

Calculus Courses' Assessment Data

Matti Pauna

Department of Mathematics and Statistics,
University of Helsinki, Finland
matti.pauna@helsinki.fi

ABSTRACT. In this paper we describe computer-aided assessment methods used in online Calculus courses and the data they produce. The online learning environment collects a lot of timestamped data about every action a student makes. Assessment data can be harnessed into use as a feedback, predictor, and recommendation facility for students and instructors. We also describe late professor Mika Seppälä's seminal work at the University of Helsinki to develop online materials and tools for learning mathematics since 2001. He also utilized these methods in Calculus teaching at Florida State University. The open online course "Single Variable Calculus" was held in Helsinki in 2004. This intensive work evolved into a complete online English Calculus curriculum starting from the Fall 2013 and soon recognized as an alternative route for taking traditional university Calculus courses in Helsinki. Automatic assessment systems of mathematical competencies, such as STACK and WebWork, can take a student's answer as a mathematical object, e.g. a function or an equation, and check whether it satisfies the requirements set for a correct answer as well as give immediate and meaningful feedback. That is a powerful tool especially for formative assessment: log data shows that many students prefer to start with quizzes and when necessary, consult lecturing materials.

Keywords: Assessment data, automatic assessment, peer assessment, Calculus

1 INTRODUCTION

Work presented here is based on the seminal efforts by late professor Mika Seppälä (Xambó Deschamps, Bass, Bolaños Evia, Seiler, & Seppälä 2006) who developed the core materials and methods used in creating an online calculus curriculum at the University of Helsinki. Seppälä also developed online content and tools used in classroom teaching at the Florida State University. Seppälä also did pioneering work on digital representation of mathematical knowledge, most notably in the OpenMath project, whose results lead to the MathML encoding standard of mathematical formulas and their semantics. MathML is not only used to encode mathematical information in web pages but also in applications, such as Microsoft Office.

At the University of Helsinki, we started on 2001 to develop new methods for online teaching of calculus, especially for producing presentation materials tailored for the web and techniques to assess students' mathematical competencies online. The main goal at the time was to produce fully online calculus courses which materialized in 2004 as the online Single Variable Calculus course. Already then, the theory and examples were presented as ten-minute talk videos focusing on a single concept,

(2017). Calculus courses' assessment data. *Journal of Learning Analytics*, 4(2), 12–21. <http://dx.doi.org/10.18608/jla.2017.42.3>

technique, application, etc. Homework practice was implemented then with MapleTA online quizzes. In 2009, we moved to an automatic assessment system called STACK whose benefit is that it is fully integrated into the Moodle virtual learning environment (Caprotti, Ojalainen, Pauna & Seppälä 2013). Online assessment is described in more detail later in this paper.

Soon we saw the need for a complete track of calculus courses corresponding to the current traditional set of calculus courses. This led us to develop courses Calculus I, Calculus II, and Advanced Calculus, which together cover the traditional undergraduate calculus curriculum thus providing another path for students to study calculus online.

2 ONLINE ASSESSMENT FOR CALCULUS COMPETENCIES

When developing online assessments, it is worth considering the types of mathematical competencies that are to be assessed, for which there are many taxonomies. We have chosen as a basis the model developed by Pointon and Sangwin (2013) and extended by Rämö, Oinonen, and Wikberg (2015) by the category “Information transfer” which requires the students to represent the objects and information in many forms during working with an assignment. The categories in this taxonomy are as follows:

1. Factual recall
2. Carry out a routine calculation or algorithm
3. Classify some mathematical object
4. Interpret situation or answer
5. Proof, show, justify (general argument)
6. Extend a concept
7. Construct example/instance
8. Criticize a fallacy
9. Information transfer

Currently, our Calculus courses contain two types of digitally mediated assessments, which we discuss next.

2.1 Automatic Assessment

Online automatic assessment systems such as STACK (Sangwin, 2013) and WebWork (Gage, current issue) provide versatile ways of practicing with calculus problems. With STACK, one can ask questions whose answer is typed in as a mathematical formula. This is, thus, superior to multiple choice quizzes, for example. The main difference is that the student has to produce the answer rather than choose (by which ever strategy) from given possible answers. The system can generate problems from a template where parameters change each time the problem is deployed. It is able to analyze a student's answer and mathematically evaluate whether it satisfies the conditions required for a correct answer. Furthermore, it can recognize, through certain conditions, or patterns, if the answer is partially correct and provide feedback accordingly.

