I wanted as many different viewpoints expressed as possible (Halkier, 2015, s. 142)

3. Analysis – Change Laboratory
In order to categorise the empiricism, I used the four learning questions, which the Finnish researcher in educational sciences Yrjö Engeström claims should be a part of any theory about learning:

- **Who is learning?**
- **Why are they learning?**
- **What are they learning?**
- **How are they learning?**

Combined with the aspects; Activity system, Multivoicedness, Historicity, Contradiction and Expansive cycles it forms a matrix (see appendix 6) based on the focus group interviews (appendix 9-16).

I will be using Engeström Change Laboratory model *Strategic learning actions and corresponding contradictions in the cycle of expansive learning* (Engeström, 2001, s. 152):

![Fig 1: Strategic learning actions and corresponding contradictions in the cycle of expansive learning (Engeström, 2001, s. 152)]
I will go through the first 3 stages with an adjustment of the appellations that is compatible with the principles of participatory design:

1. Charting the situation (recognising the need for development)
2. Analysing the situation (history and present troubles and contradictions)
3. Creating a new model (Virkkunen, 2013)

4. Charting the situation – The CEQ-system at LTH
The Swedish Higher Education Ordinance 2000 requires that all students who have completed a course, should be given an opportunity to express their opinion in a course evaluation organised by the institution. However, there is no regulation on how the evaluations should be designed. Neither what purpose they should fulfil, nor how they are going to be used. Except that the results should be made available to the students. In addition, there is no clear picture of how all this data is utilised for improvement purposes (Roxå & Mårtensson, 2011, p. 62).

4.1. The structure of the CEQ at LTH
LTH has a strong tradition for academic development practiced by the department Genombrottet. Based on their recommendations the CEQ was selected as evaluation system. It 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 (Wilson, 1997, p. 33). As a performance indicator the CEQ has especially been used in Australia, where the development of their version of the CEQ has its origin in the educational research base. The antecedent was the Course Perceptions Questionnaire (CPQ). The developers Ramsden and Entwistle intended it for use to identify factors in the learning environment that influenced how students approached their learning. However, the CPQ intended to measure how students perceived the quality of teaching at a whole course level, and the CEQ provided a source of data, which enabled performance indicators to be implemented for comparisons between institutions and over time (Talukdar, Aspland, & Datta, 2013, s. 29).

The CEQ questionnaire at LTH focuses on five areas of the teaching process, which according to Ramsden have been shown to relate to quality in student learning:

1. Appropriate workload (is the workload manageable?)
2. Appropriate assessment (is the examination supporting understanding?)
3. Generic skills (is the development of generic skills supported?)
4. Good teaching (is there support and encouragement from lecturers?)
5. Clear goals (do lecturers make an effort to help students understand what they are supposed to learn?)

Additional scales have been developed recognising that there are many factors beyond the classroom or the teacher that can have an effect on learning. E.g. the quality of the university’s learning resources and infrastructure, the student support and administrative systems, and the role played by formal and informal social contexts in which learning takes place (Lindholm, Gomez, & Nilsson, 3rd Focus Group Interview, 2018, p. 30).

The standard questionnaire at LTH contains 26 quantitative questions, where question 17 and 26 are added by Genombrottet;

Q17: The course seems important to my education
Q26: Overall, I am satisfied with this course

The different categories are mixed to obtain a more truthful response by the students, just as some questions aim at a negative reply, i.e.: Q13: It was often hard for me to discover what was expected of me in this course. It is the intention that the variety will keep the student more attentive and thus give a reply that they have been giving more thought (Alveteg, Malm, & Sjödell, 2018, p. 3).

Besides the two quantitative questions (Q17+Q26) LTH has added two qualitative questions:

1. What do you think was the best thing about this course?
2. What do you think is most in need of improvement?
4.2. CEQ and university teaching

The full CEQ-questionnaire can be seen in Appendix 1. Below is a table that illustrates how the five main categories are distributed in the questionnaire and their overall pedagogical aim. The questions in bold aim at a negative reply:

<table>
<thead>
<tr>
<th>No.</th>
<th>Clear Goals</th>
<th>Good Teaching</th>
<th>Appropriate Workload</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1, Q6, Q13, Q25 E.g.: Q6: I usually had a clear idea of where I was going and what was expected of me in this course</td>
<td>Q3, Q7, Q15, Q18, Q19, Q21 E.g.: Q19: My lecturers were extremely good at explaining things</td>
<td>Q4, Q14, Q22, Q24 E.g.: Q4: There was a lot of pressure on me as a student in this course</td>
<td></td>
</tr>
<tr>
<td></td>
<td>This scale indicates if the students regard that the teachers have established clear goals and assessment criteria. This implies aspects like the students possibility to know which quality is expected of their work, their ability to know what is expected, how easy it is to understand as well as clearly defined aims and goals. It is not sufficient to refer to course schedules. Goals have to be communicated, specified and exemplified by the teacher. Clear Goals is a separate category despite that it is among the aspects of Good Teaching, but research has shown that Good Teaching can achieve a high score despite goals and expectations being unclear (Borell, 2008, s. 15-16)</td>
<td>The scale consists of questions about how much feedback the students get. The teachers’ ability to explain things, make the subject interesting, motivate the students and understand their problems with the course. However, good teaching cannot be exactly and indefinitely defined. Reality requires a pragmatic prioritisation based on well informed reflection and it is demonstrated how the CEQ can contribute to that when they focus upon the variables that in the course are critical for the control of the teaching and through that the quality of the results (Borell, 2008, s. 15)</td>
<td>This scale is based on the assumption that overloading the students with work can have a negative effect on the quality of their learning, because it can spoil their attempts to in-depth learning and instead focus upon surface learning such as memorising.</td>
</tr>
</tbody>
</table>
### Appropriate Exam

| Q8, Q12, Q16, Q20 | As students often see the exam as defining the real goals of the course, they tend to prepare themselves in order to pass the exam. It is important that there is an accordance between the exam and the learning objectives of the course. Otherwise you risk that students establish a poor understanding with fragmented knowledge of only a part of what the course is supposed to cover (Borell, 2008, s. 16). |
| Q20: Too much of the assessment was just about facts |

### General Skills

| Q2, Q5, Q9, Q10, Q11, Q23 | This scale is designed to measure to what extent the course develops a number of general academic skills and knowledge necessary for the graduates in order to act as professionals. These skills imply problem solving, analytical abilities, group work, confidence in managing unfamiliar problems, the ability to plan one’s work and written communication (Borell, 2008, s. 17) |
| Q11: The course has improved my skills in written communication |

Rixon and Ramsden, who developed the CEQ, recommended that evaluation results became a reason to ask questions rather than to consider them as conclusions and that all comparisons and assessments implies well-informed considerations (Borell, 2008, s. 13). LTH ran a follow-up study in 2006 three years after the 2000 Higher Education Ordinance Regulation came into effect. It revealed that student evaluations were introduced in most courses, but the use of the results was defective and so was the communication to the students about actions taken as a result of the course evaluations (Roxå & Mårtensson, 2011, p. 63). Today LTH has organised a standard evaluation process with a focus upon analysing and discussing the results, which is demonstrated in the 7-step process below. As it appears students play a central role in the process.

#### 4.3. The evaluation process at LTH

The evaluation process at LTH runs in seven steps:

1. Teachers are asked if the survey has to be paper og web based and if they would like to add 1-4 supplementary free-text questions to the questionnaire
2. Students fill in the questionnaire (paper og 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 Work Report is discussed as well as how the course was implemented and how to develop it in the future. The meeting is obligatory and the participants are the course coordinator (CC), student representatives (SR) and a Programme Director (PD), who is responsible for the whole programme, of which the course is part.

6. The End Report is created containing the statistically processed data (now including the number of pass/fail) and the three concluding remarks about actions needed to be taken from CC, SR and PD respectively.

7. The End Report is published on the faculty web (available behind log-in) and sent via e-mail to all students who participated in the course (Roxå & Mårtensson, 2011, p. 66)

4.4. Recognising the need for change

Ramsden claims that the most significant and challenging aspects of evaluation comprise interpretation of results and the action that follows to improve teaching\(^2\) (Borell, 2008, s. 12). It has been the intention at LTH since the implementation of the CEQ in 2003 that emphasis should be upon quality enhancement (i.e. analysis and matching actions to improve teaching) rather than quality assurance (i.e. control based on the data) (Andersson, 2018, p. 1).

The administration at LTH has acted on their word and never used evaluations as an instrument of control, but as a tool for development. Students have been integrated in the process, teachers and course coordinators’ arguments have been listened to and evaluations have never been used as a reason for salary increase, or – worse – an excuse for dismissal. However, despite the best of intentions, evaluations are still a delicate matter. During the interviews I never explicitly asked the informants how they felt about studying their evaluations, but it was nonetheless a returning issue through all the interviews how badly they often felt when reading the students’ feedback. When an informant told about a negative evaluation comment from 2003 another informant complemented: We sit here and even though we laugh, you still remember it.

\(^2\) Unless otherwise stated it is my translation of titles or quotations that are not written in English
And you remember exactly what year it was. And I know people, who have been devastated and as a result no longer want to read their CEQs (Alveteg, Malm, & Sjödell, 2018, p. 9)

The informants kept returning to their difficulties about reading the free-text answers and apart from one informant who had been working thoroughly with statistical analysis, related to the CEQ, the rest skimmed their Work Report and practically ignored 2/3 of it. *I look at the category Good Teaching and Q26 about satisfaction and then I go to the free text. Then I look a bit at the standard deviation and some of the other figures at the front graph. Then I’m done* (Lindholm, Gomez, & Nilsson, 2018, p. 9). Bearing in mind that I only interviewed teachers with an interest in teaching and academic development, i.e. teachers that wanted to improve their teaching, so students could learn more; it is significant that they did not use evaluations as a volume of information for developing their teaching.

One of the informants was Head of Department and former AD and for several years he had tried to establish a more open dialogue about evaluations, but so far without prevalent success. He referred to a colleague, who used to say that *asking a teacher about their evaluations is equivalent to asking about their personal hygiene!* (Alveteg & Malm, 2018, p. 11). Evaluations are private and something that you only share with colleagues you rely on.

In the following chapter through the historicity of implementing the CEQ-system at LTH and present troubles and contradictions in the evaluation process, I will study what prevents teachers from using their quantitative evaluation data and which steps can be taken to change these habits. Steps which based on the interviews and research about evaluations are the most likely to succeed.

