WEPS PEER AND AUTOMATIC ASSESSMENT IN ONLINE MATH COURSES

Olga Caprotti, Johanna Ojalainen, Matti Pauna, and Mika Seppälä

Department of Mathematics and Statistics,

University of Helsinki, Finland

P.O. Box 68 (Gustaf Hällströmin katu 2b)
FI-00014 University of Helsinki

olga.caprotti@gmail.com, johanna.ojalainen@gmail.com, matti.pauna@helsinki.fi, mika.seppala@helsinki.fi

INTRODUCTION

In this paper we discuss several aspects of a learning environment called WEPS (World Education Portals) and the MOOCS (Massive Open Online Courses) held there. Our main focus here is how to provide practicing and assessing mathematical competences in a way that is both pedagogically meaningful and scalable to large number of students.

In peer assessed workshops students first solve a set of problems, and submit their solutions to the system. Once the submission deadline has passed, students get a model solution, grading template, and are randomly assigned to grade the submissions of five (or more) fellow students. In the end, the system automatically grades the grading of the students. This is a very powerful concept in learning. For the grading of the submissions of other students, students need to learn to master the materials of the task at hand thoroughly.

In addition to the peer-graded workshops, we have used a system called STACK, which generates problems (out of templates) and step-by-step solutions to these problems. STACK problems contain random parameters. A single STACK problem template may generate thousands of problems. These are particularly effective in drills of mechanical tasks.

In STACK the system will automatically evaluate students response, and give meaningful feedback. Students can learn the materials from the feedback of sophisticated STACK problems.

The virtually unlimited practice provided by the STACK system boosts students' confidence on their abilities. Some students consider these quizzes as games taking them over and over until they have mastered every problem. When they do that, they learn.

Our initial findings indicate that these assessment tools improve learning outcomes. Students work more and the assessment tools boost their learning.

These initial findings have led Johanna Ojalainen to conduct a systematic study on the effect of automatic assessment among high school students. The study reviews student achievements, confidence (more precisely, self-efficacy), and their attitudes towards mathematics. We present the goals and the design of the study.

The resources used in this study are available to any instructor or student at https://myweps.com. Instructors may use, in their own instruction, this free Moodle service with its rich resources.

BACKGROUND

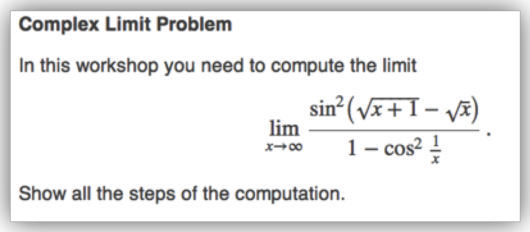
At the University of Helsinki we have over several years developed technology enhanced learning methods and materials in mathematics education as well as done research on the use of them. At 2005 a large scale EU funding was obtained to develop multilingual mathematics problem database. At that time we used the MapleT.A. system as an assessment platform. One drawback of MapleT.A., at least at that time, was its difficult integration with other learning management systems. This was one of the main reasons for moving to the open source STACK system which can be integrated with the open source Moodle learning environment. A large collection, about 2000, randomizable question templates was developed in the EU project which were converted from MapleT.A. into STACK.

Our group has actively disseminated results in the national, European and international levels. We have given talks and demonstrations of the use of the materials in teachers' conferences as well as given tutorials to students that are becoming mathematics teachers. We are actively seeking collaboration with others that are interested in using technology in science education.

PEER ASSESSMENT

In peer assessment students evaluate and give feedback to each' others work. Meaningful feedback is one of the key facilitators of learning. The purpose of feedback is to provide information about the product or understanding of a student for the purpose of helping to fill the gap between the current and desired level of understanding (Hattie & Timerley 2007). According to literature, peer assessment creates a lot of interaction between students. Furthermore, it is viewed that it is easier to give feedback to one's peers, but on the other hand students might not consider the feedback from other students as professional as from e.g. the instructor. (Boud 2000, Nicol & Macfarlane-Dick 2006, Hattie & Timperley 2007)

The open source learning environment Moodle provides a workshop activity module that enables giving homework assignments to be assessed by other students. The instructor can set various options, for example, allow self- or peer assessment, give example submissions, and set automatic allocation of the students' work for other students to assess. Students are divided into groups of five but this number can be changed. The workshop system allows giving assessment criteria, which is very important for the students to know how to assess others' work and understand the criteria under which their work is evaluated. An example of a problem used in WEPS Calculus I course workshop is presented in Figure 1.