(2017). Calculus courses' assessment data. *Journal of Learning Analytics*, 4(2), 12–21. <http://dx.doi.org/10.18608/jla.2017.42.3>

In Calculus I and II courses, STACK problems are organized in practice quizzes which students can take as many times as they see necessary to master the topic. As the problems contain complete example solutions, students actually learn by taking quizzes as well. We can see from the log data that when studying online, some students prefer to start working with the quizzes and consulting theory when needed and other students go through the presentational materials before taking quizzes. Quizzes are also used as diagnostic tests, giving students an understanding of where they stand at the beginning of a course, and as practice tests for exams. We see that continuous and meaningful feedback is indispensable for learning in an online mathematics course.

The types of competencies that can be currently assessed by automatic quizzes are typically at the lower levels of the taxonomy, i.e. from factual recall to classify some mathematical object. However, tasks such as construct example/instance can be asked with this system. Figure 1 displays an example.

1. Give an example of an odd function by typing an expression which represents it. $f_1(x) =$. Tidy question | Question tests & deployed versions

2. Give an example of an even function. $f_2(x) =$.

3. Give an example of a function which is odd and even. $f_3(x) =$.

4. Is the answer to 3. unique? (Or are there many different possibilities.)

Figure 1: An example of an automatic quiz problem with an open answer

2.1.1 Automatic Assessment Data

The underlying learning environment (Moodle) stores detailed timestamped data from every user action and provides convenient views for students and instructors. The answer(s) to each problem are stored, as seen in Figure 2.

Response history					
Step	Time	Action		State	Marks
1	12/01/14, 17:48	Started		Not yet answered	
2	12/01/14, 18:09	Saved: ans1: $9*((v-2)*(1+(v-2/3)^2)^{1/2})/6+\operatorname{asinh}(v-2/3)/2+c$ [score]		Answer saved	
3	12/01/14, 18:18	Attempt finished		Incorrect	0.00

Figure 2: Log data saved from an answer attempt to a quiz problem

(2017). Calculus courses' assessment data. *Journal of Learning Analytics*, 4(2), 12–21. <http://dx.doi.org/10.18608/jla.2017.42.3>

This timestamped data enables instructors to see common misconceptions and to adjust their teaching accordingly. Large scale item analysis would show particularly useful problems for learning and also help in improving the problem bank.

The data from quizzes includes data for each student including their time taken and total score. Many students take a quiz several times, usually until they get a perfect score, which can be seen as contributing towards to the attribute of persistency in a student’s profile. Other student attributes that can be extracted using these data include how late before the deadline of the quiz they finish the quizzes (procrastination) and, of course, how many attempts they need to get the full score (mastery).

elo	Finished	12 January 2014 5:16 PM	12 January 2014 5:41 PM	25 mins 33 secs	4.80	✓ 1.20	✓ 1.20	✓ 1.20	✓ 1.20	✗ 0.00
elo	Finished	12 January 2014 5:48 PM	12 January 2014 6:18 PM	29 mins 27 secs	4.80	✓ 1.20	✓ 1.20	✓ 1.20	✓ 1.20	✗ 0.00
elo	Finished	12 January 2014 6:22 PM	12 January 2014 6:31 PM	8 mins 56 secs	0.00	✗ 0.00	✗ 0.00	✗ 0.00	✗ 0.00	✗ 0.00
elo	Finished	12 January 2014 6:33 PM	12 January 2014 6:47 PM	14 mins 26 secs	4.80	✓ 1.20	✓ 1.20	✓ 1.20	✓ 1.20	✗ 0.00
elo	Finished	13 January 2014 3:03 AM	13 January 2014 5:10 AM	2 hours 6 mins	6.00	✓ 1.20	✓ 1.20	✓ 1.20	✓ 1.20	✓ 1.20

Figure 3: Log data from a student’s attempt in a quiz

The part of a gradebook from a single quiz and single student shown in Figure 3 tells that this student has taken the quiz five times until the highest possible grade (6.00) is reached. The student’s name in the left most column is covered. This quiz contained five problems whose scores are marked in the five right most columns. The last problem has turned out to be the most difficult and an instructor can examine each of the student’s answers by clicking the score for each attempt. We see that the student has worked with the quiz over a course of two days. The last attempt took over two hours (fourth column from the left), significantly more than others. From the data that the system gathers we cannot find an explanation for that. The problems may have taken a lot of time for the student to solve or they might have had a break from online work. In any case, this student has solved altogether 25 problems on a certain topic in Calculus.

