Improved engagement and learning in flipped-classroom calculus

Mikaël Cronhjort†*, Lars Filipsson‡, Maria Weurlander†

†Department of Learning, School of Education and Communication in Engineering Science, KTH Royal Institute of Technology, Stockholm, Sweden and ‡Department of Mathematics, School of Engineering Sciences, KTH Royal Institute of Technology, Stockholm, Sweden

*Email: mikaelc@kth.se

[Submitted October 2016; accepted April 2017]

We report on an effort to measure the effect of replacing traditional lecture-based teaching in calculus with a flipped-classroom approach. We base the comparison between the two teaching models on data from three sources: (a) a Calculus Baseline Test, designed specifically for this purpose and given as pre-test and post-test; (b) a survey measuring student engagement; and (c) student achievement on the final exam. On the Calculus Baseline Test, we found that the normalized gain was 13% higher in the flipped-classroom group. Similarly, the flipped-classroom group scored significantly higher on the engagement survey. Also, the students of the flipped-classroom group performed much better than expected on the final exam of the course, with a substantial decrease in failure rate.

1. Introduction

Interactive teaching where students participate actively through discussions or solve problems or other tasks, has become more and more common in recent years. Flipped classroom is an interactive-teaching model that has been successfully introduced in university mathematics (Jungic et al., 2015; Maciejewski, 2015; Murphy et al., 2015; Petrillo, 2015). The idea behind flipped classroom is that the information and content instruction takes place outside of class time, and class time is instead used for interactions and conceptual development, e.g. through peer–peer discussions, solving problems and applying conceptual knowledge (DeLozier & Rhodes, 2016). Although flipped classroom can be implemented in various ways, many who use it are motivated by a desire to increase students’ time for problem solving (Naccarato & Karakok, 2015). The social dimension, that students discuss or work in pairs or groups, seems to be an important aspect of interactive classrooms (Naccarato & Karakok, 2015; Weurlander et al., 2016a). The design of the flipped-classroom model is supported by evidence that students perform better in active learning classes compared to traditional lectures classes in the STEM (science, technology, engineering and mathematics) subjects (Freeman et al., 2014; Deslauriers et al., 2011; Mazur, 1997; Hake, 1998). Moreover, videos seem to be as good as ordinary lectures when it comes to presenting course content (DeLozier & Rhodes, 2016).

The literature on student learning commonly agrees that students learn when they are actively engaged in meaningful learning activities (Entwistle, 2009; Wimpenny & Savin-Baden, 2013).
These learning activities involve cognitive, affective and behavioural aspects (Trowler, 2010). Moreover, learning and active engagement is complex, involving both individual and social dimensions, as well as the wider context in which learning takes place (Kahu, 2013). However, students sometimes find learning troublesome and they struggle to understand (Weurlander et al., 2016b; Wimpenny & Savin-Baden, 2013). Interactive teaching has the potential to facilitate student engagement (Weurlander et al., 2016a). Also, student engagement has been found to be positively linked to student performance, especially for weaker performing students (Carini et al., 2006).

Several studies report on positive experiences of using flipped classroom in calculus. In a study by Jungic et al. (2015), students believed that the video lectures were helpful to prepare for class and they appreciated that they could learn at their own pace. Similar findings have also been reported by other authors (Cronhjort & Weurlander, 2016; Love et al., 2014). One student explicitly mentioned that ‘many times I have wished to pause or rewind a live lecture. You can do that with a video without disrupting the flow of the class. Also, watching a second time really reinforces the concepts’ (Love et al., 2014, p. 323). In class, answering clicker questions on mathematical problems, was also appreciated. The opportunity to first think for themselves, then discuss with peers seems to help many students to learn (Weurlander et al., 2016a; Jungic et al., 2015).

