

Improving report writing by peer assessment using Coursera

H. L. Christiansen

Associate Professor

Technical University of Denmark, DTU Fotonik

2800 Kgs. Lyngby, Denmark

e-mail: hlch@fotonik.dtu.dk

Key words: Curriculum development, New education tools for engineering education

INTRODUCTION

Report writing is a general engineering competence and it should therefore be part of any university engineering education to learn how to write a good report. Active learning methods are well-known to be effective in supporting student learning; hence it should preferably also be used for teaching report writing. In the case of report writing active learning could include peer evaluation which is what is investigated in this paper.

This paper presents a case study from the Technical University of Denmark. A course on mobile communication was redesigned to include peer evaluation as a tool for improving report writing skills. The peer evaluation process was automated using the e-learning tool Coursera. What was investigated was the improvement in report writing as well as the consistency and quality of the peer assessed grades.

1 COURSE (RE)DESIGN

The course “Introduction to Mobile Communication” [1] at the Technical University of Denmark (DTU) is a 10 ECTS course spanning a full semester. In this course all students must write three reports. Two of them are on a given task in the lab while the third one is much more open as the students can decide on a topic themselves. The reports are part of the course evaluation (30%) together with an oral exam (70%). In the previous course design, the students used to hand in all reports by the end of the course. The teacher would then grade all the reports and after the oral exam the students would be notified of their final grade. This is the traditional way of doing assessment, so called summative assessment [2].

After having attended a teacher development program described in [4] the author redesigned the course to include active learning in the form of peer assessment [3] with the aim of improving the students’ report writing skills. The intention is that the report writing develops from being purely an assessment task to being also a learning task.

Thus, there is added a learning objective to the course description: “Write and evaluate reports on mobile communication”. [1]

In order to implement this, the students must get feedback on the reports during the course. I.e., write one report, get some feedback that can then be used to improving the second report. For the second report they also get feedback and then finally they have a lot of input for the third and final report. By getting feedback throughout the course the students can gradually improve their report writing and hence do better on the final exam. This method is also called formative assessment – getting timely feedback that closes the learning loop [2].

In the new course design peer assessment is used as the method for formative assessment. It is well known from studies of peer assessment that it is a very efficient way of doing assessment as it saves on teaching resources [5]. In addition it supports the students’ learning, which is what is important in this case [3]. One interesting finding is that what you really learn from is not only getting the feedback but also giving feedback. In fact, peer assessment has been shown to be effective specifically for improving writing skills [9].

Peer assessment can be summarized in two statements that are investigated in this case study:

- The evaluation from students is as good as the evaluation from the teacher [7]
- The students learn a lot from doing the evaluation[8]

2 PEER EVALUATION SETUP

To support the peer evaluation process the e-learning platform Coursera [6] has built-in facilities for peer assessment and it was used for implementing the peer assessment for report writing in this case. The students simply submit their reports using Coursera, and Coursera then has the tools for distributing the reports, collecting the feedback and presenting the results to the teacher.

In order for the peer evaluation to work the teacher must instruct the students how to perform the evaluation. A so-called evaluation rubric must be defined and it can consist of qualitative as well as quantitative questions. Furthermore, each student must evaluate more than one report. To set up Coursera you simply specify the number of reports each student must evaluate as well as the evaluation rubric.

In this case study an evaluation rubric consisting of four quantitative (assessed on a scale from 1-5) and one qualitative question (free text comments) was used, see Table 1. All reports were evaluated by (at least) four other students.