3 PEER ASSESSMENT

While automatic quizzes mainly focus on learning procedural skills, such as differentiation techniques, not all calculus competencies can (at the moment) be assessed automatically. Students also need to learn how to present their worked out solution and justify the methods used in the solution process. This is emphasized in the Advanced Calculus Course, where homework problems ask students to prove statements. We use peer-assessed workshops to accomplish this central part of the online study activity. Students are given a set of problems weekly where they have to present stepwise calculations or derivations of results. After that, students get to assess each other's solutions with the help of an example solutions and assessment guidelines. We see, and course feedback from students support this claim, that students often learn more deeply by assessing other students' work. They have to study the example solution and other students' solutions carefully to be able to assess them. Students are also required to give constructive and corrective feedback when a solution needs improvement.

Peer workshops provide a versatile way to ask many types of mathematical questions as the solutions are assessed by humans together with assessment instructions. Also technical limitations of entering solutions are diminished because many students submit a photograph of their written papers. However, it is not worth asking simple procedural tasks, as automatic assessment can handle those and it is not always instructive for the peer to assess simple one expression answers. Furthermore, students report that they learn from following other students' reasoning and the steps of their work and augmenting their work for a correct solution when necessary.

3.1 Peer Assessment Data

From peer assessment we get the usual timestamped data, such as when students submit work, when students assess others, and the scores they give. For instance, it is interesting to see that the scores the students give to a submission are surprisingly similar. That is important for the students to feel that the assessment is fair even though it is not done by the instructor, who has designed the assessment instructions and oversees the process.

Students' submissions are mostly digital images of their handwriting with mathematical formulas like the example shown in Figure 4.

$$\lim_{x \rightarrow \infty} \frac{\sin^2(\sqrt{x+1} - \sqrt{x})}{1 - \cos^2 \frac{1}{x}} =$$

$$= \lim_{x \rightarrow \infty} \left(\frac{\sin(\sqrt{x+1} - \sqrt{x})}{\sin(\frac{1}{x})} \right)^2$$

$\lim_{\alpha \rightarrow 0} \frac{\sin \alpha}{\alpha} = 1 \Rightarrow (*) f(\alpha) = \sin \alpha \sim g(\alpha) = \alpha \text{ when } \alpha \rightarrow 0$

$$\lim_{x \rightarrow \infty} \left(\frac{\sin(\sqrt{x+1} - \sqrt{x})}{\sin(\frac{1}{x})} \right)^2 = \lim_{x \rightarrow \infty} \left(\frac{\sqrt{x+1} - \sqrt{x}}{\frac{1}{x}} \right)^2 =$$

$$= x^2 \lim_{x \rightarrow \infty} (\sqrt{x+1} - \sqrt{x})^2 = \infty$$

Figure 4: A solution to workshop problem submitted by a student

Using these images makes it rather difficult to automatically extract information from the solutions. There are a small number of students who submit their work as pdf documents written by word processors or even LaTeX. (For compatibility reasons between the different computer systems students use, pdf is the only acceptable document type.) However, the feedback students give is input in a text area in the assessment sheet. An example is shown in Figure 5.

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 $\sin^2(1/x)$, 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 5: Assessment feedback and score given by a peer

This textual data opens many possibilities for textual analysis methods, for example whether the feedback is purely factual or technical, does it contain negative/positive expressions, does it address the other student directly with a pronoun or talk in the passive mode, etc. Results of these analyses can be combined with the numerical data from scores or time taken to write the assessments. There is of course an inherent underlying problem of measuring the “time taken on a task” since for information we only have the timestamp of the starting of the process, timestamps of the actions inside the task (or ISSN 1929-7750 (online). The Journal of Learning Analytics works under a Creative Commons License, Attribution - NonCommercial-NoDerivs 3.0 Unported (CC BY-NC-ND 3.0)

timestamps of actions outside the task while student has not ended the task) and the timestamp of finishing the task. A student can very well take breaks, leave a task open for the next day, etc, and therefore the measurement of time taken on a task has some noise.

4 CALCULUS COURSE DATA

Combining all of the above data from assessments with other clickstream data enables us to find study paths, i.e. sequences of actions taken by a student while working in an online Calculus course. Our courses are modelled with the paradigm, whose focal points are workshops, shown in Figure 6.