Complex Limit Problem

In this workshop you need to compute the limit

$$\lim_{x\to\infty} \frac{\sin^2\left(\sqrt{x+1}-\sqrt{x}\right)}{1-\cos^2\frac{1}{x}}.$$

Show all the steps of the computation.

Figure 1: An example of a workshop task.

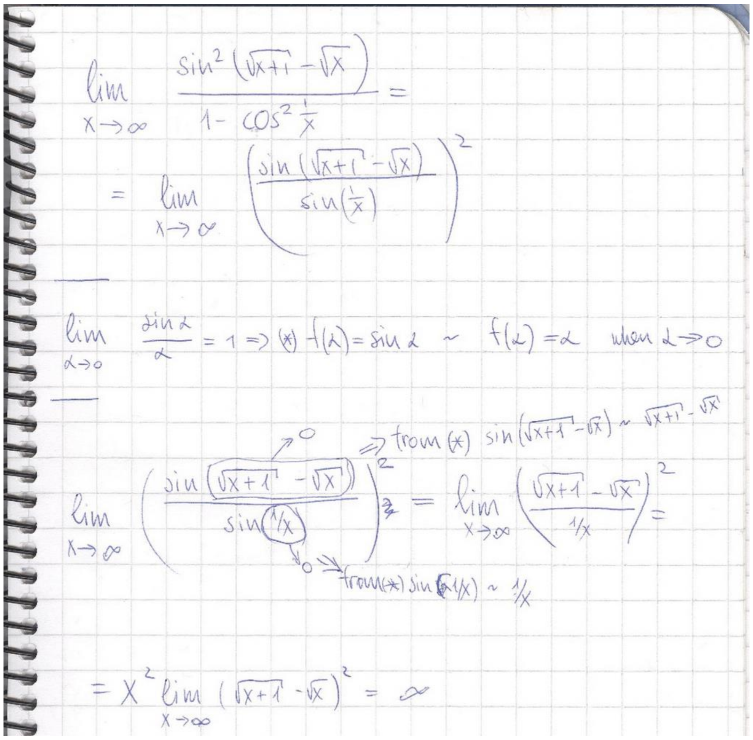


Figure 2: A student's solution to the workshop.

Students submit their solutions to the system as files or can use the mathematically enhanced text editor. They can e.g. write to paper, make a copy of it by scanning or taking a picture. In Figure 2 is an example of a student's submission.

After the deadline for submissions, the system automatically turns into assessment phase, where student submissions are allocated to other students, and assessment forms with the solutions as well as assessment criteria are available. Students are supposed to grade and give supportive feedback to other students' work, see Figure 3 for an example.

Aspect 2

Has the trigonometric formula $1 - \cos^2 \frac{1}{x} = \sin^2 \frac{1}{x}$ been properly used?

Grade 9 / 10

Comment You have used the Pythagorean theorem in the right place. On one hand you just use, i.e. you don't really explain anywhere why it is OK to change from 1-cos^2(1/x) to sin^2(1x/), but the Pythagorean theorem is rather well known, so it's OK.

By the way, I think it's better if you "over-explain" yourself, so in the future could you write down the reasons why you do certain things? Even small notes on the side of the page help the reader to understand what you are doing - or, at times, what you think you are doing.

Figure 3: An example of the feedback provided by another student.

When the assessment phase closes, the instructor can close the workshop after which the grades are calculated as averages of the five grades given by students.

Students have generally liked the peer assessment method. A survey was made in the mid-term of the Calculus I Spring 2013 course whose results are presented in Figure 4.