Table 1 - Evaluation rubric used for the peer assessment

Quantitative question #1	Possible answers
Structure Evaluate the structure of the report? Is the layout good? Does it have: <ul style="list-style-type: none"> • A front page? • Introduction? • Table of contents? • Figure captions? • A clear subdivision into sections? 	1 - No structure at all!!
	2 - Only vague indication of a structure
	3 - Still room for improvement
	4 - Only minor shortcomings
	5 - Perfect - this is exactly how it should be done!
Quantitative question #2	Possible answers
Language How well is the report written?	1 - Very poor language - hard to grasp the meaning of the sentences
	2 - Still ample room for improvement. Many sentences do not make sense at all
	3 - It all makes sense. However, many errors and typos obstructs the flow of reading
	4 - Almost perfect - only minor spelling and grammar mistakes
	5 - Perfect - Good job!
Quantitative question #3	Possible answers
Results Please, assess how well the results are documented. Typically, graphs from the simulator are a good way of providing documentation.	1 - No documentation at all - did you attend the lab?
	2 - Some results are there, but a lot is this missing. Clearly not good enough
	3 - Nice try! A decent attempt has been made, but there is room for improvement.
	4 - Almost there. There are only very few results missing
	5 - Perfect. All results are documented
Quantitative question #4	Possible answers
Link to theory Please, assess how well the results are explained by referring to the theory.	1 - No explanations at all - did you read the book?
	2 - Most explanations are missing
	3 - Has tried to explain everything, but there are misunderstandings and some explanations are not correct
	4 - Good. Only minor details are missing.
	5 - Everything is well explained! Perfect!
Qualitative question	
Any other feedback? How do you like the report in general? Is this a good way of documenting the lab work?	

3 RESULTS AND DISCUSSION

The results of the peer evaluation study presented are based on a total of 34 students participating in the class during 2014. This (small) population of students consisted of 20 bachelor students, 6 master's students and 8 international guests. There were 5 females and 29 males in the group. In the following, the grades awarded by the students are called peer grades and the students are called peer graders.

Figure 1 shows the peer assessed grade distribution generated by Coursera for report one and two. What can be seen is that the number of reports receiving the highest grades is higher for report two, which could be interpreted as an improvement. However, the variation between grades is also larger, which could be a consequence of report 2 being harder to evaluate.

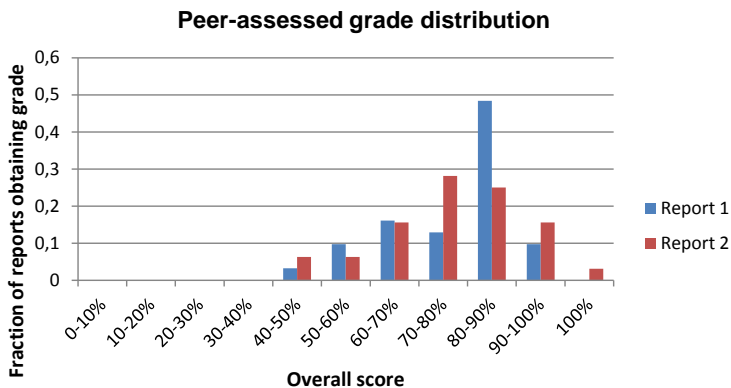


Figure 1: The output from Coursera. Grade distribution for report 1 and report 2

The following results are based on the raw output from Coursera, which has then been post processed. The raw Coursera output is the result of all the assessments, i.e., in this case four evaluations of each individual report.

3.1 Improvement of report writing skills

What was investigated here was if the ability to write reports had improved. This is interesting as report writing was added as a learning objective to the course, i.e., the main point for introducing peer assessment was an attempt to improve report writing skills. This is obviously difficult to measure but an attempt is being made here by answering the much simpler question: was report 2 better than report 1?

Figure 2 shows the average (peer assessed) grades obtained within the four quantitative questions in the evaluation rubric for report 1 and 2, respectively. What can be seen is an improvement in the category 'structure' whereas the other areas show a slight deterioration. Hence, from this we cannot conclude that the report writing skills had improved.

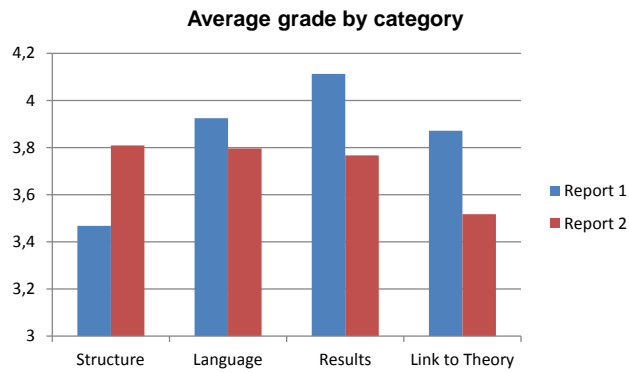


Figure 2: The average grades obtained in the four quantitative evaluation rubrics for report 1 and report 2, respectively

3.2 Consistency of peer assessment

What was investigated here was the uniformity of the peer assessed grades. This is defined here as how consistently a group of students evaluated the same report.