5. Analysing the situation

5.1. Historicity

Historicity is the third principle in Engeström’s Activity Theory. *Activity systems take shape and get transformed over lengthy periods of time and their problems and potentials can only be understood against their own history* (Engeström, 2001, s. 136).

When evaluations at higher education became mandatory in Sweden, the institutions had the autonomy to choose, which evaluation system they wanted to use and the procedures that followed them. They selected the CEQ-system and the intention was to make the working process flexible for teachers and not to waste time on completing questionnaires and writing reports (Andersson, 2018, p. 1). The purpose of the student
feedback on teaching and course evaluation at LTH is formulated in a policy document *This policy describing the system of evaluation of undergraduate education at LTH shall contribute to a process where the quality of teaching is consciously and systematically enhanced*. To achieve this, student feedback at LTH is collected for operational and reporting purposes respectively. The operational purpose (the formative evaluation) refers to the feedback that a teacher can organise during a course in order to gain a better understanding of the learning of the students and then adjust the teaching accordingly. The reporting purpose (the summative evaluation) is the data collected by the end of the course in order to produce a document describing the quality of the course. The purpose of this document is to enhance the dialogue between the programme boards, the department and the students (Roxå & Mårtensson, 2011, p. 65).

5.1.1. From questionnaire through Work Report to End Report

A teacher from LTH suggested the structure of the *Work Report*. The system administrator, who is still in charge of the CEQ-system, adjusted it, and the form has remained almost unchanged since the beginning (Andersson, 2018, p. 2).

The Work Report is computer generated based on the results from the questionnaire and has the following structure (please, see appendix 2 for a full Work Report):

<table>
<thead>
<tr>
<th>Page</th>
<th>Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>Front Page</td>
<td>Basic facts to identify the course</td>
</tr>
<tr>
<td></td>
<td>Presence at teaching (self-reported by students)</td>
</tr>
<tr>
<td></td>
<td>Average points and standard deviation from the five main categories + Q17 and 26</td>
</tr>
<tr>
<td></td>
<td>Summarising histogram of the main categories (except General Skills) + Q17 and 26</td>
</tr>
</tbody>
</table>

**CEQ-scales & special questions**

<table>
<thead>
<tr>
<th>Scale</th>
<th>CEQ-score</th>
<th>StdDev</th>
</tr>
</thead>
<tbody>
<tr>
<td>Good Teaching</td>
<td>+31</td>
<td>38</td>
</tr>
<tr>
<td>Clear Goals and Standards</td>
<td>+37</td>
<td>35</td>
</tr>
<tr>
<td>Appropriate Assessment</td>
<td>+61</td>
<td>39</td>
</tr>
<tr>
<td>Appropriate Workload</td>
<td>+10</td>
<td>44</td>
</tr>
<tr>
<td>Generic Skills</td>
<td>+28</td>
<td>33</td>
</tr>
</tbody>
</table>

**Special questions**

| The course seems important for my educ. | +82 | 34     |
| Overall, I am satisfied with this course | +61 | 39     |
As it appears, statistical data are dominating the report, which is due to mathematicians-statisticians strongly objecting to the whole idea of having evaluations about pedagogy. They considered it against nature (!) to have reports that did not comprehend statistics, which is why the standard deviations and the summary scales have been included (Andersson, 2018, p. 4).
As a starting point, SRs and teachers in a course have access to the Work Report of the given course. The Head of Department has access to all Work Reports and the system administrator keeps a summarising excel-sheet of all the quantitative data (see appendix 3) – a modified version of the excel sheet is available with a delay of one year at ceq.lth.se (see: http://www.ceq.lth.se/specialrapporter/).

The End Report (see appendix 4) has the same front page as the Work Report including the summarising histograms of Q17 and Q26, concluding remarks by SR, CC and PDs about the teaching and future actions to improve the points of dissatisfaction. When administration is monitoring the CEQ-data it is one of the central elements that all representatives have written a comment in the report though it is not always the case (see appendix 3). All End Reports are published at ceq.lth.se and accessible to all staff at LTH with a log-in (Andersson, 2018, pp. 6-7).

5.1.2. Central student roles in the development and execution of the evaluation system
One of the central decisions was that students were given considerable influence upon the evaluation process right from the start. When a working group was established in order to lay the foundation of the coming evaluation system, it included a student representative (Karim Andersson), who later went on to be the one to programme the actual system. Andersson studied Civil engineering, but never completed his thesis, as he was employed at LTH together with another student assistant (Jonas Borell) who studied Psychology and had considerable understanding of pedagogy. Borell worked with the questionnaire and how to present the information whereas Andersson attended to the programming part (Andersson, 2018, p. 1).

It has been an explanatory statement that no one else at LTH had the time (Andersson, 2018, p. 1), and it reflects that students have played an important role in the development and execution of evaluations at LTH. Paid student assistants are filtering the free-text answers in order to remove unpleasant comments that serve no purpose in terms of improving the teaching. Comparing to a Danish context raw evaluation data are considered confidential material that students do not have access to. Elected SRs take an active part in the evaluation process, and interviews have revealed that students in general demand (or expect to have) influence during a course. Students come, and if there is a disaster, I have this group of students outside my office, as an informant expressed it (Lindholm, Gomez, & Nilsson, 2018, p. 11).

Information from supervisor Lars Birch Andreasen, Aalborg University
5.1.3. Continuous Research-based studies of the CEQ-system

By autumn 2003 the CEQ system was introduced at LTH after some preliminary test rounds. Over the years suggested modifications have generally been rejected by the administration justified because the CEQ is research-based and because the administration now possesses very long data-sets ready to be analysed that they would prefer not to interrupt (Andersson, 2018, pp. 1-2).

The CEQ-system was monitored from the start including quantitative evaluations of how the system worked:

- When the evaluation changed from paper to being mainly web-based they tested if students evaluated differently depending on the media. (No difference could be registered) (Alveteg & Malm, 2018, p. 3)
- Did evaluations differ in terms of satisfaction, if students completed the questionnaire immediately or after the first or second reminder? (No difference could be detected) (Alveteg & Malm, 2018, p. 3)
- Study of a possible compliance between the quantitative and the qualitative response. An academic developer divided the free-text answers into groups that corresponded to the quantitative questions and demonstrated an accordance in the pattern of the datasets (Alveteg & Malm, 2018, p. 9).

Among others the CEQ-datasets have resulted in the following publications:

- *The evaluation system CEQ at LTH: Is the intended aim fulfilled?* (Björnsson, Dahlblom, Modig, & Sjöberg, 2009)
- *University Teachers’ experience of course evaluations – Emotions and consequences* (Roxå & Bergström, 2011)
- *Systematic Course evaluations – Academic Teachers’ experiences and the organisation’s ability to develop* (Roxå & Bergström, Kursvärderingar i system - akademiska lärarens upplevelser och organisationens förmåga till utveckling, 2013)

Moreover, researchers at LTH have carried out several gender studies based on CEQ data. This implies:

- *The Role of Gender in Students’ Ratings of Teaching Quality in Computer Science and Environmental Engineering* (Price, Svensson, Borell, & Richardsson, 2017)
Do women perform better than men at LTH, or the contrary, or neither, or both? (Hell, 2013)\(^4\)

The general interest in gender-related aspects results in returning statistical analysis with a gender focus. This has recently led to a disclosure of a #MeToo-teacher because of a follow-up action when statistics revealed that male students consequently evaluated the teacher better than the female students (Sjödell, 2018, p. 3).

5.1.4. From paper to web-based evaluations

From 2006 LTH changed from paper to web-based evaluations. Teachers can still request a paper-based evaluation, but as a starting point, questionnaires are completed electronically (Andersson, 2018, p. 2). Considering that LTH conducts approximately 1000 courses per semester, it resulted in a sizeable reduction in terms of administration, paper and waiting time. A negative consequence was a reduced response frequency, but from a cost-benefit perspective, the reduction in administration compensated for this (Alveteg & Malm, 2018, p. 2).

However, paper-based evaluations have a specific advantage because you can study the connection between the quantitative and the qualitative responses on an individual basis. In the web-based form the free text is assembled, e.g. it does not appear if the five items of complaint come from the same student. This is one of the reasons that some teachers still stick to the paper-based form or carry out their own paper-based evaluations (Sjödell, 2018, p. 2) and (Lindholm, Gomez, & Nilsson, 2018, pp. 2-3).

5.1.5. Evaluations as part of the psycho-social working environment

During the past 15 years there has been an awareness at LTH of evaluations’ possible influence upon the psycho-social working environment among teachers, e.g. (Roxå & Bergström, 2011) and (Roxå & Bergström, 2013). When it turned out that many teachers had their holiday ruined because of negative free text responses, the publication of the Work Reports was postponed until after the break, (Lindholm, Gomez, & Nilsson, 2018, p. 16). The front page of the Work Report is now dominated by a summarising histogram, which in 2011 replaced a pie chart that illustrated the response to Q26 (Overall, I am satisfied with this course) divided into the three categories dissatisfied, neutral, satisfied.

\(^4\) On the Usefulness of Course Evaluation Data in Quality Assurance and The Role of Gender in Students’ Ratings... are the original titles. The other titles are my translation
If all students were dissatisfied, the teacher would be facing a black diagram at the very first page, which left many teachers unnecessarily offended. A working group was established with the indicative name “Operation Schwarzwald Cake” that suggested the summarising histogram as a more diplomatic illustration (Alveteg & Malm, 2018, p. 8). In general, through its staff policy LTH has demonstrated an extensive consideration for teachers’ feelings regarding evaluations. Nevertheless, it was a recurrent issue in all interviews that teachers were extremely affected by evaluations and especially the free text answers. I consider it one of the main reasons that teachers do not study the quantitative data. They disclaim evaluations as soon as they can and instead rely on the information they obtain through immediate reactions from the students during class and informal dialogues with their colleagues and students during breaks and feedback sessions.