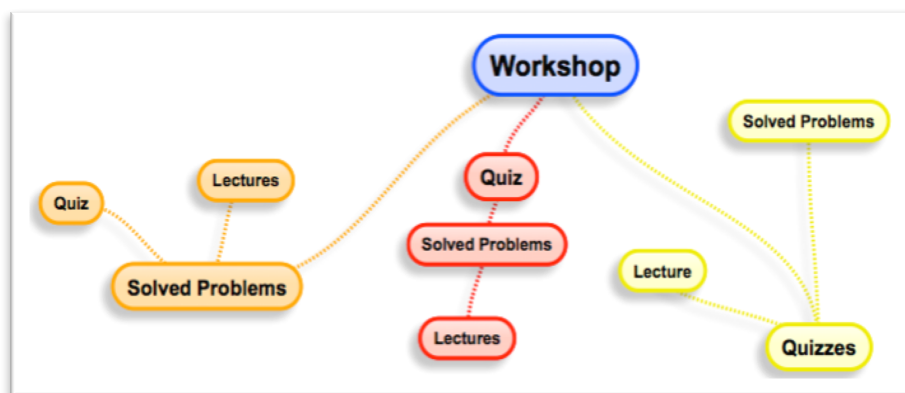


Figure 6: Calculus course architecture supporting individual student learning paths

The model shown in Figure 6 allows a student to follow their preferences for which order to study the resources and tasks. For instance, a seemingly traditional way would be to study the theory, then look at examples and applications, take a quiz to check learning, and finally submit the workshop problems. On the other hand, we see that many students first start taking quizzes, perhaps to check whether they already master the subject (e.g. from previous studies) and consulting written materials when needed. All the activities are aimed at supporting students to solve the more involved workshop problems.

All these logged activities enable us to build study paths, i.e. a click stream data of an individual student of their actions while working on an online course. Graded activities define control points in the path: we see what kind of materials a student used before taking a quiz or workshop and how successful they were at the task. This gives us a method of assessing the quality and relevance of materials. But most importantly, combined with data of the student profile we could see whether that path was effective for that type of student. By analysing all this data and extracting models using machine learning or topological methods, it would be possible to build a system that gives feedback and recommendations to a student as to which turns make in their study path.

5 CALCULUS COURSES AT THE UNIVERSITY OF HELSINKI

Finally, we describe teaching of online calculus courses at the University of Helsinki, as the author has access to that data. Online Calculus is offered as an alternative path to the more traditional and popular lectured courses. As the courses are offered in English they are popular among exchange students and e.g. students who have limited access to the campus. We run Calculus I, Calculus II, and Advanced Calculus courses that together correspond to the traditional Calculus curriculum. The rotation is so that Calculus I and Advanced Calculus are run in the Fall and Calculus II in the Spring allowing the student to take these courses in three consecutive semesters. The first full rotation (Advanced Calculus was added next Fall) started Fall 2012. As of end of Fall 2015 altogether 390 student registrations have been in these courses. 110 students have taken the midterm exam of any of these courses. 82 students have passed a course. We list some descriptive statistics of various types of scores in these courses in Table 1:

Table 1: Descriptive statistics of scores for different types of assessments in Calculus Courses

	N	Min	Max	Mean	SD
Course Grade	82	0.00	5.00	3.32	1.34
Exam 1	82	7.00	24.00	19.52	4.12
Exams C Total	80	18.00	47.00	34.60	7.67
Quiz C Total	82	0.00	100.00	79.73	18.84
Workshop C Total	82	9.36	96.03	65.12	18.56

The maximum for course grade in Helsinki is a 5 and a student must earn 75 percent of the total points to achieve it. For the lowest passing score, a 1, a student must earn 50 percent. The midterm exam (Exam 1) and the final exam each give maximum of 24 points totalling 48 across both exams (Exams C Total), a maximum score that none of the 80 students got. Here, quizzes are the automatic assessment part of the course and workshops are the peer assessment part.

It is interesting to ask which type of assessment activity might be most beneficial for learning Calculus. We understand that no conclusions from this data cannot be yet made and more data is needed. Some initial data relating performance on quizzes and workshops to exam scores are pictured in Figure 7.

