

# Extreme Use of Student Peer Assessment

Lars Bendix, *Department of Computer Science, Lund University, Faculty of Engineering LTH*

**Abstract**—Students’ desire (and need) for ever more feedback and many students’ lack of good study techniques caused the author to dig into Student peer assessment (SPA). He wanted to discover if it could provide more feedback in a sustainable way and if it could provide students with a generally useful study technique.

Preliminary (bad) experience made him realize that he had only scratched the surface of SPA. This motivated him to dig much deeper to get a real understanding of the necessary prerequisites to make SPA work – and to explore if there were limits to how and for what SPA could be used.

After a more careful study of literature about SPA theory and experience, the author analysed two of his own courses for student deliveries that could be suitable for SPA. At the time of writing, the author has theoretical experience with creating SPA instances for different types of student deliveries. At the time of the Round Table, he will also be able to report practical experience from using concrete examples of SPA with students on three different courses.

In the Round Table discussions, the author would like to focus on the aspects of “cost-effectiveness” (for both students and teachers) and “invasiveness” of using SPA.

**Index Terms**—Student peer assessment (SPA), Feedback, Study techniques, Round table.

## I. INTRODUCTION

FEEDBACK is an essential part of all learning. It can be practical feedback in the way of burning your fingers when touching a hot pot on the stove, or at a more theoretical level when your teachers (or parents) tell you how you could have done things better. Feedback tells you whether you have understood something correctly – and if not, good feedback can make you reflect on why. That is why we constantly seek feedback – and students are no different.

I was very surprised when a colleague some years ago calculated that almost 80% of the resources on his course were spent on feedback and assessment – and still the students wanted more feedback. However, it wasn’t until I ended up in a situation where heavy feedback to students on two different courses happened at the same time, that I started thinking about how to solve the “feedback problem” in a smarter way than me working late hours.

Student peer assessment (SPA) seemed like the obvious solution – also because SPA wasn’t just about getting free labour from students since they themselves would learn from assessing other students’ work. So it would be a win-win situation. I had heard about SPA through colleagues and gave it a go on a course with a limited number of students.

However, initial experience was not very good and it caused me to reflect on the reason(s) why.

Further investigation into the “theory” behind SPA revealed some fundamental mistakes in the way that I had used SPA – mistakes that may be common also to other people. Furthermore, I discovered that I needed a better understanding of what is required to make SPA work. The investigation also triggered many ideas for how SPA could be put to good use on another course having more students and more types of “deliveries” from students that might be made “assessable”.

In the following, I will first give more motivation and background for using SPA followed by a short analysis of the possibilities for SPA occasions on two specific courses and a presentation of a number of actual proposals. Then I present preliminary experience and lessons learned from a more systematic use of SPA and some reflections on why things went how they did and what can be done to improve, before I finally draw some conclusions from this adventure.

Warning: Since this paper is “an invitation to a lively Round Table discussion”, it may contain: provocative statements, un-informed opinions, un-supported claims and non-motivated conclusions.

## II. MOTIVATION AND BACKGROUND

There are limits to how much feedback a teacher can provide to students in the hours he has for a course. This author hit that limit and looked to SPA for an answer to his problems. In this section, I will first describe what were the causes that triggered the interest in and need for SPA, then I will briefly present the theory underlying SPA and some experience from others.

### A. Initiating problem

For more than a decade, the author has given a course on “Configuration management” for up to 50 students per year. The course has changed a lot during the years and there has been more and more feedback on things that students produce during the course and after (the project is done after the exam). Students have been happy with the feedback they get, but have always asked for even more (and more detailed). This far, students get feedback on lab reports (4 reports from around 15 groups), individual paper reviews (from around 50 students), a synopsis (an “extended project proposal” - from around 15 groups) and a project report (from around 15 groups). All feedback is given by the course responsible.

Two years ago, the author took over on a course on “Coaching of agile teams” running in the same seven weeks as the above course. The previous teacher had given feedback on the essay that students (around 20) had to hand in every week. I wanted to continue that and “hit the wall” during the eight day where I had to provide feedback for 20

essays, 50 paper reviews and then 20 more essays – in parallel with teaching on the two courses. This was really a wake-up call. The way I had been giving feedback to students might have been appreciated, but it was not sustainable and certainly not scalable.

Last year, I did some preliminary experiments with SPA on the course “Coaching of agile teams” where the feedback on the first four essays were provided by the teacher and on the final three essays by the students. Experience was not that good and the goal for this year's “exercise” is to use SPA better and more systematically – and on both courses (with some preliminary “testing” on a “Configuration management” course taught all autumn at the IT University in Copenhagen). The claimed results of this systematic use of SPA would be: the same amount of feedback for less teacher resources (and in a way that is scalable), more student learning for very little extra student effort, and the mastering of a useful general study technique for the students.