5.1.6. Sub-conclusion – the historicity of the CEQ-system at LTH
The overall aim of the evaluation system is to enhance teaching and student learning, but a local study at LTH (Roxå & Mårtensson, 2011) revealed that the aim had not been received and integrated by a large section of teachers within the faculty. They criticised the system for not being sensitive to disciplinary differences and for having a hidden agenda of controlling teachers at the time when the system was implemented (Roxå & Mårtensson, 2011, p. 67). I cannot recognise this pattern when I study the historicity of the CEQ-system at LTH. It generally reveals an extensive consideration for the teachers and for respecting their wishes. When the system was developed, the Work Report where the CEQ-data was assembled, had an overwhelming focus on statistics to comply with the mathematicians-statisticians who had particular reservations about the evaluation system. When the pie-chart that visualised student satisfaction on the front page of the Work Report turned out to arouse negative feelings, it was considered a psycho-social working environment problem and the pie-charts were replaced with a more balanced histogram. Continuous research based on CEQ-data substantiates the professional standards of the system. The management has given the students a democratic voice in the process. Generally, the implementation
and consolidation of the evaluation system reveals a willingness to listen to employees and students and try to meet their requests. Changing from a paper- to a web-based system can be argued was a top-down decision. A decision that can be justified by the time-consuming administration that was saved, when replies no longer had to be registered by hand. It reduced the response frequency, and teachers lost the possibilities of seeing the connection between quantitative and qualitative replies on an individual basis. However, teachers still have the opportunity to execute the evaluations on paper so that wish was also accommodated by the management.

5.2. Empirical analysis – What are the present troubles and contradictions?
The fourth principle of Activity Theory is the central role of contradictions as sources of change and development. They are not the same as problems or conflicts. Contradictions are historically accumulating structural tensions within and between activities (Engeström, 2001, s. 137).

Some people are more difficult to please than others, as an informant sarcastically said, when he referred to a lecture given by an external teacher, who had begun by asking the students, if they knew The Planning and Building Act? When no one said anything, he gave a summary of the law, but several students wrote in their evaluation that the teacher had wasted their time since they had been taught The Planning and Building Act in the previous semester! (Alveteg, Malm, & Sjödell, 2018, p. 17). Through the interviews I have identified the following troubles and contradictions related to the summative evaluations at LTH, which will be discussed and clarified in the subsequent chapters. The topics are:

1. The consequences of negative evaluations.
2. New teachers’ special challenges with exposure to student feedback.
3. Problems related to student participation.
4. Teachers designing their own evaluations
5. How to interpret the category Appropriate Workload.

Again - being affected by negative student feedback was a returning issue when the different subjects were discussed.

5.2.1. Negative evaluations – who cares?
Several informants articulated that they did not know enough about the CEQ-system in itself. Some were not even aware that it was research-based (Lindholm, Gomez, & Nilsson, 2018, p. 9). Others had difficulties
seeing the point of distributing the same questionnaire repeatedly when it did not seem to be used for any statistical analysis, which should be the reason for reiterating it:

Informant I: These CEQ-questions are a small extract of the original CEQ, which is far more comprehensive. There are so many questions that it’s getting too much. In Australia, they use it at many educational institutions, but in which form and which context we don’t know.

Informant II: That’s interesting, because if I could choose, I would probably select some of the other questions, because I feel that LTH is demanding standard evaluations so they can compile the statistics. Then it is important that we get the same question in order to be able to compare, but we don’t really know what we want to use the statistics for⁵ (Alveteg, Malm, & Sjödell, 2018, p. 11).

None of the informants had been confronted by the management due to negative evaluations (or at least no one revealed so). This could be because I was interviewing teachers that were very dedicated to their teaching and all had a strong interest in pedagogy. Less enthusiastic colleagues might have experienced otherwise, but there definitely did not exist an official procedure about evaluations. It appeared more like a fire alarm than systematic control:

The system administrator had an excel-sheet (appendix 3) with all the quantitative data from the CEQ sorted in descending order according to student satisfaction (Q26). It appeared that one teacher had a score of ±90 (out of ±100) and even managed to achieve the same impressive average the following semester! The system administrator answered evasively when I asked if teachers knew that this type of document existed and if he would report such a negative result. I assume that the system administrator despite (or maybe because of) his reluctance to answer my question gives a hint to the relevant people (e.g. the Head of Department or Programme Director), but the absence of an official procedure underlines the feeling of a fire alarm. In the meantime I have found the excel-sheet at ceq.lth.se (see: https://www.ceq.lth.se/specialrapporter/ - CEQ-data för samtliga läsår), but despite the document being available, you have to know what you are looking for (e.g. the course code) in order to be able to deduct information about teacher performance.

⁵ The interviews with Mattias Alveteg, Jan-Olle Malm, Charlotte Sjödell and Karim Andersson were all conducted in Danish/Swedish. All quotations from these interviews are my translations. The interview with Federico Gomez, Christin Lindholm and Sandra Nilsson was conducted in English and quotations are original.
Despite the fact that informants with a management position would read all the Work Reports related to their department or programme, a lot of insights about evaluations are apparently heard through the grapevine (Alveteg, Malm, & Sjödell, 2018, p. 17) and dealt with through alternative channels:

*I talk about (my) evaluations in the coffee room. When something has grated a bit I try to talk about it with the teacher, but there are also course coordinators who come to me on their own initiative, and then we try to talk about it. As Head of Department I get all course evaluations. I go through them and if I have heard rumours I look specifically for something, but I have never had to go to see somebody because of bad evaluations* (Alveteg & Malm, 2018, p. 12).

Considering the different types of monitoring and the lack of official procedures called forth because of negative evaluations, it is difficult to say, if continuous negative feedback has any other consequences than optional courses closing, because no one wants to attend (Andersson, 2018, p. 7). However, it is a possibility that all the unofficial procedures are in fact the best way of dealing with evaluations if you wish to obtain the overall aim of improving the teaching. When you have the frame of an official evaluation system, then small talk in the corridors and chit-chats with colleagues and the students about the methods and content of the teaching can be the most efficient tools. Just as a systematic procedure where negative evaluations automatically lead to AD supervision or an invitation to the manager’s office will be considered stigmatising and as a result have a negative impact on the teaching, regardless of the intentions.

5.2.2. New teachers are particularly vulnerable

Experienced teachers appear to be more relaxed or less affected by student comments, but admit that even after many years of teaching and corresponding evaluations, negative remarks can still ruin their weekend (Alveteg & Malm, 2018, p. 12). *I suppose that I’ve read most of what can be said, but for younger and less experienced teachers it is more difficult. In addition, because it’s a document that circulates a bit, it becomes even more important that inappropriate responses are removed* (Alveteg & Malm, 2018, pp. 4-5).

Expressed by a teacher, who has taken it upon himself to remove inappropriate comments, before the student feedback is passed on to new colleagues. Just as he makes sure that he goes through the evaluations with them, so they are not left on their own to deal with them. Again, this is not arranged through official channels, but a personal responsibility that he has assumed.
Don’t take it too hard, was the best advice that (young) informants would give to colleagues that had just begun teaching, and they both said that they had not received any support when it came to relating to evaluations (Lindholm, Gomez, & Nilsson, 2018, pp. 15-16).

Whereas experienced teachers working for several years in the same institution, usually have a network to reassure them, when the negative feedback is overwhelming (Alveteg & Malm, 2018, p. 12), teachers with only a few years of experience in the same workplace do not have these networks, and may not even know the value of them yet. Systematic AD supervision offered to new teachers should be able to compensate for the missing network and it will not be stigmatising as long as it is done consistently and not only because of negative feedback.

5.2.3. Student participation – a double-edged sword
Whereas all informants often engaged in dialogue with students about their teaching with the intention to improve it, their attitudes differed regarding student participation and especially the extensive influence that students have at LTH.

From the very beginning, students were given a central role in the implementation of the evaluation system. Student assistants had access to raw evaluation data when filtering the free-text comments. Student representatives had access to Work Reports and participated in the CEQ-meetings along with course coordinators and Programme Directors. Moreover, student assistants programmed and developed the pedagogical presentation of the data. Even though they were student assistants, who knew the institution as they had been working there for several years, had special skills through their educational background and as student representatives had been engaged in different working groups, they were nonetheless student assistants. One could argue that it was a challenge without the responsibility that comes with being a permanent member of staff.

5.2.3.1. Upgrading of student skills is required
Whereas nobody problematises that student assistants have access to raw evaluation data that can contain very inappropriate comments about teachers that they know, informants’ opinions about students’ abilities to filter out unsuitable negative feedback differed. Some informants acknowledge that they do it to take care of the teachers (Alveteg & Malm, 2018, p. 4), but others pay attention to how much actually slips through (Alveteg, Malm, & Sjödell, 2018, p. 9) and (Lindholm, Gomez, & Nilsson, 2018, p. 6). I have had to call the system administrator a couple of times and told them that they are too careless in terms of the text.
that gets through, says one of the informants and suggests that the instructions that the assistants presently get from the system administrator are supplemented with specific examples of unacceptable comments. This will give the assistants an indicator of the kind of text that should be removed (Alveteg, Malm, & Sjödell, 2018, p. 12)

An upgrading of student assistant skills was a recurrent suggestion – often in connection with the presentation of prototypes that I had aimed at teachers. When I presented the prototypes for two video tutorials “Be specific and constructive”, the first video contained examples of inappropriate comments shared by the teachers from LTH who had experienced them. The second video had an identical set up, but showing constructive examples (see Appendix 5). Especially the first video was intended for teachers with the underlying message “you are not the only one”, but an informant immediately suggested that this would be a good tutorial for all students. For student assistants it was a reminder of what they had to delete and it could be sent with the questionnaire as a reminder to all students to use a proper tone (Lindholm, Gomez, & Nilsson, 2018, p. 7).

Another informant wanted to upgrade the student representatives. She pointed out that they needed to become aware of proportions; They look at lot at the free-text and consider it the truth. They don’t take into consideration that it may only be three comments and that there are 200 students in the class (Lindholm, Gomez, & Nilsson, 2018, p. 7). Other informants expressed satisfaction with their cooperation with student representatives – especially at the CEQ-meetings, where they could clarify points of complaints and explain that this was a widespread feeling in the class – or the contrary that the specific remark only represented a few of the students. When representatives had been elected, there was also the possibility that they could carry out the mid-term evaluation and that there would be a continuous dialogue with them during the course. This made it possible to identify contingent problems while the course took place (Alveteg & Malm, 2018, p. 5).