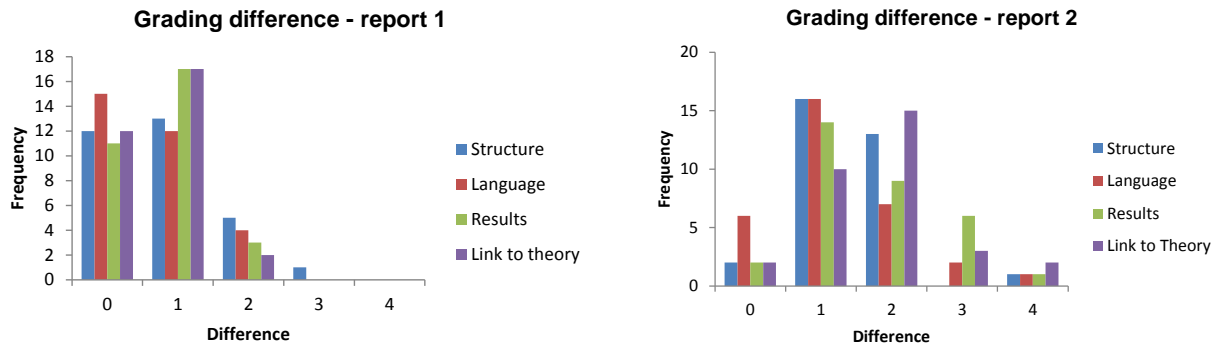


Figure 3: Consistency of grading for report 1 and report 2, respectively.

Figure 3 shows histograms of the difference between the minimum and maximum grades awarded to the same report within the four categories by the peer graders for report 1 and report 2, respectively. Hence, lower numbers show higher consistency (agreement) among the peer graders. A large difference indicates disagreement but could also be interpreted as an indicator for how difficult this rubric is to assess for the peer graders.

What can be seen is a difference between the first two and the last two grading rubrics. The first two rubrics (structure and language) are very generic and require (almost) no knowledge of the underlying theory, whereas the two last grading rubrics (results and link to theory) are harder to assess if the peer grader does not have a thorough understanding of the theoretical background.

By looking at Figure 3 a number of interesting findings can be made: The peer grades for report 1 were more consistent than for report 2. For report 1 only very few peer grades differed by more than 1 grade, whereas report 2 showed a much larger variation in the peer grades. Especially the “hard” rubrics requiring understanding of the theory show a large difference. This can be explained by the fact that report 2 covered a lot more theory than report one. Report 1 was done after only two weeks of teaching and therefore was very rudimentary and relied on a limited amount of theory. Report 2 was done after 7 weeks of teaching and required knowledge of all material covered until then. Clearly, good students are able to assess how well the theory has been explained whereas poor students cannot.

3.3 Quality of peer assessment

It is often debated how good peer evaluation is [7]. Is it actually as good as the professor’s grade? This has also been investigated here.

All reports in this case study were also evaluated by the author and his Ph.D. students. Hence, the results from this “professional” evaluation can be compared to the grading from the peer assessment (calculated as the average of all the peer grades). The result is shown in Figure 4. The difference is calculated as the instructor’s grade minus the peer grade. I.e., a positive difference indicates that the instructor’s grade was higher than the peer grade.