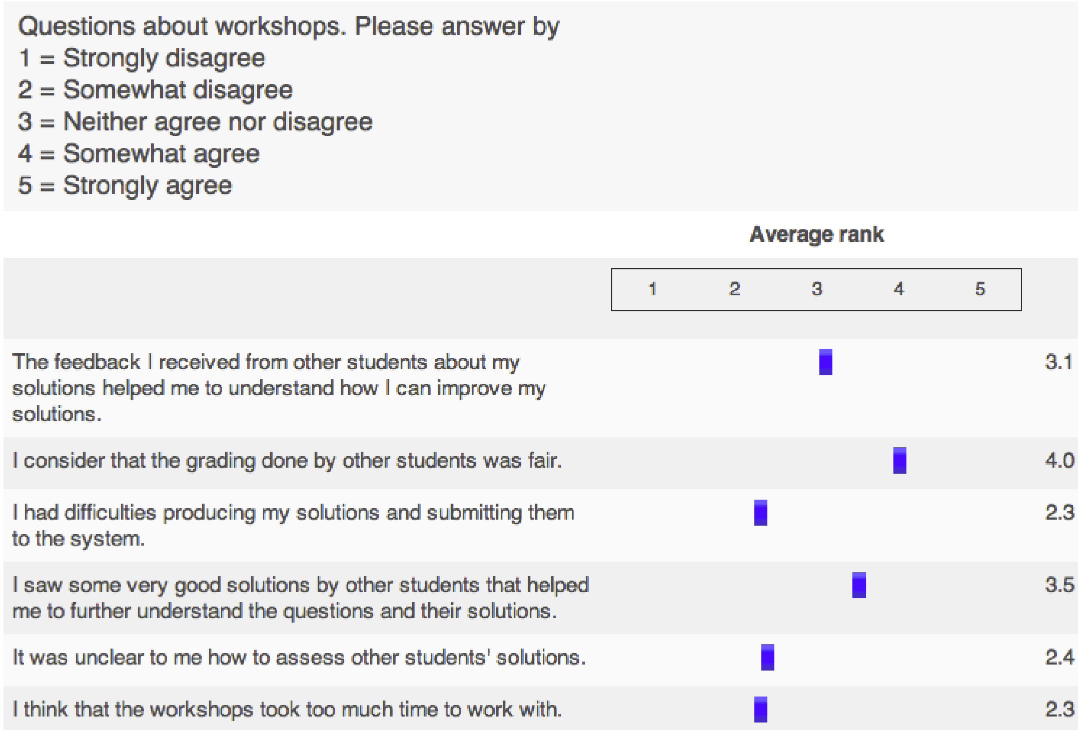


Figure 4: Students' feedback about peer assessment, n = 10.

Students consider that the grading done by other students is fair, that they learned from seeing other students' different solutions. Here is a spontaneous comment from a student about peer assessment, which captures the benefits and possible drawbacks of workshops in student's point of view:

> Workshops seems to be a great way to learn! When I graded other people solutions I think I really learned a lot about how the excercises should be proved. It really makes you think about the solutions when you have to grade someone else's work.
>
> Downside of the peer-grading is probably that it is too easy to be too strict on giving points and probably my bad english and bad math skills don't mix up very well with this. So hopefully no-one who receives my grading don't take it too seriously.
>
> Teachers: Do you have any suggestions how I should grade the future excercises. Is it better to be too strict or not? I tried to point out everything that I thought was not relevant or didn't make sense to me.
>
> One thing I noticed is that it's hard to decide what to do when one needs to give points on "did the reasoning contain any unnecessary parts" if I thought that the reasoning is incomplete or doesn't prove the initial statement at all. I think I gave full points almost every time but I'm not sure if this is the way to go.

Overall our experiences thus far with peer assessment are positive. We see the great potential of peer assessment in education because it enhances interaction and provides meaningful feedback to students in an online course as well as it is scalable to massive courses.

AUTOMATIC ASSESSMENT

Testing and grading mathematics skills automatically can dramatically reduce teachers' time taken in routine tasks. More importantly, it allows students to practice and get meaningful feedback at any time and pace they desire. Perhaps in terms of learning the automatic assessment systems are most powerful at formative assessment, i.e. assessments that supports learning while it happens. Summative assessment can also be successful in proctored situations.