5.2.3.2. The experience of student participation might depend on gender, experience, status and/or nationality

However, the informants that were most content with the cooperation with student representatives were both very experienced teachers. One was professor, the other Head of Department. They were men and they were Swedish. I cannot generalise based on six informants, but it is a possibility that they by virtue of their age, experience and nationality are being treated more respectfully than teachers with fewer years of
experience. Especially if they are also from another country, which was the case for two of the informants. Statistical analysis of the CEQ at LTH (Price, Svensson, Borell, & Richardsson, 2017) cannot substantiate my statement, but I nonetheless would like to accentuate the recurrent frustration expressed by the other four informants about the hardness in the comments and the need to tighten up for all students (Lindholm, Gomez, & Nilsson, 2018, pp. 6-7).

Though the attitude to student participation in the evaluation process varied, all informants expressed a genuine interest in seeing their students develop. All referred to the more informal dialogues as more educational in terms of improving their teaching: Talking about the informal chat, sometimes I do that, because my course is the last one before the master-thesis, so some of the students are working in the labs and continue after the teaching...and if they have passed the course, sometimes they feel more free to give me their personal opinion. Sometimes we have nice chats about what was more rewarding both when it comes to happy and unhappy people (Lindholm, Gomez, & Nilsson, 2018, p. 5). Another teacher: I talk to students during breaks. They come and ask for help. And they ask questions. I talk to teaching assistants, so I mean the information gets to me in one way or another. I try to talk to students about the course (Lindholm, Gomez, & Nilsson, 2018, p. 5).

5.2.3.3. Communicating through informal dialogues is less binding
Teachers’ hesitation towards student participation concern the fact that student assistants often are not doing their job properly (i.e. too many inappropriate comments slip through). Secondly, that students’ influence does not come with a matching responsibility. They are free to criticise (often anonymously), representatives participate in CEQ-meetings and their comments are registered in the End Report that’s public. However, nothing happens if they do not show up at the meetings, do not write any comments or do not make an effort to increase the response frequency among their fellow students (Alveteg & Malm, 2018, p. 6). Informants generally preferred the informal dialogues, which seemed to have a much bigger imprint on their teaching than the official dialogues. One can argue that informal dialogues are less binding, i.e. teachers are free to select the suggestions and comments that they agree with compared to official evaluations where the suggested actions in the End Report should preferably be executed. It is understandable that teachers prefer the type of communication where they are free to choose rather than official meetings where they have less influence, but nonetheless must follow the decisions passed.
5.2.4. When teachers go solo – Improvements or alternatives to the CEQ

It was a recurrent issue in the interviews that the CEQ might be too general to be informative for a course. Is it the standard that rules rather than what is best for the programme? However, LTH encourages interdisciplinary courses and in general that students can follow courses outside their programme. As a consequence, there are usually students from different programmes in the same course and it is important that everybody completes identical questionnaires (Alveteg, Malm, & Sjödell, 2018, p. 11).

5.2.4.1. Students are reporting to the teacher’s superior

Evaluations are mandatory for courses with more than 25\(^6\) students – and optional for the ones that have fewer. Students are requested to specify to what extent you have participated in the various course activities (0\%, 20\%, 40\%, 60\%, 80\% or 100\%) and the quantitative questions have teaching or the teacher(s) as their starting point. They are formulated so students are forced to reflect upon their own learning (e.g. Q21: The teachers on the course worked hard to make the subject interesting). However, the free text questions set the stage for reporting in third person to a higher court rather than in second person addressing the teacher who prepared and taught at the course:

Informant III: I can see from the language and the form that they are writing that they really don’t expect that it is me who reads it. Because of how they write about me it is written for someone else, so maybe the students could be made aware of that it is actually the teacher reading it.

Informant IV: …students write to the board or to the student representatives. I agree with that (Lindholm, Gomez, & Nilsson, 2018, p. 7).

The above dialogues reflect the problem of writing about the teacher in the third person, which can affect the tone as it is more likely that students would modify their strictures against a teacher if they were addressing them directly. Moreover, it indicates a second problem: if there is more than one teacher in the course, students rarely specify to whom they are referring.

\(^6\) During interviews with the system administrator and informants that are involved in the administration of the CEQ-process the limit of 25 students is consistently being used. However, official documents at ceq.lth.se consequently write 30 students. I have mailed the System Administrator, but not received an answer, so in the meantime I stick to the 25 students, but am aware that it could be 30.
5.2.4.2. Teachers make their own evaluations
When teachers cannot deduce the required information from the evaluations that they want to, they either stick to the informal dialogues during class (Lindholm, Gomez, & Nilsson, 2018, p. 5), or they make their own evaluations: If the CEQ is for the students, we cannot differentiate which answers come from my students about my part of the course, so I send them my own evaluations. Then I summarise their evaluations and send them to the other teachers and to the board .... Now we have separated the course, so from this year I'll have my own course and my own CEQ and then I can look at the CEQ, because I will know that it is only my students that have been giving feedback (Lindholm, Gomez, & Nilsson, 2018, p. 2).

Another informant had made her own questionnaire by combining some of the learning objectives of the course with specific questions from the CEQ, the reason being that she could use the results to a greater extent than the replies from the standard CEQ (Alveteg, Malm, & Sjödell, 2018, p. 10). It almost caused an argument during the interview as one of the other informants in his capacity as academic developer had had to act upon the courses, where no students had completed the CEQ. Students cannot be bothered to complete two questionnaires, and in situations when a teacher has his or her own, it is the CEQ that is deselected by the students (Alveteg, Malm, & Sjödell, 2018, p. 10). From a programme perspective, it upsets the strategy of uninterrupted datasets of the CEQ for all courses at LTH with more than 25 students.

5.2.4.3. The possibility of adding 1-4 free text questions
Course coordinators receive an e-mail a couple of weeks before the course ends reminding them that they can add 1-4 extra free text questions. One informant had made use of that and asked very specifically to how students had improved:

1. Their creative skills
2. Their ability to communicate orally
3. Their ability to communicate visually
4. Had they had sufficient access to the facilities at the school in order to be able to complete the course elements (Sjödell, 2018, p. 2)

She expressed the problem of having to resend the same questions every semester and not being able to keep them as a permanent part of the CEQ-evaluation for her course (Sjödell, 2018, p. 2).
None of the other teachers had made use of the possibility of adding extra questions. One had encouraged students to write about specific issues in the free-text section (Alveteg & Malm, 2018, p. 4), but in general, it appeared that emergent questions were better clarified immediately by asking the students in class.

5.2.4.4. Should students be reminded about the learning objectives?

Informat II: it is a fact that the present CEQ-evaluation is not combined with the learning objectives, and in that case the CEQ doesn’t help anything. The internal evaluations do, because they are goal oriented. The CEQ has a number of general questions, but nothing about whether specific objectives have been met. It can even be that some students believe that certain goals have been fulfilled, but they don’t know, because they can’t refer to the course description.

Informat I: This is one of the points of criticism about the validity of the questionnaire to ask the students if the learning objects have been met, because a correct answer implies that students are aware of what these are, which is not always the case (Alveteg, Malm, & Sjödell, 2018, p. 10).

Presently it is only the aim of the course that figures at the head of the questionnaire and there was a debate if students should be reminded of the learning objectives of the course and the usefulness of evaluating according to those (Alveteg, Malm, & Sjödell, 2018, pp. 10-11). Adding the learning objectives to the questionnaire (e.g. students are presented to the objectives before they begin to complete the questionnaire), does not imply that students actually read them.

I perceived the discussion between the two informants as an illustration of the disciplinary differences between teaching a subject and a subject serving as a tool for further study. Learning objectives concentrate on competences and thus it was in the interest of the informant who was a teacher from Industrial Design to have increased focus upon the tools or competences she was teaching. The first informant was a teacher from the Institute of Chemistry, who had more interest in the subject on its own and thus attached less importance to the learning objectives.

5.2.5. Appropriate Workload – when is my evaluation positive?

The reason that the students are asked about their workload is based on the knowledge that if the work pressure is too heavy, students stop studying in order to understand and instead memorise the subject-matter. Except for one informant everyone could agree that they were in doubt about how to interpret the category “Appropriate workload” – in Swedish Lämplig arbetsbelastning which is not exactly the same. The system administrator suggested the term inte for hög arbetbelastning (Andersson, 2018, p. 5), but the
informants agreed that it would not solve the doubt about what to be considered good – or at least acceptable results. An informant also pointed out that as teacher she would define projects where students worked extremely hard to reach the goals they had set for themselves, but it was not her, who had demanded that they worked flat out (Alveteg, Malm, & Sjödell, 2018, p. 2).

According to an informant who previously worked as AD, the theory behind the CEQ says that the bigger the workload the better, but his opinion/experience is that it should not exceed +50 and it should not be below 0. In connection with this he called attention to that especially in terms of the category Workload student evaluations are biased based on their experience from previous courses, i.e. if the workload was very easy in the first course, they might experience the following course as overloaded. If they felt overloaded in the previous course, they will experience the following course as easier, despite the teaching and the required workload remaining unchanged (Alveteg, Malm, & Sjödell, 2018, p. 3). How students experience and thus evaluate workload is thus uncertain and the only thing that is possible is to have a guideline that indicates the acceptable interval. The informant told that when he was Programme Director, he would react if the Workload was less than ÷20 and above +40. A statement that underlines the uncertainty since it was the same informant, who earlier pointed out that the interval should be between 0 and +50!

5.2.6. Subconclusion – troubles and contradictions
In the article Doing course evaluations as if learning matters most, Dr. Kristina Edström argues that the purpose of an evaluation process must be determined from the beginning, because quality enhancement and quality assurance on a practical level need rather different investigations. Edström quotes Patton, who has observed that: The same data seldom serves both purposes well (Edström, 2008, p. 99). If the purpose is enhancement then teachers will try to identify problems where there is a need for improvement, but if the purpose is quality assurance then the aim is to create a basis for fair judgment. As a consequence, it can be dangerous to focus on problems as they can be turned against the teacher. If the two objectives are combined in the same evaluation activity, the quality assurance part will unavoidably become the focus, because of the risk. In this case it does not matter much what the aim of the institution is, if that is the way that the individual teacher sees it (Edström, 2008, p. 100).