Do students learn more in flipped mathematics classrooms? There are some studies that indicate that they do. Murphy et al. (2015) concluded that students in a flipped linear algebra class scored considerably higher on the final exam compared to students in a traditional lecture class. A closer look at two questions on the final exam revealed that students in the flipped class explained more in their solutions compared to students in the lecture class. This suggests that students in the flipped class not only solved the problems but understood what they were doing and why to a greater extent. Similarly, Petrillo (2015) reported on an increase in students’ achievement in the flipped-classroom setting. Also, the study showed a decrease in failure rate in the flipped classroom compared to the lecture setting. In line with these findings, Love et al. (2014) reported that students’ average change in exam score was significantly higher for students in the flipped class than for students in the traditional lecture environment. Findings in a more large-scale study, involving 690 students in first-year calculus for life science majors, corroborate this pattern: students in the flipped class outperformed students in the traditional lecture class (Maciejewski, 2015). However, another recent study showed that students’ performance were only slightly better in the flipped class than in the traditional lecture settings, and this difference related more to conceptual rather than procedural problems (Wasserman et al., 2015). Moreover, it has been found that students with good basic mathematical skills, but with low knowledge in calculus seem to benefit the most from the flipped-calculus classroom (Maciejewski, 2015).

At KTH Royal Institute of Technology most engineering students study the course SF1625 Calculus I. Students are divided into several lecture groups. The course contains 42h lectures, 26h tutorials and 12h seminars. Study programmes taking the course simultaneously have a common examination. Typically, the failure rate at the written exam is between 15–50%, depending on the study programme. In traditional teaching of calculus at KTH, lecturers present the major theorems and proofs of the course and provide some examples illustrating how the theorems can be applied. There are, of course, significant differences between different lecturers in how they use examples or stimulate communication with their students, but common is that little time is spent on peer discussions and students rarely are given time to work hands-on with the materials.

In this setting, we have conducted some experiments with interactive teaching in calculus since 2012. In one lecture group, we began using a teaching model based on peer instruction (Mazur, 1997). The model was adapted to mathematics, by increasing the diversity of clickers questions: some had focused on understanding concepts, but others on performing limited calculations or understanding proofs. Although this was our first experiment with peer instruction, we received some encouraging
results indicating improved student learning (Cronhjort et al., 2013). In a survey, we asked about the students’ experience of the teaching method, and found that the engagement and alienation were meaningful concepts for describing their relationship to the teaching method (Weurlander et al., 2016a). Many students were enthusiastic about learning calculus in the peer-instruction setting; others were more sceptical and preferred a mix of traditional lectures and peer-instruction sessions. However, some students were frustrated and found learning difficult using peer instruction, possibly due to higher demands and that their preferred view of how calculus should be taught and learned was challenged (Weurlander et al., 2016a). The following year, preparatory films were included in the course in order to facilitate students’ preparations for the interactive in-class activities. The films had integrated online quizzes on the web platform Scalable Learning (www.scalable-learning.com). An investigation, based on focus-group interviews of students with experience of different implementations of flipped classroom, identified five key elements that were perceived as important by the students: preparatory films with focus on basic concepts, quizzes connected to the films to stimulate critical thinking, individual responses to the quizzes that make students feel seen and encouraged to prepare, in-class interactivity offering challenges to performance and feedback and suitable degrees of difficulty providing confidence as well as challenges (Cronhjort & Weurlander, 2016).

The research reviewed above shows that flipped-classroom models can have a positive impact on student learning in calculus. Still, little is known about the significance of different components or implementations of flipped classroom, how flipped classroom influences learning, and how different students are affected. There has been a call for more research investigating students’ performance in flipped classrooms compared to traditional lecture classrooms, especially using objective measures (Naccarato & Karakok, 2015; DeLozier & Rhodes, 2016; Maciejewski, 2015). The study of this article is a contribution to this research.

The aim of the present study was to compare students’ learning between a flipped classroom implementation and traditional teaching. Specifically, we investigated quantitative results from a Calculus Baseline Test given as pre- and post-test, data from a survey focusing explicitly on student engagement, as this seemed to be connected to improved learning in flipped classroom, and results from the final exam of the course.

In the fall semester of 2015, eight study programmes were taking the course in calculus simultaneously. Lecturers were assigned to different study programmes by the director of studies, and based on lecturers’ preferences four of the programmes were offered interactive teaching and the other four were offered traditional lectures. Thus, we obtained 226 registered students from four study programmes taught in a flipped-classroom setting using interactive engagement sessions, and 413 students from four study programmes taught primarily by lectures in three lecture groups. All other parts of the course were kept the same, e.g. textbook, tutorials, seminars and final exam. The examination was composed of nine open problems and full solutions were required.

