Evaluating a "Flipped Classroom" Experiment

Ole Jacob Bergfjord and Tarjei Alvær Heggernes
Department of Business Administration, Bergen University College, Møllendalsveien 6-8, 5020 Bergen, Norway

Keywords: Flipped Learning, Video Lectures.

Abstract: In this paper, a "flipped classroom" experiment is evaluated using three different datasets. We use student evaluations of the "flipped classroom" experiment in particular, in addition to regular course evaluations and exam results for the past three years in order to allow for statistical comparisons. Overall, the results from the experiment are quite positive. Interesting effects include that students report that they prepare better for lectures, are more satisfied with the course overall, and achieve slightly better grades. In particular, much fewer students get very low grades. On the one hand, we argue that our results support more experiments with technology to improve education. On the other hand, we also hope that our analysis could be useful as a reference for evaluating such experiments.

1 INTRODUCTION AND LITERATURE

Technology has evolved many an industry during the last few decades. More often than not, these industries have been watching from the sideline, believing they were exempt from the impact of new technology, only to be proven wrong and forced to catch up or disband.

Higher education is an example of this. The last few years have seen rise to several different forms of educational technology: MOOC's (Pappano, 2012), learning intelligence (Shayer and Aday, 2002), adaptive learning (Dawid, 1996) etc. Technology has also facilitated the subject of this paper: The flipped classroom.

So, how to define the flipped classroom? Tucker (2012) describes the flipped classroom as follows:

While there is no one model, the core idea is to flip the common instructional approach: With teacher-created videos and interactive lessons, instruction that used to occur in class is now accessed at home, in advance of class. Class becomes the place to work through problems, advance concepts, and engage in collaborative learning. Most importantly, all aspects of instruction can be rethought to best maximize the scarcest learning resource—time

The data presented in this paper were collected after a pedagogical approach like the one defined here was used. A fair amount, but not all of the direct instruction moved out of the classroom, and time in the classroom was for using concepts on solving assignments.

New technology is not an absolute must for flipped learning. By reading the curriculum and preparing for lectures, one could still make lectures more interactive by solving problems. In our experience however, many students do not prepare for lectures, making it harder achieve good discussions in class.

Still, new technology does facilitate flipped learning. Examples of videos facilitating flipped learning can be viewed at the website for Khan Academy (Thompson, 2011), where short instructional videos on a myriad of subject are available. The videos are made in an "old-school" blackboard style, with narration coupled with drawing explains the subject. Tools to create such videos are now plentiful, easy to use, and free or of low cost for educators.

I addition to this, the internet is a convenient and stable distribution channel for the videos, both the open internet and via closed LMS-platforms like itslearning.

2 DATA AND METHODS

The experiment was conducted in a Strategic Management class for 2nd year students in a BA program in economics and business administration

Bergfjord O. and Heggernes T..
Evaluating a "Flipped Classroom" Experiment.
DOI: 10.5220/0005412102050208
In *Proceedings of the 7th International Conference on Computer Supported Education* (CSEDU-2015), pages 205-208
ISBN: 978-989-758-108-3

at Bergen University College, Norway. The class consists of approximately 100 students. Entry requirements for the program are fairly strict, so the students are generally well qualified and highly motivated.

Although not done deliberately for research purposes, it is fortunate for comparison that the course has been virtually identical for at least three year, except from the flipped classroom experiment. This includes the student group, the teacher, the curriculum and the exam form, difficulty and grading. Hence, we will argue that the flipped classroom experiment is likely to be an important factor behind any significant changes.

Our analysis is conducted based on three different datasets. The first dataset consists of the results from a specific survey among the students regarding the "flipped classroom" experiment (see Appendix 1 for a translated version of the survey). Students were asked to complete this survey electronically and anonymously after the course and the exam. 45 students completed the survey. Although a higher response rate would have been beneficial, we still think the results are fairly reliable. Each respondent should consider 14 different statements; only two of (45*14) = 630 values are missing. Also, most students have supplemented their quantitative responses with sensible and interesting verbal comments. These are not directly included in our analysis, but the overall impression is very well aligned with the quantitative results.

Our second dataset is based on general student evaluations of the course over a period of three years – 2012, 2013 and 2014. These surveys are done every year after classes are finished. The survey consists of a number of different questions about each course, the general learning environment etc, but for our analysis we have singled out one single question:

"How was your learning outcome from the lectures in the course Strategic Management?" (Scale from 1 to 5).