The five present troubles and contradictions that I have identified during the interviews reveal insecurity about what is an acceptable standard to achieve in your evaluations. Or rather an insecurity about when
your teaching has been good enough. The management has in all ways kept its original promise and has not used the evaluation system for quality assurance except for aiming at high professional standards, when it came to statistical analysis of the large CEQ-datasets. By giving teachers the responsibility to enhance the quality of their teaching in their own way, they have simultaneously left especially new teachers in a vacuum, where they find it hard to navigate and know when their teaching and the students learning have been satisfactory. This tendency is reinforced because teachers in general keep their evaluations to themselves and only share in private with a few trusted colleagues. And it is suddenly even more understandable that the two experienced teachers are the ones that are most satisfied – also with the cooperation with the students. The two teachers were professor and Head of Department respectively – in other words, they had received an official stamp of approval from the management that they do a good job. If the management presents a number of quality parameters, you know which criteria to meet in order to say that you have done well. If you disagree with the criteria – at least you can agree to hate the management. It is more difficult to define your own set of criteria for good teaching – especially because you have to re-develop them if you want to improve your teaching. At LTH it can be very difficult to hate the management, simply because they haven’t defined criteria of quality. For want of something better, you can always agree to be annoyed about the students.

In the next chapter I will be analysing the evaluation system at LTH through Yrjö Engeström’s third generation of Activity Theory with two interacting Activity Systems (in this case the teachers and the management) as the smallest unit of analysis. By studying who is learning and why, what and how they are doing it, I can identify the dynamics of the learning in the organisation (LTH). By means of the acquired knowledge I can create models/changes that can pave the way for expansive learning because of quantitative evaluation data being used to a greater extent to improve the teaching and thus the student learning.

6. Engeström’s third generation of Activity Theory

In the article Expansive Learning at Work: toward an activity theoretical reconceptualisation, the educational researcher Yrjö Engeström introduces the third generation of activity theory where he takes two interacting activity systems as its minimal unit of analysis (see fig. x). This makes it possible to focus research efforts on the challenges and possibilities of inter-organisational learning (Engeström, Expansive Learning at Work: toward an activity theoretical reconceptualization, 2001, s. 133).
In order to examine activity theory and its concept of expansive learning I use the four central questions that according to Engeström any theory of learning must answer:

1. Who are the subjects of learning – how are they defined and located?
2. Why do they learn – what makes them make the effort?
3. What do they learn – what are the contents and outcomes of learning?
4. How do they learn – what are the key actions or processes of learning?

Engeström presents five central principles of activity theory, i.e. the activity system as a unit of analysis, the multi-voicedness of activity theory, historicity of activity, contradictions as a driving force of change in activity, and expansive cycles as possible forms of transformation in activity – in this case aiming at extended use of quantitative evaluations to improve teaching. The four questions and five principles form a matrix that I have completed based on the interviews (see appendix 6)

6.1. Who are the subjects of learning?
In this case, the subjects of learning are the teachers and the management/administration that represent the two sides of the activity system as they are facing each other with what could be conflicting interests of quality enhancement (teachers) and quality assurance (administration) though both share the common objective that students learn more.

![Yrjö Engeström’s Activity Theory](image)

**Fig. 3**
Yrjö Engeström's Activity Theory with two interacting Activity Systems as its smallest unit

![Engeström's Activity System](image)

**Fig. 4**
Engeström’s Activity System. Who is learning is marked by red circles.
When using the Activity Theory it is possible to add a separate activity system. Students are of course a subject of learning too, but since they are not the subject matter of the research and because it is the institution that defines their learning objectives, they will not be included as a separate activity system. The focus will be upon teachers and the administration.

Other subjects that are learning during the evaluation process belong to the category “division of labour” that on the management side consists of the management, the student assistants, Academic Developers and the system administrators. The teacher side consists of teachers/Course Coordinators, Academic Developers (AD) and Student Representatives. Academic Developers are an integrated part of both systems, though mainly the management since they implement research-based standards for educational theory and practice that becomes official guidelines for good teaching.

6.2. Why do they learn? – What makes them make the effort?

For situated learning theory (Lave & Wenger) motivation to learn stems from participation in culturally valued collaborative practices in which something useful is produced (Engeström, 2001, s. 141). This could be the case for new teachers entering into a cooperation, but Engeström also points out that novices gradually gaining competence in relatively stable practices is not equivalent to motivation for risky expansive learning process (Engeström, 2001, s. 141-142). In contrast to classroom teaching with specific learning objectives, it is not clearly defined what we have to learn in a workplace. The object is more often implicit than explicit and it is more like an intention rather than a specific goal. To pave the way for expansive learning is to identify the objects and how changing or adjusting other aspects in the Activity System can affect them.

Identifying why the subjects are learning depends primarily on the objects and secondly the rules – in this case that evaluations are mandatory, which also implies that one’s evaluations will be presented to the management. In the subsequent chapters, I will be discussing the objects that I have identified from the teachers and management side respectively. From the management side this implies an interest in ensuring the good reputation of LTH, from the teachers’ side an element of vanity. A common objective that students learn as much as possible and an interest in seeing progress – though what is considered progress depends on the subject.
6.2.1. Teachers’ Object: Evaluations will be presented to the management
Whereas several informants every now and again expressed an irritation over students’ non-binding participation in the evaluation process, no one said anything about a possible problem of regularly having to discuss your evaluations with a representative from the management. Even if it is true that none of them feared losing their job due to negative evaluations, it can still be a central reason for optimising your teaching that you are being assessed by the end of each course and that the management and the other teachers in your course will know about your results. You will also have to participate in CEQ meetings where student representatives are present. To qualify as a senior lecturer you have to pass your portfolio project, where teachers are encouraged to refer to their evaluations. Whether teachers like it or not, they have to present their evaluations to somebody else (including authorities), which most likely will make them make the extra effort to do well.

6.2.2. Teachers’ Object: Quality enhancement because of vanity. The devil’s favourite sin
During the interviews, the informants were presented with a Work Report containing very positive evaluations and one of the informants commented: *If I’d gotten a report with these results I would take it easy, because it doesn’t appear as if there are any dramatic problems in this course. I would read the report when I’m in a good mood* (Alveteg & Malm, 2018, p. 7). It reflects an element of vanity, which teachers have and which should not be neglected as a driving force for improving your teaching. It is good, but not sufficient, if they can see that students have learned a lot. It is also important to achieve the acknowledgement that you did well. To be confirmed personally by students’ learning. An informant pointed out, that student representatives should be reminded of proportions if they focused on three...
negative free-text comments, when there were 200 students in the class. And she later said that it was also necessary that she reminded herself of it (Lindholm, Gomez, & Nilsson, 2018, p. 7). Vanity can be a driving force for making better teaching, but there also appears to be an unspoken consensus among teachers that it is acceptable to be upset by even minor complaints. A whining that in other workplaces is more likely to be considered annoying and socially unacceptable. However, in the LTH staff room the outcome (Object(2)) of quality enhancement is improving the teaching and as a consequence better learning for students (Object(3)).

6.2.3. Management Object: Evaluations contributes to ensuring the reputation of LTH
Object(1) from the management side is among others quality assurance in order to maintain the high professional standards at LTH to consolidate the good reputation of the institution within their subject fields (Object(2)). Evaluations are a means to carry out quality assurance.

One of the means of achieving this is by systematically monitoring the teaching through (among other things) evaluations. By means of ADs they can identify possible needs for supervision and academic development among the teachers (Lindholm, Gomez, & Nilsson, 2018, p. 3).

6.2.4. Shared object: It makes sense to everyone that students learn as much as possible
Everybody that I spoke to expressed a genuine interest in his or her students learning as much as possible. It was a meaningful objective to them. The question was if evaluations were a sufficient measurement for the extent of student learning. Some considered it as bricks in a larger puzzle that constituted the picture of the student learning (Alveteg & Malm, 2018), whereas others found it hard to see the proportion between the large datasets that were accumulated and the (few) statistics that were pulled on the basis of it (Alveteg, Malm, & Sjödell, 2018, p. 11). Two teachers had made their own questionnaire. One in order to evaluate her own specific learning objectives (Alveteg, Malm, & Sjödell, 2018, p. 10), the other because he wanted to get his own evaluations knowing that they were specifically addressed to him. He was a part of a larger course and it did not appear from the CEQ-data which feedback concerned him (Lindholm, Gomez, & Nilsson, 2018, p. 2).

Both teachers and management are driven by the object that students learn as much as possible whether the initial object is quality development or quality assurance (object(1)) with an outcome either to improve the teaching or ensure the reputation of LTH (Object (2)) resulting in increased student learning (Object(3)).
6.2.5. Studying evaluations are closely related to looking for progression. The question is *which one?*

It has to be noted that several of the informants belonged to both parts of the activity system as they were at the same time teachers and (vice) Programme Directors. The interviews revealed a common interest in using evaluations to monitor a progression, but the kind of progression they were looking for varied. Whereas teachers would look into details about specific elements in their teaching, the administration were looking for more general progressions in the main categories for a longer time span (Lindholm, Gomez, & Nilsson, 2018, p. 13). This will be discussed further in chapter 7 as one of the solutions in order to increase the use of quantitative evaluation data is an interface or a dashboard, where teachers as well as management can compare data on an individual basis in order to examine the particular progress that they are looking for.

6.3. What do they learn? – What are the contents and outcomes of learning?

According to the informants, they only skimmed the quantitative part of the Work Report, but that does not necessarily imply that the quantitative data are not affecting their attitude to teaching – both in terms of current reminders of the standards for teaching and learning at LTH and as incitement for the free-text response. In the Activity System the focus is upon Object 2 and 3 as the outcomes of Object(1), which particularly is about improving the teaching and to recover from inappropriate comments in the evaluations.

6.3.1. Evaluations as recurrent reminders of guidelines for learning and teaching

Object(1) is for many teachers – among other objects – that their students learn as much as possible. When students regularly have to complete the same questions, which require that they reflect upon their learning, students are repeatedly reminded of a focus in their learning process (e.g. the general academic...
skills they must acquire and that they are supposed to understand, not memorise the subjects they are introduced to). Students answer the 26 quantitative questions before they reach the free-text questions, which means that their mindsets are set on the standards they have responded to. This does not prevent them from adding inappropriate comments and expressing their feelings whether justified or not. Speaking from own experience, removing the quantitative questions, only to maintain the qualitative ones, will result in fewer and more unfocused comments.

As for teachers, when they regularly have their teaching performance evaluated, they are reminded of the (CEQ) standards of good teaching. That they need to specify the learning goals. That they need to teach their students general academic skills. That they should aim at constructive alignment between the teaching and the corresponding exam and finally that they do not overload their students so they start memorising rather than studying to understand the subject.

