Evaluation of a “Flipped Classroom” Approach in Management Education

Ole Jakob Bergfjord
*Bergen University College*, ojb@hib.no

Tarjei Heggernes
*Bergen University College*, tah@hib.no

Follow this and additional works at: [http://ro.uow.edu.au/jutlp](http://ro.uow.edu.au/jutlp)

Recommended Citation
Evaluation of a “Flipped Classroom” Approach in Management Education

Abstract
In this paper, a “flipped classroom” approach is evaluated using three different datasets. We use student evaluations of the “flipped classroom” 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 are quite positive. Among the interesting effects, students report that they prepare better for lectures, are more satisfied with the course overall, and achieve slightly better grades. In particular, fewer students receive 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 and new approaches.

Keywords
Flipped classroom, video lectures, management education, evaluation, instructional design
Introduction and literature

Over the past decade, the higher education landscape has seen a rise in disruptive technological innovations, such as MOOCs (Pappano 2012), learning intelligence (Shayer & Aday 2002) and adaptive learning (Dawid 1996). One aspect of this has been the advent of the “flipped classroom”. The flipped classroom involves “inverting” the educational space, such that “events that have traditionally taken place inside the classroom now take place outside the classroom and vice versa” (Lage et al. 2000). While Tucker (2012) suggests there “is no one model” of the flipped classroom:

“... 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 basis of the flipped classroom method is the importance of active learning, as opposed to passively digesting information. Problem solving, small group discussions and short written assignments are examples of active learning (Bonwell & Eison 1991) mention, and by moving the lecture out of class, these kinds of exercises can be performed in class, under the lecturer’s guidance. Several studies suggest deeper learning as a result of introducing active learning activities before class (Brani et al. 1991) and during class (Hake 1998).

The most common way to use technology within a flipped classroom methodology is for the teacher to prepare videos for the students to watch at home. Examples of videos used to facilitate learning can be found at the website for Khan Academy (Thompson 2011), where short instructional videos on a myriad of subjects are available. The videos are made in an “old-school” blackboard style, where narration coupled with drawing explains the subject. Tools to create such videos are now plentiful, easy to use, and are free or of low cost for educators. The internet is also a convenient and stable distribution channel for videos, both the open internet and via closed LMS-platforms. New technology is not necessary for the flipped classroom, however. By reading the course material and preparing for class, lectures can still be made more interactive by solving problems – a concept similar to traditional tutorials or seminars. In our experience, however, many students do not prepare for lectures, making it harder to achieve in-depth discussions in class.

Although the “flipped classroom” concept is fairly new, there is a significant and growing body of literature on the practice. A meta-analysis by Means et al. (2010), based on more than a thousand empirical studies, concludes that in general, online learning seems to work relatively well. On average, students receiving online education performed slightly better than those receiving traditional face-to-face education. Interestingly, the strongest positive effects were found in experiments combining online and face-to-face education, however, as many such experiments include increased learning time in general, as well as other educational elements, it is hard to pinpoint how much of the positive effect is due to the online media.
A number of papers discuss the design principles and expected benefits of the flipped classroom (e.g. Bishop & Verleger 2013; Johnson 2013; Kim et al. 2014). Important principles include student-centered learning; flexibility; social as well as cognitive presence in class; and continuous assessment, both of the instruction and individual student learning. A number of studies also report various results from trials of the flipped classroom (e.g. Amresh et al. 2013; Bates & Galloway 2012; Butt 2014; Davies et al. 2013; Johnson 2013; Kurtz 2014; Mason et al. 2013; Zappe et al. 2009). The results are fairly positive overall, but generally difficult to generalize; this is because they originate from such a wide spectrum of institutions and courses, the research methodologies differ between studies, and different objectives are emphasized.

Goodwin & Miller (2013) cite a number of studies outlining the benefits of the flipped classroom as a teaching method, including Hamre and Pianta’s (2005) findings of better interaction between teacher and student, Greenberg et al.’s (2011) findings of opportunities for real-time feedback, Bergmann and Sams’ (2012) and Medina’s (2008) finding of student engagement, and Hattie’s (2008) findings of more self-paced learning. The findings of these papers are from both primary and higher education, and were used as a basis for our survey, specifically evaluating the flipped classroom method.

Despite the positive empirical results cited above, there are still a number of objections to the use of flipped classrooms, in particular related to the teacher’s role. For example, Ash (2012) suggests that many teachers are reluctant to use new technology, or, more generally, to radically change their pedagogical approach. If teachers are forced to apply an approach they feel uncomfortable with, neither student satisfaction nor learning outcomes (for instance measured by grades) are likely to improve. A related point is that the empirical results from the scientific literature are likely to be based on experiments conducted by highly motivated, tech-savvy teachers. This introduces a potentially important bias with regard to more general implementation, which in turn makes it more difficult to measure true effects.

A second objection is discussed by Forsey et al. (2013), and relates to what they call “the symbolic and ceremonial aesthetics of the lecture“ (p. 480). They argue that if lectures are substituted by videos to watch at home, and mainly individual or small group activity during traditional “lecture hours”, this might lead to a sense of fragmentation, both from the student body, the university organization as a whole, and university tradition and values. As noted by Forsey et al. (2013), however, this effect could be reduced by retaining some lectures. Such effects are unlikely to be captured in limited experiments, but they are potentially important and dangerous in the longer term.