2. Interactive teaching sessions

For students in the interactive group, we used a teaching format that included elements of both peer instruction and flipped classroom.

To begin with, the students were required to watch a short online video introduction before each lecture. In the videos, we had integrated quizzes on the material. The quizzes gave feedback to both students and teacher on the level of preparation before the class.

In class, a substantial amount of time, at least 50%, was spent on active student work or student discussions. Most of the remaining time was used to sum up these student exercises and student discussions.
Typically, a lecture, or rather teaching session, would start off with a short introduction and maybe also some clarifications regarding the preparatory video. The students could pose questions in the video forum, and some of these questions were excellent starting points. This would normally take 5–10 min. After that, the students were put to work. Sometimes they were then given multiple-choice questions to discuss and answer, much in the style of peer instruction. Sometimes they were asked to solve exercises or even applied problems, using pencil and paper, and discuss their ideas and computations with each other. As they did so, the teacher walked around and listened to the student discussions and answered questions. After such an exercise, there would always be a thorough interactive but teacher led review to the solution of the problem, sometimes including the underlying theory. On average the full cycle, of the students solving a problem and the teacher reviewing it on the blackboard, would take ~15 min, and so in a 90-min lecture there was time for about five or six such problems.

In the course, there were also problem-solving sessions where typically a teaching assistant would solve problems on the blackboard, and seminars for which the students needed to prepare solutions to various problems and where the students were supposed to be active. Tutorials and seminars were run in the same way for all students in the course and, hence, were not changed in any way for the students that were offered interactive-teaching sessions instead of lectures.

3. Method of investigation

The investigation of the learning and the engagement of the different student groups was conducted in the fall term 2015 as follows.

In order to measure the learning taking place in the different student groups during the course, we designed a Calculus Baseline Test, consisting of 15 multiple-choice questions divided into three subcategories: pre-calculus, calculus concepts and calculus theory and formalism. This test was given to the students both as a pre-test, at the beginning of the course, and as a post-test, at the end of the course. For each student taking the pre-test as well as the post-test, we calculated the normalized gain $g$ in the standard way, as

$$g = \frac{\text{post-score} - \text{pre-score}}{100 - \text{pre-score}}$$

i.e. the improvement divided by the maximum possible improvement (the grading was done on a scale from 0% to 100% correct answers). The average gain was calculated for the flipped-classroom group as well as the control group, and used as an indication of how much was learnt in the respective groups during the course. This procedure is similar to the way the Force concept inventory has been used in physics education research (Halloun & Hestenes, 1985).

At the occasion of the post-test, we also administered a survey aimed at measuring the level of engagement of the students in the different groups. Our underlying assumptions of engagement are in agreement with Kahu (2013), and we view ‘student engagement as a psycho-social process, influenced by institutional and personal factors, and embedded within a wider social context’ (Kahu, 2013, p.768). We designed our own items, relevant for our context, but building on previous research (Gunuc & Kuzu, 2015; Kahu, 2013; Handelsmann et al., 2005) and similar surveys (ISSE, SCEQ and LEQ).1 Our survey consisted of 15 items, in the form of statements, and the students were asked to rate to what degree they agreed with those statements, on a scale from one to five, where one meant they totally

1 ISSE: Irish Survey of Student Engagement (studentsurvey.ie), SCEQ: Student Course Engagement Questionnaire (Handelsmann et al., 2005), LEQ: Learning Experience Questionnaire (developed by Dan Borglund at KTH).
disagreed and five that they completely agreed. The items covered three themes as follows: (a) cognitive engagement, (b) emotional engagement and (c) active participation. The first theme concerned students’ intellectual activity, i.e. whether they worked hard, tried to understand, solved problems, etc. The second theme involved statements about whether the students were interested, felt motivated to learn or if they felt that the teacher valued them and their effort. The third theme included items regarding student interactions with teachers and other students. In order to compare the engagement level between the flipped-classroom group and control group, we applied the Mann–Whitney U test to identify significant differences, and we calculated the mean engagement ratings for the different groups, respectively. In this case, means were used for describing the differences between the groups instead of medians, as steps between the answer alternatives are large compared to the variation.