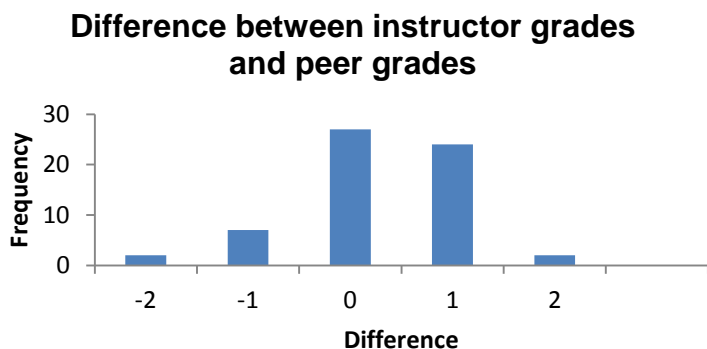


Figure 4: The difference between instructor’s grade and peer grading

From Figure 4 we can observe the following: Generally the “quality” interpreted as the correspondence to the professor’s grade of the peer assessed grades is very good. This is in line with other studies [7]. There is a tendency that the professor’s grades are generally a little bit higher than the peer grades. That can be concluded from the fact that the frequency of positive differences is higher than for negative differences. Apparently, the students are very tough when evaluating each other’s work.

3.4 Overall findings

Overall, our findings are that the consistency of the peer grades are quite good, however, it does depend on the difficulty of the report. More difficult reports seems to be

harder for the students to evaluate, which results in a larger variation in the peer grades awarded to the same report.

Also some questions in the evaluation rubric are clearly easier than others to evaluate by all students. Generic questions are easier for the students to evaluate compared to questions that require in-depth knowledge of the theory the report relies on.

Finally, the “professor’s grades” seems to be not much better than the students. In this case the professor had been a little nicer (resulting in higher grades) compared to the students, but there was generally a good relation between the peer grades and the professional grades.

The findings related to improving writing skills were inconclusive, which can mainly be attributed to the limited size of this study and to the fact the difficulty level of the two reports was very different making them hard to compare.

Guidelines for peer assessment derived from this:

- Reports based on a limited amount of theory are easier to evaluate by all students.
- Generic report characteristics such as structure and language are easier to evaluate by all students.
- Reports showing a large disagreement among peer graders are good candidates for being re-graded by the professor.

4 CONCLUSIONS AND FURTHER WORK

The findings are that the consistencies of the peer grades are quite good, and that there was generally a good relation between the peer grades and the professional grades. Furthermore, analysis of grading distributions from all assessments in the course has been made in order to investigate the improvement in report writing, i.e., the applicability of peer assessment as a learning tool. The report writing did improve marginally to some extent, but the reports are difficult to compare as they are very different.

However, this study is quite small as only 34 students participated. Therefore, the applicability of the results is also limited and the conclusions are not carved in stone. Instead they should be seen rather as an addition to the growing set of papers dealing with peer assessment.

An obvious further work is to continue the peer assessment in the coming years and compile all the results into a larger study (with a larger group of students) of peer assessment and its applicability as a learning tool.

REFERENCES

- [1] <http://www.kurser.dtu.dk/34330.aspx>
- [2] Biggs, J. and Tang, C., "Teaching for Quality Learning at University", Forth edition, McGraw Hill, 2011
- [3] Topping, K.J. , "Peer Assessment", Theory Into Practice, vol. 48, Issue 1, 2009
- [4] Andersson, P. H., & Hansen, C. T. (2014). "Experiential learning as a method to enhance experiential teaching". Proceedings of Sefi the 42nd Annual Conference.
- [5] Ng, Wing-Shui. "The Impact of Peer Assessment and Feedback Strategy in Learning Computer Programming In Higher Education." Issues in Informing Science and Information Technology 9 (2012): 17.
- [6] See www.coursera.org
- [7] Falchikov, N., & Goldfinch, J. (2000). "Student peer assessment in higher education: A meta-analysis comparing peer and teacher marks". Review of Educational Research, 70(3), 287–322.
- [8] Cassidy, S. (2006). "Developing employability skills: Peer assessment in higher education." Education and Training, Educ. Train, 48(7), 508–517
- [9] Crossman, J. M., & Kite, S. L. (2012). "Facilitating improved writing among students through directed peer review." Active Learning in Higher Education, Act. Learn. High. Educ, 13(3), 219–229.