### B. Student Peer Assessment

The present way we assess students and provide feedback has many problems. Davies et al. state that teacher-provided feedback does not scale well, they point out that much feedback is really wasted (given at the exam it is too late) and they call for a better alignment between assessment, learning and teaching [5]. Short feedback loops as a way of continuously working step-by-step towards perfection is a general technique that is used in many contexts. Feedback is one of the five fundamental values that eXtreme Programming [4] (a popular and widely used software development methodology) is built on. EXtreme Programming and its focus on frequent immediate feedback has served as inspiration for the teaching framework eXtreme Teaching [1].

Student Peer Assessment is a technique where students assess and provide feedback to other students on the same course [8]. SPA has been shown to heighten students' awareness of their work [9], to encourage students to take more responsibility for their own learning [6] and to develop the students' critical thinking [10]. For this experiment - and for the Round Table - my interest in SPA is more to use it for providing feedback than for assessment, and also to investigate its possibilities as a learning technique (for the assessing student).

Previous local experience at LTH has shown that students do better at exams when SPA is used (learn more?) [2], that their assessments can be trusted [3] and that it can also be carried out on large courses (200+ students) [7]. The setup that I have differ from the reported usage in that due to logistic difficulties, I will not use pair (but individual or group) SPA and feedback will not be immediate and face-to-face, but slightly delayed and via email.

## III. ANALYSIS AND PROPOSAL

In this section, I will first analyse what went wrong with my first attempt at using SPA and then analyse the two courses I am responsible for to discover their potential for future use of SPA. This is followed by a few sample proposals to show in a more concrete way how I intend to carry out SPA for different aspects of the courses.

### A. Problem analysis

Given that other people had reported successful use of SPA, I was determined to make it work for me too. So it was important to reach an understanding of why it had gone wrong for me and what are the underlying pre-requisites for making SPA work.

On the Coaching course SPA was used during the final three weeks for the weekly essay and during the following reading period for the preliminary project report and for the final project report. For the first four weekly essays I provided feedback, both in written form and orally at the following seminar where the essays were discussed. My intention had been to show the students different ways of providing feedback (I used different setups for different seminars). Unfortunately I had not been clear from the beginning that SPA would be used in the latter part, so students were not prepared to pick up inspiration for how to give feedback. Furthermore, there was no explicit assessment template (only the teacher's tacit knowledge) to help guide the students in what they should focus on in their assessment. For the preliminary and final project reports these problems had been dealt with and results were much better. However, there another tendency that showed both for the essays and the reports - students were better and more comfortable with giving oral feedback than written. In all cases feedback had to be given in written form, but for the first essay and for the final report the written this was supplemented by oral feedback (for the essays only a randomly selected set of students).

So for SPA to work two important things have to be in place: general explicit awareness and training in SPA for the students; concrete support from guidelines and assessment templates for each specific instance of SPA.

In order to explore the potential use cases for SPA, I will now analyse the two courses where I am responsible for what student deliveries there are. Besides the *essays* and *preliminary and final project reports*, the coaching students also hand in a *synopsis*. This is a 3-4 page extended project proposal and presently feedback is given by the teacher only (in written form due to logistic reasons). On the Configuration Management course there are more deliveries. There is an individual *paper review*, which could be compared to the essays on the Coaching course, though the requirements (and guidelines) for form and contents are more extensive and explicit. Students also produce four 3-4 page *lab reports* that are done in smaller groups. Since they carry out a project too, they have a synopsis and a final project report (no preliminary report is produced). Finally, they have six *exercise sessions* where they discuss a number of assignments in smaller groups and where the outcome is a plastic OH where they put down the most interesting result of their discussions (and where teacher feedback is given on a selected set at the following lecture).

### B. Suggested solutions

I will very briefly sketch the design of two specific instances of SPA to give a better idea of how it could look like. Both cases differ from [2], [3] and [7] in the usage of SPA in two ways: assessment was not done in pairs, but individually or in groups; assessment results were communicated with a time delay with respect to the

production of the delivery to be assessed.

The first is the group lab reports on the Configuration Management course. A separate webpage was created for the SPA and access to the webpage and its contents (the lab reports) was password protected for privacy reasons. On the webpage it was stated which group should assess which group and it was chosen to do that in a round-robin fashion since that scales to an odd number of groups (which there was on this specific instance of the course). There was also a link to the detailed assessment instructions that had to be followed – including how to communicate the assessment.

The second is the individual paper review on the Configuration Management course. The general outline was the same as for the lab reports (also to create "recognition effect" in the students). However, there were two things that required particular attention. Firstly, students have six different papers to choose from when doing the review and two students in the same groups must not review the same paper. So for assessing, I had two options: to let a student assess another student from the same group (but a different paper), or to let a student assess another review of the same paper. I chose the latter option since these students had no prior experience with doing SPA. Secondly, I had to consider how to scale the logistics of communicating assessment results from the present 10 students to the future 50 students. The pragmatic solution was to have students include their email in the review (which was the time consuming part of setting up the webpage).