The "flipped classroom" experiment was conducted in 2014, so we have results from two years without the experiment and one year with the experiment. As mentioned above, the course has otherwise not changed, so the results should be comparable. Again, the surveys were completed electronically and anonymously, after all lectures, but before the final exam, in order to avoid any (dis)satisfaction with the exam to influence the results. The response rates were similar to the experiment-specific survey, with 46, 38, and 50 students completing the surveys in 2012, 2013 and 2014.

The third and final dataset simply consists of historical exam results. For administrative purposes, the Strategic Management grade is merged with another grade before published as the official grade for a larger course. The grades we have studied are the unpublished, separate grades for Strategic Management, i.e., the course area where the experiment was conducted. We use the grade distribution (A-F) as our measure of grades, and we have grades from around 100 students each year. There will of course always be fluctuations in exam results, depending on both the exam itself and the student group.

Hence, we think that this dataset is the least reliable source of the three. However, the exam form has remained the same, the student quality should be fairly similar, and the exam has been made and graded by the same person each year, aiming for the same level across the three years. So we will argue that although they are not conclucive on their own, differences in grades could provide some useful information as part of a larger picture. More fundamentally, grades also seem a better measure of what we really are interested in. It could be (and is) discussed whether grades are a good reflection of actual learning outcome, but we assume that there is at least a clear positive correlation, such that a student who has learned more in course on average gets a better grade in the course. The fundamental purpose of any pedagogical or technological experiment should be to improve learning; not per se to introduce methods students like.[1] Thus, it is natural to include grade patterns if possible when evaluating such experiments.

3 RESULTS

In this section, we will present the key findings from the three datasets. First, consider the survey where students specifically evaluate the experiment. Table 1: Students evaluation of the flipped classroom experience.

The figure above presents the main results. Overall, the numbers suggest that the experiment was relatively successful, with most averages between 3,2 and 4,2. First, a few comments to the low scores for statement 1) and 6).

[1] Student evaluations of experiments can be important for further implementations; a factor more thoroughly discussed in the conclusion.

Table 1: Evaluation of the experiment.

Statement	Average score	Std dev
1) Available online videos increase motivation to attend lectures	3,04	0,9
2) More discussion in lectures increases motivation to attend lectures	3,36	1
3) Available videos online makes it easier to prepare for lectures	4,2	1,01
4) Available videos online made me prepare better for lectures	3,69	1,18
5) Available videos online makes it easier to repeat material after lectures	3,91	1,04
6) Available videos online made me repeat material better after lectures	2	0,95
7) Available videos online makes it easier to prepare for the final exam	4,22	1,08
8) Available videos online made me prepare better for the final exam	4,13	1,08
9) The experiment in general has changed my preparation for lectures	3,82	0,96
10) The experiment in general has changed how time is spent in lectures	3,45	0,82
11) The experiment in general has supported more effective learning	3,77	1,05
12) The experiment in general has increased my knowledge and understanding of the course area	3,76	0,74

Table 1: How much do you agree with the following statements on a scale from 1 to 5? 1=completely disagree, =completely agree.

One "traditionalist" objection to such experiments is that online material takes away students' incentives to attend lectures. Thus, it is not surprising that the online videos not increase motivation to attend lectures. Nevertheless, the average score is 3,04, which indicates that the presence of online material neither increased nor decreased motivation. We are positively surprised by this result, as some decline in motivation would be in line with the conventional wisdom. One speculative reason could be that brief videos online, if made well, could work as "teasers", and increase motivation to come to the lectures to learn more. The other statement with a low average score is statement 6. Although students indicate in statement 5 that the online videos makes it easier to repeat material after lectures, they admit that they rarely actually do this. Although better repetition after lectures would be a beneficial effect, it is again not very surprising that student admit that little has changes here.

If we look at the other statements, students repost that the online material has made it much easier to prepare for both lectures and the final exam; and that they actually do this too. Finally, the scores for statement 11) and 12) are worth a closer look, as these give an overview of whether the general objectives have been achieved. Average scores of 3,76 and 3,77 to these to statements respectively indicate that the experiment was relatively successful in supporting more effective learning and increasing knowledge about the course area.

The second dataset consists of general student evaluations, rating their learning outcome of the Strategic Management lectures.

Table 2: Learning outcomes from participating in lectures.