As for the management, their object(1) of quality assurance has an outcome that they focus upon - or learn what students think about the teacher’s performance. Through the CEQ-meetings they learn if student representatives and course coordinators agree to the feedback or if there are deviations in need of further investigation. Bearing in mind that negative evaluations do not necessarily imply that students have not learned anything. On the contrary – e.g. a heavy workload often results in negative evaluations, but also increased student learning.

An outcome of object(1) that students learn more is listening to reminders of what good teaching performance implies and as a further outcome what teachers should focus upon to integrate in their teaching, if they want to obtain better evaluations.

6.3.2. Teachers learn to survive through their informal networks

*My colleague has also been working many years as a teacher, but we had to hold a debriefing where we patted each other’s backs and re-assured one another* (Alveteg & Malm, 2018, p. 12). A CEQ-meeting lasts 20 minutes, but the main part of the evaluations is being discussed in informal networks where a part of it is re-assurance. These informal networks are gradually developed and it is something that new teachers learn in the course of time that they can find reassurance among colleagues whether they are new as themselves (Lindholm, Gomez, & Nilsson, 2018, pp. 15-16) or an elder colleague takes the responsibility upon him/her self (Alveteg & Malm, 2018, p. 5). In this case, object(1) is the wish to survive/not to be too
deeply affected by evaluations and the outcome (Object(2)) is learning to seek help among colleagues that you trust in.

6.4. How do they learn - What are the key actions or processes of learning?
Since Vygotsky defined the original Activity Theory the cultural-historical approach was primarily a discourse of vertical development toward “higher psychological functions”, but Engeström accentuates that the third generation of Activity Theory needs to develop conceptual tools to understand dialogue, multiple perspectives and networks of interacting activity systems (Engeström, 2001, s. 135). In this case I am aiming to move from object(1), which is quality enhancement when the subject is the teachers. The outcome (object(2)) being better teaching. When the subject is the management, object(1) is quality assurance and the outcome (Object(2)) is to consolidate the good reputation of LTH. Both subjects meet in a shared perception, which is that students learn more (Object(3)).

6.4.1. The Management is learning through formal communities
In his role as Head of Department an informant talks about especially looking for evaluations, where he has heard rumours (Alveteg & Malm, 2018, p. 12). Through their access to all evaluations in their department, the CEQ-meetings that follow them and possible dialogues with academic developers, a substantial part of the management’s learning about the evaluations take place in formal communities were the results are being processed. It is also through the formal information procedures that the management gets an impression of how students have developed academically – judged according to the distribution of grades or the averages that a class of students has obtained.
6.4.2. Teachers learning about student development primarily happens through informal communities

Grades reveal the academic level of the student, but they do not reflect the learning curve. When it comes to processing students’ learning from a course, teachers’ learning primarily has its source in the sensing that takes place in the class when they are teaching (Lindholm, Gomez, & Nilsson, 2018, p. 11) and otherwise in the informal communities that are gradually being developed. This implies mediating artifacts such as the dialogues with students as part of the paperless mid-term evaluations, during the breaks and cosy chats as part of supervision (Alveteg & Malm, 2018, p. 1). Or incidental meetings after a course has ended (Lindholm, Gomez, & Nilsson, 2018, p. 5). It is conversations with teaching assistants during lunch-breaks about the progression of a course (Lindholm, Gomez, & Nilsson, 2018, p. 5) and maybe small confidences to trusted colleagues by the coffee- or copy-machine.

One of the informants was a former AD and had tried to establish informal chats about evaluations during breaks, but never succeeded (Alveteg & Malm, 2018, p. 12). There is an unwritten rule among teachers that you do not interfere in your colleagues teaching without being asked. A rule that includes evaluations too. There exists a clear need to be in control of whom your evaluations are shared with. If anyone at all. Most evaluations are studied in private. In the private setting teachers look at the basic facts – especially the self-reported presence – to see, if there is an accordance between the answers and their own impression during the course (Lindholm, Gomez, & Nilsson, 2018, p. 9). Despite the fact that most of them are reluctant to study the quantitative replies, some tell that they can be looking for specific questions if they have taught something that relates to it (e.g. written communication) (Alveteg & Malm, 2018, p. 9).

It is a common denominator that studying and dealing with feedback on your teaching is something that if it is not done in private then at least in informal communities, where any response is non-binding. The teacher remains in control and can select the response and suggested actions that he or she personally prefers. Just as they at any moment can break off a conversation they do not want to continue. This is in contrast to formal meetings where presence is required and response and decided actions are registered in writing as a binding point of reference.
6.4.3. Writing comments for the End Report forces teachers to relate to their data

A complete End Report requires that Course Coordinator as well as Student Representatives and Programme Director write a comment based on the Work Report and the CEQ meeting. At the meeting central issues are being discussed, and it is decided which actions must be taken in order to make the teaching and learning better. That CC, SR and PDs have to write those comments is a requirement that is monitored and considered one of the major issues about evaluations (Andersson, 2018, p. 6). Having to be present at a meeting where evaluation results are being discussed and afterwards having to comment on actions that will be taken makes the End Report an ideal mediating artifact for the management as it forces the participants to relate to their data. Perhaps that is the best way for the management to ensure that evaluation data is being studied and used to improve the teaching so that students learn more.

Based on the interviews and the following clarification of who is learning, why, what and how they are learning I can now specify Engeström’s Activity Theory when using it for the evaluation process at LTH:

![Engeström’s third generation of Activity Theory with two interacting activity systems as its minimal unit of analysis](image)

Fig 8 Engeström’s third generation of Activity Theory with two interacting activity systems as its minimal unit of analysis.

In the third generation of Activity Theory the minimal unit of analysis is two interacting activity systems, which makes it possible to focus research efforts on the challenges and possibilities of inter-organisational learning (Engeström, 2001, s. 133). Notice that the lines in the illustration are in fact arrows that point in both directions as an indicator of the mutual dependence between all the different aspects. The objects are depicted as ovals, the shared object (3) being that students learn more, which emphasises that the object-
oriented actions are always characterised by ambiguity, surprise, interpretation, sense making, and potential for change. Just as the object(2) that is Better teaching (teachers)/consolidate the good reputation of LTH (management).

<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>

**Mediating artefacts / Tools and signs**
- Mid-term evaluations
- Personal assessment of the learning of the class
- Informal chats with students
- Self-evaluation
- Supervision
- Peer-reviews
- Exams/assessments
- CEQ-system
- Academic Development unit (Genombrottet)
- Exams/official assessments
- The website ceq.lth.se containing the majority of documents related to the CEQ since 2003!

**Rules**
- Unwritten rules:
  - Teachers do not discuss evaluations unless somebody explicitly mentions it
  - Teachers do not interfere in each other’s teaching unless explicitly requested to
  - Evaluations are something that you have to get through as fast as possible
- Written/official rules:
  - Evaluations are mandatory
  - Focus in the evaluation process is supposed to be upon analysis and discussion of the courses with the aim of attaining better teaching. Time should not be spent on gathering data and writing reports (Andersson, 2018, p. 1)
  - CC, SR and PD must write comments for the End Report based on the CEQ-meeting.

**Community**
- Informal networks
  - Colleagues
  - Other teachers
  - Academic developers
- Formal network
  - Teachers
  - Academic Developers
  - Management

**Division of labour**
- Course coordinators
- Student representatives
- Student assistants
- (Academic Developers)
- Programme directors
- Course coordinators (teacher)
- Student representatives (Academic Developers)
A general danger about evaluations is that the farther away you are from the actual data collection the more likely you are to take them for granted (Edström, 2008, p. 99). This has so far not proven to be a problem at LTH as the management since the implementation of the CEQ-evaluation system has kept their promise of not using the evaluation results as a reason for promotion, salary increase or dismissals. It should remain so if teachers should continue to focus upon quality enhancement and improvement of their teaching. Preferably also by using their quantitative data as a useful mediating artifact. In the following chapter I will suggest a number of solutions that should pave the way for expansive learning regarding evaluation data.

7. Creating a New Model
The fifth principle of Activity Theory proclaims the possibility of expansive transformation in activity systems. According to Engeström Activity Systems move through relatively long cycles of qualitative transformations. As the contradictions of an activity system are aggravated, some individual participants begin to question and deviate from its established norms. In some cases, this escalates into collaborative envisioning and a deliberate collective change effort (Engeström, 2001, s. 137). In this case the aim is not an expansive transformation, but to prepare the way for it by making quantitative evaluations more useful to teachers so they actively begin using them as a rich data source or mediating artifact for making their teaching better so students learn more. To make small changes that will hopefully reduce the number of inappropriate comments in the Work Report and make the quantitative data more accessible by being able to work with them in order to study possible progressions. In general, creating a more comfortable atmosphere around evaluations by reassuring teachers that they are not the only ones to have been exposed to inappropriate comments. Hopefully to pave the way for greater openness in especially informal communities.

It is not the evaluation system in itself that produces the positive effects. Studies have shown that improvement of teaching is possible when the evaluation is combined with interventions such as support, for example academic development supervision to teachers, or otherwise is used as part of an aligned set of strategies (Edström, 2008, p. 102). This could include a systemic alignment, where Edström suggests that instead of including quantitative data in portfolios in order to illustrate a progress in one area or another, teachers should instead be encouraged to discuss the actions they have taken to improve their teaching based on their evaluation data (Edström, 2008, pp. 103-104).
Based on my interviews and analysis of the data, I will suggest a number of tools or mediating artefacts for the teachers where some of them will also be useful for the administration. Other tools are aimed at students with the intention to affect their feedback to teachers so among others free-text answers become more constructive and useful for planning and executing the teaching.

During the interviews, the possibility of arranging workshops about evaluations as part of academic development was discussed several times. Every time the idea was dismissed on the ground that it was not very likely that people would like to share their evaluations (e.g. (Lindholm, Gomez, & Nilsson, 2018, p. 8)). They already felt uncomfortable about the evaluations being shared with several during the CEQ-process and the End Report being available to all their colleagues (Lindholm, Gomez, & Nilsson, 2018, pp. 5-6). Experienced teachers revealed that they would keep the Work Report to themselves to make sure that new teachers were not on their own when receiving it. Instead, they would go to see them and talk about the response they had received – the credit, when it was good and the reassurance and how to move forward, when the feedback was less positive (Alveteg & Malm, 2018, p. 5). Making it a permanent arrangement (e.g. like a mentor) did not seem to appeal to anyone.