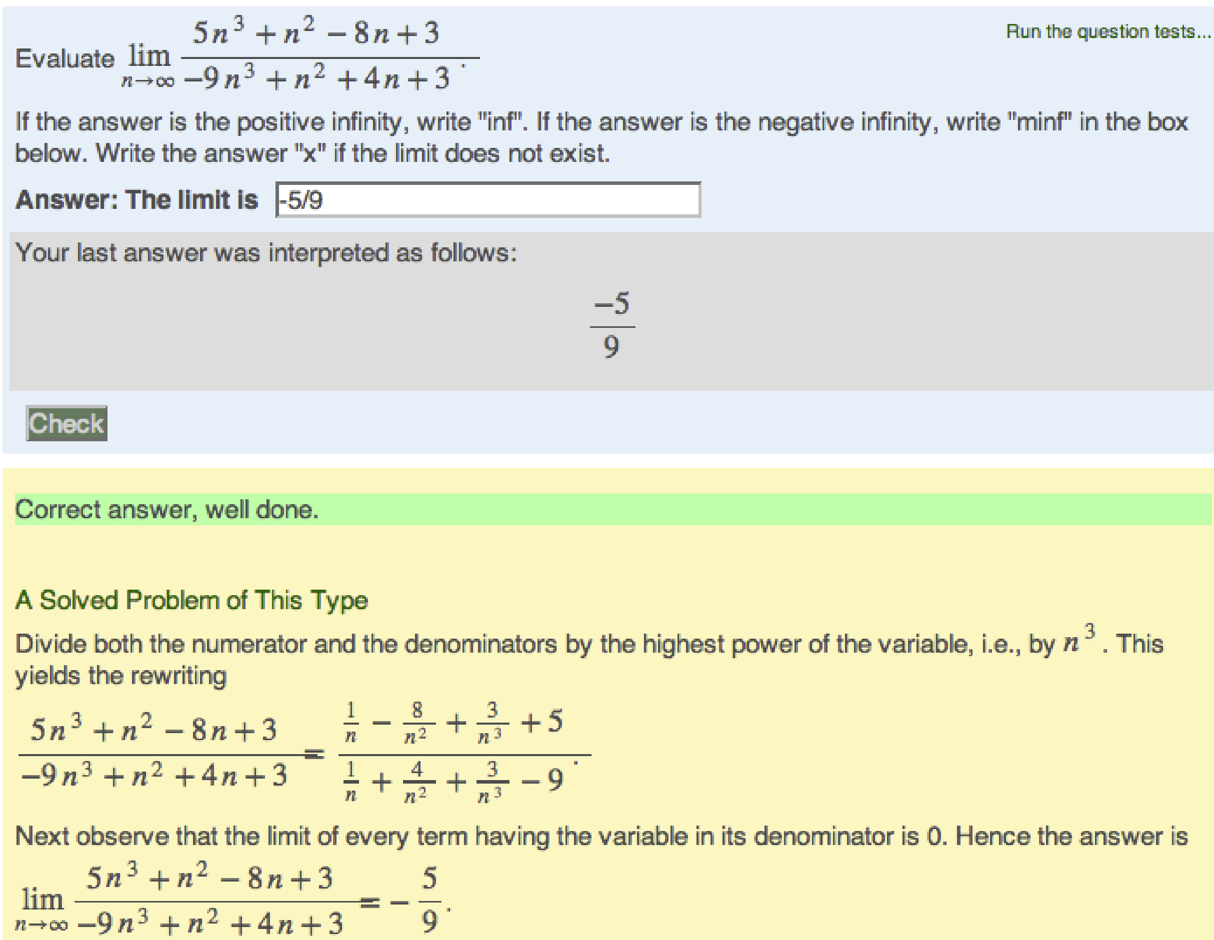


Figure 5: An example of a STACK problem and its feedback

Most important features of the STACK assessment system are its ability to generate random instances of a same question type, ability to not only to detect whether an answer is correct or incorrect but also if it is partially correct and its ability to give detailed feedback on the question and student's attempt. STACK is developed by Dr. Chris Sangwin at the Birmingham University (Sangwin 2008). An example of a STACK problem together with its solution is presented in Figure 4.

In mathematics, practice is mandatory for learning. In the beginning stages of college education, a student is required to learn the basic computational techniques (e.g. addition of fractions or simplifying algebraic expressions) before entering into more sophisticated fields (e.g. problem solving and mathematical argumentation). Automatically graded exercises allow practicing the basic methods at student's own speed and as many times until he or she has obtained mastery the material.

For many students it may happen, that they forget the basic mathematical toolbox when they move from e.g. high school to college. Also, mathematics might not be the most motivating subject matter for students who enter to study economics or engineering at college or university level. It can tie lot of resources of the institution to remedy the gaps in basic mathematical competences. Automatic assessment can first of all provide a diagnostic test where the basic knowledge of different mathematics areas are measured and by the results, the student can be led to take remedial practice exercises. With well-thought feedback the students can quickly relearn the basic techniques they were supposed already to master before entering the new school.