Finally, the results on the final exam of the course were analysed, regarding failure rate and the best grade A. Because of inherent differences in aptitude and pre-knowledge of students at the various programmes, it is not relevant to compare the different programmes with each other: programmes with high-admission requirements would normally outperform programmes with lower requirements, regardless of teaching method. Instead, we compared the grades on the final examination with the grades on the final examination previous year for the different programmes, respectively. The previous year, the teaching was lecture-based for all corresponding study programmes.

Participation in our investigation was voluntary for the students, and about 30% of the students chose not to participate in the Calculus Baseline Test and in the engagement survey. Even though the situation was similar in both groups, and there is no apparent reason to believe that there is some systematic difference in the missing data, this loss of data nevertheless makes it difficult to draw any statistically significant conclusions. Also, the variance is quite large in our datasets from the baseline and engagement tests. On the other hand, when it comes to the final exam, all students in the respective groups are included in the study.

4. Findings

The normalized gain on the Calculus Baseline Test was 13% higher in the flipped-classroom group than in the control group. Table 1 shows the normalized gain and the number of students doing the pretest as well as the post-test in the different groups.

<table>
<thead>
<tr>
<th>Calculus baseline test (n)</th>
<th>Average normalized gain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flipped (79)</td>
<td>46.5</td>
</tr>
<tr>
<td>Lecture (165)</td>
<td>41.0</td>
</tr>
</tbody>
</table>

There were 15 items on the engagement survey. Table 2 shows the items, results of the Mann–Whitney U test, and the mean results for each item of the engagement survey for the different student groups. According to the Mann–Whitney U test, there are significant differences ($p < 0.05$) between the two groups on most items (except item numbers 2, 4, 5, 11, 15). Item numbers 1–6 are statements regarding intellectual engagement, item numbers 7–11 emotional engagement and item numbers 12–15 about participation. Figure 1 shows a chart of the mean results for each item. An answer value 1 represents I don’t agree at all, and 5 means I fully agree. The difference between the two groups is greatest for item numbers 9, 10 and 13. These items are all about the relationship between teacher and students.
The results on the final examination in 2015 were compared to the corresponding results of the previous year (2014) when all involved study programmes had lecture-based teaching. Table 3 shows the failure rate in 2014 and 2015 for the flipped-classroom group (using flipped classroom in 2015) and the control group using lectures. The intervention group performed significantly better in 2015 with flipped classroom. The failure rate on the final examination decreased by 56% (from 40.6% to 17.7%). The control group also performed better, but the improvement in this case was only 28.7% (from 25.1% to 17.9%).

The top students also seem to have improved their results with flipped classroom. In the flipped group, the proportion of students receiving the best grade A increased by 37.7% (from 10.6% to 14.6%). The proportion of grade A in the lecture group increased by 22.2% (from 15.8% to 19.3%).

It is interesting to note that both the flipped group and the lecture group performed better in 2015, even though the improvement is bigger in the flipped group. The general improvement from 2014 to 2015 could indicate that students were better motivated or prepared, or that the examination was easier in 2015.

5. Discussion and conclusion

Both the intervention group (flipped classroom in 2015) and the lecture group are composed of several study programmes. As described above, it is difficult to compare different programmes. Here, details
concerning different study programmes are not discussed, but we intend to study this in future, where the results obtained in 2016 will also be included.

In this implementation of flipped classroom, the interactive sessions not only included multiple-choice questions, but also exercises or applied problems that were solved using pencil and paper. In addition to these activities designed to promote interaction, there were teacher-led reviews, sometimes including theory. It is possible that a broad range of activities is suitable in interactive sessions in mathematics, not only multiple-choice questions, for the students to develop the desired mathematical skills. There has been a lot of research in the last 10–15 years into what skills constitute mathematical knowledge and according to this research knowledge of mathematics can be split into a number of distinct competencies (Niss & Jensen, 2002). Therefore, it seems reasonable to use a variety of activities to develop these competencies.