In general, evaluations are private and to pave the way for evaluations as part of organisational learning, it is a starting point that privacy has to be respected, so teachers only have to share to the extent they feel comfortable about it (except for the CEQ-process that is already taking place). Below are listed a number of solutions that all have privacy as starting points:

7.1. Adjustment of questionnaire and associated reports

Adjustments of CEQ-questionnaire – Are you satisfied with your own performance?
The quantitative questions remain unchanged, but the initial quantitative questions (e.g. sex, year of birth, year of study etc.) also contain the question: To what extent have you participated in various course activities? One could add: Are you satisfied with your own work performance? (Answer: 6-step scale from 0-100%). The aim of the added question is to make students reflect upon their work performance and later be able to study possible statistical correlations between answers and the students’ attitude to their own performance.

Adjustment of CEQ-questionnaire – Extending the free-text question
In order to encourage students to be more considerate when answering the free-text questions, the question What do you think was the best thing about this course? could be rephrased:
What do you think was the best thing about my course?

Or

What do you think was the best thing about our course (please, address the teacher you are writing to)?

The second question What do you think is most in need of improvement? could be rephrased:

What do you think is most in need of improvement and why? (Alveteg, Malm, & Sjödell, 2018, p. 12).

The rewordings invite students to give reasons for their complaints and they will be more likely to write in a proper tone when addressing teachers directly.

It was suggested that students were reminded of the learning objectives (e.g. they were presented to them in the beginning of the questionnaire). Based on experience that students do not read them anyhow I will not recommend it.

Adjustment of the Work Report – Remove the Summary Scales

Remove the Summary scales divided on satisfaction. All informants were asked to explain the scales and no one except the system administrator was able to do it. The most experienced teacher had an understanding of the information you could deduce from the scales – the rest said that they never looked at them. It was a routine common to all informants to skim the front page, read the free-text answers, maybe skim the histograms for the quantitative replies. The scales added to the feeling that the Work Report contained an overload of unnecessary information. Since the scales were introduced to please math-statisticians 15 years ago and the majority of the ones that protested has retired in the meantime, it does not seem a loss to remove them.

Adjustments of the End Report

By the time the End Report is ready to be published (i.e. the CEQ-meeting has taken place and the participants have written their comments), students have been graded and it is settled how many have passed or failed the course. It will put the feedback into perspective to add this information to the End Report. To challenge the students workwise can increase their professional development, but the heightened pressure of work often results in (temporarily) dissatisfied students and negative evaluations.

7.2. Systematic AD supervision – especially for new teachers

The interviews reveal that new teachers receive no systematic support when it comes to addressing their evaluations. It is only if more experienced teachers independently take the responsibility upon themselves that they are given any assistance. Don’t take it too hard, was the best advice one of the informants could
give to new teachers, which reveals that it is tough reading one’s first evaluations. Especially if you are left on your own (Lindholm, Gomez, & Nilsson, 2018, p. 15). A solution could be systematically to offer new teachers AD supervision about how to interpret and use the quantitative data (including that a response between $\pm 20$ and $+40$ is acceptable for the category “Appropriate Workload”). How to use the constructive comments in the free text to improve the teaching. And to be aware that everybody receives negative comments in their evaluations and they should focus upon the constructive and positive comments.

Note: Vanity figures among the seven deadly sins and the supervision is not intended as “a shoulder to cry on”. We can also demand that teachers as grown-ups have a sense of proportion and do not attach great importance to one drop of criticism in an ocean of positive remarks.

7.3. Video tutorials aimed at teachers and/or students

*Tutorials in Swedish and English available at Kompetenceportalen at LTH.SE*

An informant suggested to look to ICA for inspiration, at they had their own system for upgrading the skills of their employees consisting of small tutorial videos, which were compulsory to watch (Alveteg, Malm, & Sjödell, 2018, p. 14). Another informant added that videos should last no longer than 3 minutes and be available in both Swedish and English. A surprising number of the students did not speak English (Alveteg, Malm, & Sjödell, 2018, p. 18).

The following subjects are suggested as a starting point:

<table>
<thead>
<tr>
<th>Content</th>
<th>Aim</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>What is the Course Experience Questionnaire (CEQ)</strong></td>
<td>An introduction to the CEQ aimed at both teachers and students to understand the research that the system is based on in order to study in-depth learning. Several informants expressed that they did not know what justified the CEQ as evaluation system. By understanding the research and intentions behind the CEQ it is the aim that both teachers and students will develop a greater interest in using it. (Please, see Appendix 7)</td>
</tr>
<tr>
<td>• Research based – what is it measuring?</td>
<td></td>
</tr>
<tr>
<td>• Why did LTH select the CEQ-system?</td>
<td></td>
</tr>
<tr>
<td>• The main categories and why LTH selected the specific five categories</td>
<td></td>
</tr>
<tr>
<td>• Why did LTH add the questions about relevance (Q17) and satisfaction (Q26)</td>
<td></td>
</tr>
</tbody>
</table>
### The CEQ-system at LTH

1. Mandatory to do evaluations when more than 25 students in a class (optional if fewer students)
2. Optional to add 1-4 extra free-text questions
3. When is the CEQ carried out?
4. What does a Work Report contain?
5. The CEQ-meeting
6. The End Report (including remarks by SR, CC and PD)
7. Where to find the CEQ documents (ceq.lth.se)

### Be specific and constructive (negative)

One by one, teachers working at LTH provide a specific example of an inappropriate comment they have received in their evaluations.

The stories are mainly aimed at students to make them understand that malevolent comments does not improve the teaching. The second target group is other teachers – indirectly communicating that they are not the only ones to have received negative response. (Please, see Appendix 5)

### Be specific and constructive (positive/constructive)

One by one, teachers working at LTH provide a specific example of useful comments that had helped them improve their teaching.

Students are the primary target group with a direct request for constructive criticism with specific suggestions for the teacher about what they think would make the teaching better. Second target group is other teachers – it is a possibility to show the tutorial just before the CEQ is completed – or when the CEQ is mailed to the students to attach a link to the tutorial. (Please, see Appendix 5)
An informant suggested the tutorials could be available at *Kompetenceportalen* (Alveteg, Malm, & Sjödell, 2018, p. 19). The advantage of tutorials is that you can go back to them when you get the evaluations and re-study them in private when it is relevant for you (Lindholm, Gomez, & Nilsson, 2018, p. 8).

**7.4. Upgrading of Student representatives / Student assistants**

As discussed in the chapter about **present troubles and contradictions** the problem with student participation was a recurrent issue in the interviews and though the system administrator gave the student assistants lessons before they started filtering the comments, apparently too many inappropriate comments slipped through (Alveteg, Malm, & Sjödell, 2018, p. 12). A start could be to watch the tutorials “Be specific and constructive” – both the negative and the positive, to get an understanding of which comments should be removed.

Secondly, a workshop where the participating students are given all the free-text answers from a course. Each student makes their own filtering and afterwards the participants compare and discuss their selections and rejections mutually. An academic developer could supervise the session.

**7.5. Interface – teachers work with their own data**

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 it would be relevant to compare (e.g. the summarising histogram or the histogram representing a single question). Quantitative data becomes more informative when you can compare them, but it was 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.

Based on the interviews I have made a draft of the services this interface could offer. After submitting this thesis I found out that this interface already existed in the form of a supplementary dashboard-unit to the technical evaluation system Blue developed by the Canadian company *explorance*. In the following chapter I will present the suggested service and subsequently compare it to the dashboard-unit from *explorance*. 
7.5.1. Introductory questions

<table>
<thead>
<tr>
<th>Inquiry</th>
<th>Aim</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name and password</td>
<td>Giving the user access to their specific dataset(s) and noone else’s</td>
</tr>
<tr>
<td>Language - Swedish or English</td>
<td>Service available in Swedish and English as a common international language. Not all students speak English, so it is not possible to have an English version only.</td>
</tr>
<tr>
<td>Course code</td>
<td>The specific data that the user wants to work with and testing if he/she has access to this particular dataset⁷</td>
</tr>
<tr>
<td>Time interval (from-to)</td>
<td>The timespan to compare the data. It should be possible that “from” and “to” represent the same year if the user wants to study one particular category or question and compare to other averages at LTH</td>
</tr>
<tr>
<td>Datatype - what CEQ-question</td>
<td>Which particular question or category that the user would like to study</td>
</tr>
<tr>
<td>- which CEQ-category</td>
<td></td>
</tr>
</tbody>
</table>

7.5.2. Visualisation of quantitative data

In order to visualise the quantitative evaluation data I will go through the main principles as listed by Stephen Few – author of the book *Show me the numbers* – who works with data visualisation for analysing and communicating quantitative business information.

According to Few, messages that involve quantitative data always include related categorical data that identifies what the numbers measure, e.g. the year the data was registered. There are three types of categorical data based on how they relate to one another:

- Nominal (i.e. no intrinsic order)
- Ordinal (i.e. intrinsic order)
- Interval (i.e. conversion from a quantitative scale by subdividing a range of numeric values into a set of smaller ranges) (Few, 2004, pp. 1-2)

⁷ Using “Course code” to access a particular dataset can constitute a specific problem if you want to compare data during several semesters because course codes regularly change (Lindholm, Gomez, & Nilsson, 2018, p. 14)
Few describes a set of preattentive attributes, which combine to form what we perceive as objects. These attributes include 2D location, line, length, size, shape, orientation and colour, but Few only considers 2D location and line length as highly effective means to visually encode quantitative values. The three graphs below illustrate how these two attributes can encode quantitative values:

![Graphs illustrating 2D location and line length encoding](image)

Fig. 9: How 2D location and line length encode quantitative values in graphs (Few, 2004, p. 3)

As all three graphs have the same quantitative scale along the vertical axis and the same categorical scale along the horizontal axis it is apparent that it is the graph and the vertical bars that communicate most clearly, in contrast to the isolated points that appear visually weak.

Example: Studying the single question
The following illustrations and data are all fictitious – some of the illustrations are copied from LTH Work Reports, but are not related to the specific question.

Supposing a teacher wanted to compare the progress of a certain CEQ-question (e.g. Q3) from 2015 to 2018.

It is possible to compare the replies as they are distributed between the 5 categories from highly dissatisfied (+100) to highly satisfied (+100), but it also appears that it is very difficult to differentiate between the data.