In online courses and especially in massive open online courses automatic assessment together with peer assessed workshops provide a feasible and scalable way for practicing and learning. For example, in a basic Calculus course all *homework assignments* can be implemented as automatically graded exercises. There are ongoing Calculus I and Calculus II courses at myweps.com where instruction is delivered as YouTube videos and practice problems as workshops and STACK exercises.

RESEARCH ON THE EFFECTIVENESS OF AUTOMATIC ASSESSMENT IN LEARNING

Initial findings based on our experience on using web-based exercises in teaching over the years show that they have the potential to enhance learning in many situations. These observations have led Johanna Ojalainen to conduct a Ph.D. research project to systematically study the effects of web-based exercises in learning, especially in students' self-efficacy beliefs and learning achievements.

This interdisciplinary research aims to categorize high school students according to in which ways they benefit from using web-based exercises. Moreover, we aim to find out how the computer-generated feedback builds students' confidence in their own abilities. The main goal is to identify the characteristics of students who get positive experiences of self-efficacy by learning mathematics using automatic assessment and feedback. By using results of this study we can further develop automatic assessment environment to improve mathematics and science learning. The design and goals of this study have been prepared in cooperation between the Department of Mathematics and Statistics and the Department of Teacher Education of the University of Helsinki.

This design research combines theoretical, developmental and experimental phases in cycles. The aim is to improve computer-aided teaching practices in high schools trough flexible and iterative design, development and implementation. In addition, we use quasi-experimental model (Cook & Campbell, 1979): students are divided into experimental and control groups. The experimental group consists of 150 students, ages 16-19. At the moment the control group consists of 32 students, ages 16-19. The experimental group learns by using web-based math exercises and the control group does their homework from a textbook. Survey data are collected by a questionnaire at the beginning and at the end of the course in both groups. The experimental group also answers questions about the usability of math exercises and the environment. The procedure is presented in Figure 6.

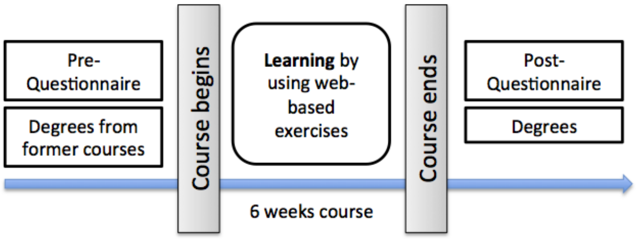


Figure 6: Design of the study

The questionnaire we use is based on Modified Fennema-Sherman Mathematics Attitude Scales, FSMAS, (Fennema & Sherman, 1976). Additionally we have created claims for measuring competencies of using computers and the Internet. We measure learning achievements as course degrees.

CONCLUSIONS

With well designed automatic assessment of basic mathematical techniques together with peer assessed workshops concentrating on more varied mathematical competencies, it is possible to develop self maintaining online math courses. Both of these assessment forms encourage students' constant practicing, self-regulation and interaction skills. The initial effort to produce such pedagogically rich learning materials may be high but once developed, can lead to courses that require only marginal monitoring and maintenance effort. The calculus and logic courses available at the moment at myweps.com are available for the use and further development for all instructors.

REFERENCES

A. Bandura, 1977, "Self-efficacy: Toward a Unifying Theory of Behavioral Change", *Psychological Review 84 (2), 191–215*

David Boud, 2000, "Sustainable Assessment: rethinking assessment for the learning society", *Studies in Continuing Education, Vol. 22, No. 2, 2000*

Cook, T. & Campbell, D. 1979. Quasi-experimentation: Design & analysis issues for field settings.

Fennema, E. & Sherman, J.A. 1976. Fennema-Sherman mathematics attitudes scales. JSAS Catalog of Selected Documents in Psychology 6, 31 (Ms. No. 1225).

John Hattie and Helen Timperley, 2007, "The Power of Feedback", *Review of Educational Research, Vol 77, No. 1, March 2007, pp. 81-112*

David J. Nicol and Debra Macfarlane-Dick, 2006, "Formative assessment and self- regulated learning: a model and seven principles of good feedback practice", *Studies in Higher Education Vol. 31, No. 2, April 2006, pp. 199–218*

Sangwin, C. J. 2008. Assessing Elementary Algebra with STACK. International Journal of Mathematical Education in Science and Technology, 38(8).