The findings of this study indicate that the implementation of flipped classroom was beneficial for the students. This conclusion is supported by all the measures taken. The Calculus Baseline Test indicates a higher gain in the flipped-classroom group compared to the lecture group. Both groups improved their results on the written examination from 2014 to 2015, which could be explained by an easier examination or better pre-knowledge in 2015, but the group with flipped classroom in 2015 improved their results more than the control group, indicating that they learned more. Thus, this study corroborates earlier findings, and together this suggests that student performance increase in a flipped-classroom setting (Murphy et al., 2015; Maciejewski, 2015; Petrillo, 2015; Love et al., 2014). The Calculus Baseline Test does not reveal any differences with regards to different mathematical competencies. In order to explore further if flipped classroom is more beneficial for some mathematical competencies, as indicated by Wasserman et al. (2015), more research is needed.

Do flipped-classroom models benefit all students? We found that the failure rate decreased more and the highest grade increased more in the flipped classroom compared to the lecture group, indicating that the intervention was beneficial for low as well as high-performing students. A decrease in failure rate in flipped-classroom setting has also been found by Petrillo (2015). A more nuanced picture was described by Maciejewski (2015), where students with little prior knowledge in calculus, but with good basic mathematical skills, increased their performance the most in the flipped-classroom setting. Active learning and flipped-classroom settings require students to become more self-directed and take more responsibility for their own learning, which many students seems to adjust well to. However, it may be too demanding for some students (Weurlander et al., 2016a). Also, there are dyslectic students and other students with special needs who can find certain aspects the flipped-classroom setting difficult (Cronhjort & Weurlander, 2016). On the other hand, flipped classroom may also present advantages, such as the opportunity to see video clips several times, and at home. How different students are affected by flipped-classroom settings and the particular demands that this puts on students, need to be further investigated.

The engagement survey showed significant difference between the two groups in our study on 10 of the 15 items in the survey, indicating that students felt more engaged in the flipped group. The items

<table>
<thead>
<tr>
<th>Table 3. Failure rate on the exam 2014 and 2015</th>
</tr>
</thead>
<tbody>
<tr>
<td>Failure rate on examinations</td>
</tr>
<tr>
<td>Flipped</td>
</tr>
<tr>
<td>Lecture</td>
</tr>
</tbody>
</table>

In 2014, all students had lectures. In 2015, the intervention group had flipped classroom.
that differ most between the two groups indicate that the relationship between teacher and students is the most important difference regarding engagement. This is in agreement with earlier results, where students appreciated feeling more seen by the teacher and receiving more feedback with flipped classroom (Cronhjort & Weurlander, 2016). In flipped classroom, where more time is spent on discussions and interactions between peers and between teachers and students, students seem to feel that the teacher care about their learning, take their questions seriously and that they receive more support. These social interactions may help students to feel a sense of belonging and a connection to their learning environment (Kahu, 2013; Wimpenny & Savin-Baden, 2013). The higher engagement in the flipped classroom can be one explanation to why the students achieve more compared to the traditional lecture setting. This would be in keeping with earlier research, showing a positive link between student engagement and student performance (Carini et al., 2006). Moreover, the social aspects of peer discussions in the teacher-led sessions of flipped classrooms seem to improve student learning (Smith et al., 2009).

Here, we present results from one intervention only, in a specific context. Measuring the effects of teaching is a difficult task, and learning is dependent on many factors. We see no reason to believe that flipped classroom would always give similar results, regardless of context. To increase the robustness of our study, we used three different methods of measurement. Our results show a clear pattern, and indicate that interactive teaching methods can be successful and should be considered and investigated further.

REFERENCES


Mikael Cronhjort is a lecturer, researcher and educational developer at KTH Royal Institute of Technology, Stockholm, Sweden whose background is in theoretical physics and mathematics, and field of research is Engineering Education, with a special interest for Mathematics Education; and he has many years of experience from teaching physics, mathematics and educational science.

Lars Filipsson is an associate professor of mathematics at KTH Royal Institute of Technology, Stockholm, Sweden and his mathematical research is in several complex variables with many years of experience teaching undergraduate mathematics, primarily to engineering students.

Maria Weurlander is a lecturer, researcher and educational developer at KTH Royal Institute of Technology, Stockholm, Sweden having a background in biology and chemistry, and she has many years of experience from both teaching in natural science subjects and educational development work and research focus is on teaching and learning in Higher Education.