#### IV. EXPERIENCE AND REFLECTION

In this section, I will report preliminary lessons learned and discuss some ideas for future steps in using SPA.

##### A. Lessons learned

At the time of writing I have experience with creating and using SPA instances on a course with 10 students at the IT University in Copenhagen. These students have no prior exposure to SPA and it had not been possible to make SPA a compulsory part of the course. The experience is based on one group lab report and the individual paper review.

Only one out of three groups carried out their assessment of a lab report. For the others, one group had not had time to do it yet and the other had completely overlooked the activity. Since the SPA occasion came in a period of "inactivity" on the course more should obviously have been done to signal that there were still things to be done. In the same period other courses that the students followed had pretty heavy deliveries, so more should have been done to convince the students that the time invested in doing the assessments would be amply returned.

For the individual paper reviews the results (and reasons) were almost the same: only three out of 10 students had carried out their assessment at the time of writing. However, the students' general attitude to this assessment was more positive, exemplified by this student comment: "I learned more from doing the assessment than from the feedback". This indicates that it is better to "sell" SPA to students from a "better learning" perspective than a "more feedback" one.

As for my personal experience with the workload from creating SPA instances there has been a rather heavy start-up cost in getting to grips with how SPA should be done to

work. However, for the specific instances actually created the workload has only been slightly higher than in previous years (primarily because there were quite clear "learning objectives" for the activities) and I expect that much can be reused next year (with a little editing of the webpages).

##### B. Discussion

For the future, I plan to implement the lab report and paper review on the course given at LTH. Scaling up to four lab report assessments should not cause "fatigue" in students and scaling up the paper review to 50 students should be sustainable for the teacher. Other activities will not be subject to SPA on this course. In part because that would mean 13 occasions of SPA in 7 weeks, in part because learning objectives for exercise sessions are (presently) not so clear. or the coaching course, I plan to repeat what I did last year - but done better this time. Since this course has a more "slow" pace (14 weeks) I will also use SPA for the synopsis and for four instead of three essays. In general, using SPA has also been beneficial for the course contents and teaching in that it revealed that learning outcomes and goals were not always that explicitly formulated.

#### V. CONCLUSION

There are many things that could be discussed regarding the use of SPA, but for this round table the author would like to focus on these two main points:

How do we make sure that it is **cost-effective** for both teachers and students to use SPA? If we cannot make that absolutely clear to the students it might be difficult to get real student buy-in. When students use a number of hours on SPA activities they should save the same (or more) number of hours on other activities.

What is the **limit** for how far can we go? In fact we double the number of "deliveries" that students have to produce when using SPA. How often can SPA be used in a 7-week period?

#### REFERENCES

- [1] R. Andersson, L. Bendix, "eXtreme Teaching – a Framework for Continuous Improvement", in *Computer Science Education*, Volume 16, Number 3, September 2006.
- [2] A. Axelsson, "Delmål och kamratgranskning – erfarenheter från en grundkurs i programmering", in *Proceedings of the 2:a Pedagogiska inspirationskonferensen*, Lund, Sweden, May 2004.
- [3] A. Axelsson, "Kan man lita på kamratgranskning?", in *Proceedings of the 3:e Pedagogiska inspirationskonferensen*, Lund, Sweden, May 2005.
- [4] K. Beck, "*Extreme Programming Explained – Embrace Change*", second edition, Addison-Wesley, 2005.
- [5] S. Davis, B. Smith, K. Sambell, "Can we improve feedback without additional resources?", *ASKe Pedagogy Research Centre*, Oxford Brookes University, 2007. Available: [http://www.brookes.ac.uk/askedocuments/s\\_davis.pdf](http://www.brookes.ac.uk/askedocuments/s_davis.pdf)
- [6] F. Dochy, M. Segers, "The use of self-, peer-, and co-assessment: a review", in *Studies in Higher Education*, Vol. 24, No. 3, 1999.
- [7] M. Kihl, R. Andersson, A. Axelsson, "Kamratgranskning i stora klasser", in *Proceedings of the Utvecklingskonferens*, Lund, Sweden, September 2007.
- [8] K. Kwang, R. Leung, "Tutor versus peer group assessment of student performance in a simulated training exercise", in *Assessment & Evaluation in Higher Education*, Vol. 21, No. 3, 1996.
- [9] P. Orsmond, S. Merry, K. Reiling, "The importance of marking criteria in the use of peer assessment", in *Assessment & Evaluation in Higher Education*, Vol. 21, No. 3, 1996.
- [10] C. Rust, "A briefing on assessment of large groups", LTSN Generic Centre Assessment Series No. 12, York, LTSN, 2001.