Year	5	4	3	2	1	Avg.	Resp
2014	11%	57%	26%	4%	2%	3,69	46
2013	0%	46%	38%	10%	5%	3,26	38
2012	2%	22%	56%	16%	4%	3,02	50

Table 2: "How was your learning outcome from the lectures in the course Strategic Management?" (Scale from 1 to 5; 1=very low, 5= very high).[2]

We see that the student satisfaction has increased over time, from 3,02 via 3,26 to 3,69 in 2014. As mentioned above, all other parts of the course have remained more or less identical throughout the period. Hence, it seems reasonable to attribute at least some of the improved student satisfaction to the flipped classroom experiment. One speculative explanmation for the stepwise improvement is that whereas the experiment was started in 2013 - and improved the results somewhat already that year - the teacher needed another year to improve the online material, resulting in a further improvement from 2013 to 2014.

Finally, we have gathered exam results from the course over the same three years. Again, the exams were made to be comparable, and were graded by the same person in the same way, so any differences could be attributed to the experiment.

[2] The percentages are based on all students participating in the survey. Each year, 1-2 students typically would answer "not relevant".

Table 3: Exam results.

	2012	2013	2014
A (5)	6,00%	4,00%	6,00%
B	26,00%	32,00%	24,00%
C	47,00%	53,00%	61,00%
D	18,00%	10,00%	8,00%
E (1)	4,00%	1,00%	1,00%
F	0,00%	0,00%	0,00%
Average	3,11	3,29	3,26

Table 3: Exam results in the course Strategic Management, 2012-2014.

The differences here are much smaller than in the previous datasets, and changes in the average are not likely to be statistically significant. However, when inspecting the material more closely, another interesting detail appears. Whereas 22% of students got low grades (D/E/F) in 2012, only 11% and 9% got low grades in 2013 and 2014. The low number of students at these grades again makes it difficult to get statistically significant results, but a speculative hypotheses could be that the experiment has proved particularly helpful for weaker students. One possible effect is that whereas many of these students before neither prepared for lectures nor repeated after lectures, the online material has made them watch short videos before lectures, thus increasing their independent course work during the lecture period from “close to zero” to “relatively small”. This, combined with more effective exam preparations, could be enough to raise many of these students’ grades from a D or an E to a C.

4 CONCLUSIONS AND IMPLICATIONS

Basic findings

Our analysis is based on three separate datasets:

1) A student evaluation of the “flipped classroom” experiment in particular, and its impact of various aspects of the learning situation.
2) Regular course evaluations for the past three years, with results before and after the flipped classroom experiment for comparison.
3) Exam results for the past three years, again with results before and after the flipped classroom experiment for comparison.

Overall, the pedagogical method seems to have worked fairly well.

Dataset 1 provides the following main insights:

The flipped classroom has in particular improved students’ preparation for lectures. Student also report that the flipped classroom has made it easier to repeat material after lectures, but students report that they still rarely do this. Students report that the flipped classroom has been particularly beneficial for repetition for the final exam. Overall, the flipped classroom has improved students’ evaluation of the lectures compared to previous years. Finally, students’ answers suggest that the flipped classroom might be more suitable for some courses than for others - this could be supported by the student comments.

Dataset 2 suggests that students are significantly more satisfied with the course after the introduction of the experiment than they were before. Although the teacher, the curriculum and the main structure of the course has remained the same during the investigated period, we do of course realize that it is difficult to conclude that this improvement is due to the experiment alone.

From dataset 3, we can observe a small improvement in average grades during the period. The most interesting effect is however the reduced frequency of low grades (D and E). One hypothesis could be that the flipped classroom experiment, with the opportunity to watch videos at home, makes it easier for weak students to get a grasp of at least the most fundamental and important parts of the course.

REFERENCES

Pappano, Laura. "The Year of the MOOC." *The New York Times* 2.12 (2012): 2012.

Shayer, Michael, and Philip Adey, eds. *Learning intelligence: Cognitive acceleration across the curriculum from 5 to 15 years*. Open Univ Pr, 2002.

Dawid, Herbert. *Adaptive learning by genetic algorithms: Analytical results and applications to economic models*. Springer-Verlag New York, Inc., 1996.

Tucker, Bill. "The flipped classroom." *Education Next* 12.1 (2012): 82-83.

Thompson, Clive. "How Khan Academy is changing the rules of education."*Wired Magazine* 126 (2011).