![Histograms comparing CEQ-question progress](image)

Fig. 10: Comparing histograms for the same CEQ-question over a period of 4 years
The visualisation of the quantitative data becomes clearer when it is only the averages that are compared:

The line connects a series of individual data points and is simultaneously displaying the trend of a series of data points. The line shows the movement of values with a clear representation of change from point to point and thus reflects the trend in the feedback to the particular question (Few, 2004, p. 4). In comparison bars are a visual heavyweight and are better when it comes to visualising the individual numbers. The 3D visualisation reveals that adding the extra dimension is unnecessary as it does not communicate the data better. On the contrary, it overloads the memory by adding more information than strictly necessary.
Example: Comparing the averages of the same category

In the figures below I have worked with different ways of visualising the same (fictitious) data in the category Good Teaching over a period of five years from 2014 to 2018.

In fig x + y it appears that adding colours to the columns does not make the visualisation clearer.

In the figures below, I have added a tendency line – the first line in fig z reflecting the total average of the category during the five years that are being compared. In fig zz the line reflecting the total average is repeated as well as a line reflecting a fictitious LTH average for the same category during the same time interval.

A tendency line showing the average reflects an overall pattern so that the separate years can be compared. However, two lines in the same figure make the columns become invisible and as a consequence unimportant, because you as viewer only see and compare the two averages – and your focus is upon, which has the highest. If a tendency line is to be included it thus has to be only one at a time. Or you should be able to compare the tendencies without being disturbed by individual data columns.
In figure xx and yy I am pretending to compare the course average with the total LTH average of the category in question for each year.

The figures illustrate that columns are stronger, when it comes to communicating individual data, whereas lines has the ability to visualise a tendency. It is a matter of preference, which type of visualisation is the most useful in the given situation.

**Example: Ranking data**

Quantitative messages always reveal a relation between the data. In this case I am using Q26 (Overall, I am satisfied with this course) and how especially horizontal columns are useful to visualise the rank between data.

In contrast to ordinal scales that have an intrinsic order, nominal scales have not, so it does not matter whether we use the horizontal or the vertical axis. When it is an ordinal scale like a timeline – the time has to be on the horizontal axis and the quantitative data on the vertical axis (Few, 2004, p. 2).

In the figures below I am comparing the individual – or course result to the fictitious averages of the institute and LTH for Q26:
Because it is nominal data it is possible to rank it and it appears that it is easier to see the difference in size when using horizontal columns compared to vertical columns. In the context of ranking data, horizontal columns are visually stronger.

Based on my experiments with data visualisation, the following functions would be relevant in an interface where teachers can work with their own data:

- Solely focusing on averages – maybe adding standard deviation
- Always to be able to choose if data should be visualised as either columns or lines/graphs
- Being able to compare and rank ones data to averages at the institute and at LTH
- Being able to compare a data progression to a given average over the same period of time

It is a very simple interface, and it offers the need for comparing evaluation data, which the teachers and Programme Directors that I interviewed requested. In the following paragraph I will introduce a similar system that is already available, but which I was unaware, when sketching the guidelines for this interface.
7.6. The Dashboard-unit from *explorance* as a supplement to the evaluation system Blue

When this thesis had been submitted and just before the oral exam defending it, I came to speak to Senior Lecturer and PhD Torben K. Jensen who is the Head of Center for Teaching and Learning at Aarhus University (AU). I was informed that AU was just about to invest in a Dashboard unit to their present evaluation system that gave the teachers the possibility to work with their own data. I will present the Dashboard-unit very briefly in terms of the functions that AU requested and what the system has to offer. My knowledge is based on my talk with Torben K. Jensen as well as the sales material from *explorance* that he was kind enough to send to me.

AU wanted a system that was able to:

- Do statistics out of choice on the questions the user decided to focus upon
- Cross-tabulate with grades
- Study the development over time
- Compare with information from the study administrative system (e.g. is there a connection between the satisfaction with the teaching and the subsequent grading)
- Maintaining anonymity – the user should not be able to connect evaluations on an individual level

The sales material promises the following:

**Timelines:**

- Fig. 18: The average response to a specific question over time
- Fig. 19: The average response over time including Standard Deviation and min. and max values

The red circle in both figures highlights the function that the user can always change between columns and graphs.
Comparing subcategories of a dataset:

Fig. 20: Comparing the average response to a specific question showing the difference between full time and part time students

Fig. 21: Comparing the response to the teaching by different nationalities. The colours represent different nationalities and visualise how, on average, they are responding to the same teaching

Comparing results from other courses based on student filters or course filters

Fig. 22: Comparing course data based on a student filter (Nationality, Gender, Year of study, Full time or Part time and Graduate). Comparison illustrated by numbers

Fig. 23: Comparing course data based on course filters: Course code, Department, Course name and Division. Comparison illustrated by horizontal bars.

One can argue that teachers will only present the data where there is a progress in their results, but while they are comparing the data (maybe in search of a progression) they will also be facing the ones without a progression thus being reminded (but in private) of areas they should attend to.
8. Conclusion
As written in the introduction – Life may not be the party that we hoped for, but while we are here, we might just as well dance! It has been a recurrent issue during my research that the teachers I have interviewed certainly don’t consider evaluations a party and those feelings re-occur in peer-reviewed publications about the topic. However, evaluations are mandatory and in order to make the most of them – or at least more of them, I proposed the following research question:

How can we make the quantitative evaluation data more useful to teachers so as to increase their use of or work on evaluations?

Since it is often the inappropriate student comments in the qualitative replies that stop teachers from studying their evaluations thoroughly, the framework becomes essential in order to change the attitude to evaluations as “something that just needs to be overcome”. To prepare the way for expansive learning at LTH I used Engeström’s Change Laboratory and examined the historicity of the CEQ-system as well as the present troubles and contradictions.

The historical study revealed a very thorough approach to selecting and developing a research-based evaluation system when it became clear that evaluations would be mandatory at all higher education institutions in Sweden. The CEQ system was chosen and the pedagogical unit Genombrottet began a long-term cooperation with the Australian researchers that had developed the system. During the years, the evaluation data was examined to test its validity and the CEQ-system was studied with the aim of academic development. Genombrottet is also one of the few pedagogical units where ADs have done research about how teachers are affected by evaluations. In addition, the Management at LTH appears to have stuck to their word and not used evaluations against teachers. Nor is there any history of job cuts that makes teachers fear to lose their job and thus no anxiety that the management would use evaluations as a reason for employee termination in the future. There appears to be a solid basis for using evaluations as quality enhancement and thus for teachers to focus upon identifying the constructive negative feedback in order to improve their teaching.

In my empirical analysis I identified troubles and contradictions, which generally reflected an insecurity among teachers about when ones teaching was actually good enough. Particularly bearing in mind that
students are not always capable of assessing their own learning and consequently negative evaluations do not necessarily reflect bad teaching or lack of learning.

It was thought-provoking to realise that on one hand teachers wanted explicit procedures about what would be the consequences of bad evaluations and on the other hand the informal chats appeared to be the best way of dealing with them. Official procedures would most likely result in unnecessary fear and the feeling of stigmatisation and neither would lead to better teaching. Not surprisingly new teachers are particularly exposed to negative feedback and especially to avoid stigmatisation by selecting the ones that get especially bad evaluations, it is better systematically to offer AD supervision for all newly educated teachers.

Student participation was a double-edged sword as they had considerable influence upon the whole evaluation process, but no real responsibility. The only sanctions to give students that do not manage their task properly is to dismiss the assistants that continue to let inappropriate comments slip through. And to prevent student representatives to get re-elected if they do not show up at CEQ-meetings, do not write comments for the End Report or do not admonish their fellow students to use an appropriate tone. Instead of sanctions I would recommend an upgrading of skills for both student assistants and student representatives, especially before they begin carrying out their tasks.

The suggested solutions always aimed at maintaining the privacy around evaluations, so teachers only have to share anything related to the feedback when they want to. The recommendations included video tutorials where some of the tutorials should be a basic introduction to the CEQ-system and the process as only a few of the informants knew that it was a research-based evaluation system and the many academic reflections that lie behind them. This will hopefully lead to a greater motivation for using it.

All of the suggested tutorials address both students and teachers, but whereas the videos about the CEQ-system and -process are intended to inform everybody, the purpose of the tutorials "Be specific and constructive" is to reassure teachers that they are not alone. And to remind students that if they want better teaching they have to be specific in their feedback and that inappropriate comments do not contribute to that. Because this is what it is all about. To make teaching better, more efficient and preferably mutually rewarding. To develop pedagogically so that we 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.

My solutions have also implied adjustments of the questionnaire, the Work Report and End Report – everything aimed at respecting the need for privacy about evaluations – hopefully in the longer run to create a broader comfort zone around evaluations so they would be shared with more than a trusted few.

Dealing with evaluations happens in informal networks, which can be an excellent way to seek reconsolidation, but also implies the danger that you can handpick the replies that you want to hear and are not committed to follow up on some of the more pressing problems. Just as well as it can be only too easy to reconfirm each other in that it is the students’ fault – even though actions can actually be taken to improve the students’ efforts in studying or to raise the academic level. Something that is unavoidable through more formal and documented activities.

It will be interesting to register the spin-off effects that the dashboard or interface will result in. All informants expressed an interest in such a tool and would rather have it at their disposal sooner than later.

CEQ-evaluations aim at constructive alignment in the teaching, but to optimise the benefit from evaluations there must be a systemic alignment at a macro-level too. This is already the case in terms of the CEQ-process, data analysis by Genombrottet and the management keeping their promise not to use negative evaluations against teachers. It can be further consolidated through an encouragement to enclose End Reports (i.e. actions taken to improve teaching based on evaluations) in portfolios and maybe in the long run smaller, voluntary workshops with teachers to improve the suggested solutions. The next steps in Engeström’s Change Laboratory are Examining the new model, Implementing the new model and then reflecting upon the process.

The suggested solutions are intended to prepare the way from private to organisational learning regarding evaluations, hopefully enlarging the comfort zone so evaluations can be spoken about more openly paving the way for greater knowledge sharing. However, 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 any more than indications. Evaluations represent the students’ perception of the teaching and learning and we cannot assume that the teacher shares this comprehension, which is why evaluations are only a starting point for questions to be asked rather than the
nominal truth about the teaching and learning in a course. Definitely a reason for creating an environment that stimulates organisational learning with evaluations as an integral part.
Bibliography