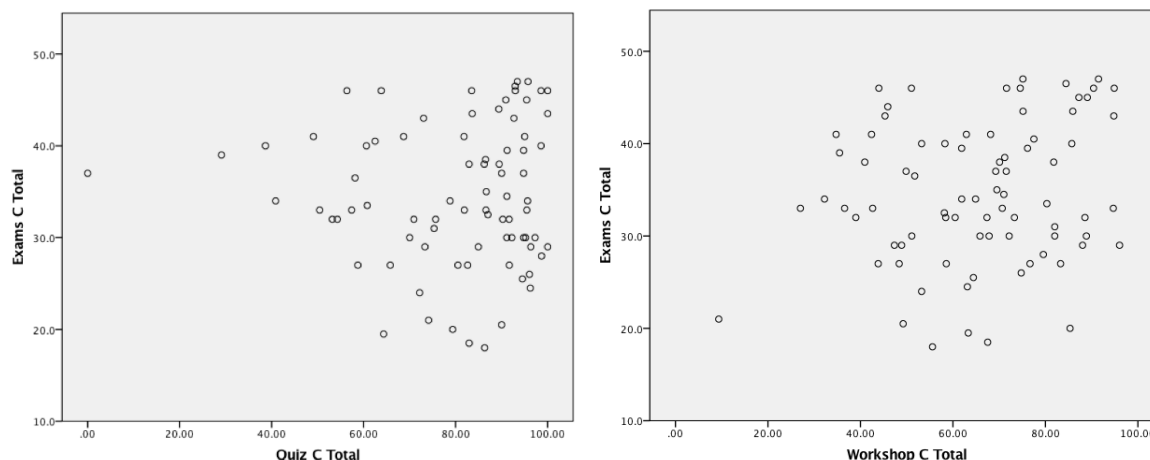


Figure 7: Scatter plots for correlating Exam scores with Quiz scores (left) and workshop scores (right).

Correlation coefficients show that there is no correlation between quiz and exam scores ($r = .01, p = .95$). This may be because a quiz can be retried until full points are achieved. There is also a non-significant relation between workshops and exams ($r = .19, p = .10$), but the lack of statistical significance is likely due to the small sample size and with more data this relation may be important. There is a significant correlation between quiz and workshop scores ($r = .43, p < .001$), likely because quizzes were planned to prepare for solving workshop problems.

6 CONCLUSION

In this paper, we have described online Calculus courses and computer-aided assessments used in them, namely automatic assessment with STACK quizzes and peer reviewed assessment. These assessments can be also thought of as homework or practice. In the first type, students get automatic, immediate, and meaningful feedback from the computer, and in the second type, students get to see authentic solutions from other students and give and receive constructive feedback towards a better solution to a math problem.

We have started an initial analysis of various types of assessment and course activity data that can be gathered from the online learning environment. Studying these data should enable us to find best ways of using e-assessments to support learning in online courses.

REFERENCES

- Caprotti, O., Ojalainen, J., Pauna, M., & Seppälä, M. (2013). WEPS Peer and Automatic Assessment in Online Math Courses. *Electronic Proc. 21st Int. Conf. Technology in Collegiate Mathematics*. Retrieved from <http://archives.math.utk.edu/ICTCM/i/21/S026.html>
- Gage, M. (2017). Methods of interoperability: Moodle and WeBWorK. *Special Section: Shape of Educational Data Proceedings. Journal of Learning Analytics*, 4(2), 22–35. <http://dx.doi.org/10.18608/jla.2017.42.4>

(2017). Calculus courses' assessment data. *Journal of Learning Analytics*, 4(2), 12–21. <http://dx.doi.org/10.18608/jla.2017.42.3>

Pointon, A., & Sangwin, C. J. (2003). An analysis of undergraduate core material in the light of hand-held computer algebra systems. *International Journal of Mathematical Education in Science and Technology*, 34(5), 671–686. <http://dx.doi.org/10.1080/0020739031000148930>

Rämö, J., Oinonen, L., & Vikberg, T. (2015). Extreme Apprenticeship – Emphasising Conceptual Understanding in Undergraduate Mathematics. In K. Krainer & N. Vondrová (Eds.), *Proceedings of the Ninth Congress of the European Society for Research in Mathematics Education* (pp. 2242–2248). Prague: Charles University in Prague, Faculty of Education and ERME.

Sangwin, C. (2013). *Computer Aided Assessment of Mathematics*. Oxford University Press.

Xambó Deschamps, S., Bass, H., Bolaños Evia, G., Seiler, R., & Seppälä, M. (2006) e-Learning Mathematics. *Proc. Int. Conf. of Mathematicians*. Madrid, Spain: ICM.