To further contribute to this literature, this paper reports on a trial of the flipped classroom method in a Strategic Management class for second year students in a BA program in Economics and Business Administration at Bergen University College, Norway. A significant proportion of the direct instruction was moved out of the classroom in this trial, and time in the classroom was spent using concepts to solve assignment problems. The paper outlines methods and data, and discusses how the primary objective (improved learning) could be measured, and then we show how different measures could be used to study the effects of the teaching method.
Data and methods

The Strategic Management class each year consisted of approximately 100 students. The class language was Norwegian, so no international/exchange students participated. The entry requirements for the program are fairly strict, so the students were generally well qualified and highly motivated.

Ten videos were recorded, explaining the main theories used in the course. The videos were between 5 and 25 minutes in length, and most were around 12 minutes. Typically, a couple of days before the lecture addressing the theories, one or two videos (per lecture) were made available for the students via the LMS. The suggested pre-class activity consisted only of viewing the videos, no other activities were given.

The face to face class would start with a short recap of the theories from the videos, and then one or several assignments would be given for the students to work on during the lecture period. Work on the assignments would be timed to 5 or 10 minutes, with a longer period used to present and discuss the results of the assignment.

One part of the course involves the resource-based view of firms - how different kinds of companies use different resources as a basis for their competitive advantage. A video published before class discussed the theoretical principles of a resource-based view of the firm. During the face-to-face component of the class, the class was split in two: one half identifying the most important resources for Facebook, and the other half identifying the most important resources for Norwegian Airlines. The results of the discussions were written on the blackboard, and the following discussion among the whole class first addressed whether the listed suggestions were, in fact, resources for the two companies in question, and thereafter, the differences and similarities between the resources of the two companies.

The main part of the class activities might involve using the theory from the videos to discuss a case, first in smaller groups, and then as a whole class together. Different cases would often be given to different groups, or different groups would discuss different parts of the theory on the same case.

The platform used for hosting these specific videos between 2012 and 2014, Camtasia Relay, is no longer in use at our institution, so the statistics are no longer available to us, and we were not able to make a backup before the subscription ended. Later incarnations of the course used YouTube, and on average 30-50% of the attending class watched the videos before class. Over the whole course, each student watches each video twice on average – and much of this is exam preparation, rather than viewing ahead of the relevant lecture. The figure below shows viewing statistics for 2016 (YouTube-based), but we assume the pattern is also representative of the courses using the Camtasia platform. In 2016, the 18 videos were viewed a total of 4071 times, by 110 students completing the course. Each of the weekly pre-lecture peaks are 30-60 views.
Although not done deliberately for research purposes, it is fortunate for comparison that the course was identical between 2011 and 2015, except for using the flipped classroom method. This includes the composition of the student group, the teacher, the curriculum and the form, and the difficulty and grading of the exam. The flipped classroom method is thus 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 student survey about the “flipped classroom” approach (see Table 1 for a translated version of the survey with results). Based on the results categorized by Goodwin and Miller 2013, Statements 1-2 measure engagement, and Statements 3-8 measure any development in student pace and method of learning.

Students were asked to complete this survey electronically and anonymously after the course and the exam. Forty-five students completed the survey. Although a higher response rate would have been beneficial, we still think the results are fairly reliable. Each respondent considered 14 different statements; only two of (45*14) = 630 values are missing. Most students also supplemented their quantitative responses with sensible and interesting verbal comments. The overall impression is very well aligned with the quantitative results.

“We learn the theories from the videos, so we can use the theories for discussions and assignments in class.”

“In hectic times a “flipped classroom” is a good thing, as it might be difficult to attend lectures.”

“It’s easier to get the points from videos than from lecture notes.”

“It’s great to be prepared for lectures without spending too much time preparing.”

“This method makes the subject more engaging than the regular method, I talked to several other
students who have become more interested in the subject."

“If you don’t have the time to prepare, there was no reason to attend class because you can’t join
the discussions.”

From these comments, it is reasonable to believe that for these students at least, there has been a
stronger engagement in the subject, and that some of them appreciate the opportunity for self-
paced learning by learning from videos outside the lecture hall.

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 performed every year after classes are finished.
The survey consists of a number of different questions about each course, the general learning
environment, and so on, but we singled out one question in particular for our analysis:

“What was your learning outcome from the lectures in the Strategic Management course?” (Scale
from 1 to 5).

The “flipped classroom” approach was used in 2013 and 2014, so we have results from one year
with traditional lectures and two years using the flipped classroom. 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
prevent any (dis)satisfaction with the exam from influencing the results. The response rates were
similar to the flipped classroom-specific survey, with 46, 38, and 50 students completing the

The third and final dataset simply consists of historical exam results. For administrative purposes,
the Strategic Management grade is merged with another grade before it is published, as the official
grade for a larger course. The grades we have studied are the unpublished, separate grades for
Strategic Management, the course area where the flipped classroom was used. 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, and thus we think that this dataset is the least reliable source of the
three. The form of the exam has remained the same, however, student quality should be fairly
similar (based both on entry requirements and our general impression), and the exam has been
created and graded by the same person each year, aiming for the same level across the three years.
We will thus argue that although they are not conclusive on their own, the differences in grades
could provide some useful information as part of a larger picture. More fundamentally, grades also
seem to be a better measure of the subject in which we are really interested. The extent to which
grades are a good reflection of the actual learning outcome can be debated, but we assume that
there is at least a clear positive correlation, so that a student who has learned more in a 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 to introduce methods that students
like per se\(^1\). Student evaluations of courses or teaching methods are also notoriously difficult to

\(^1\) Student evaluations of experiments can, however, be important for further implementations; if an experiment is very
unpopular among students, it is unlikely to be sustainable, regardless of whether the learning outcome has improved.
Such “unpopularity” could be picked up by specific surveys on the methodology; by decreasing scores either on the
official course satisfaction surveys; or by formal complaints to the lecturer or the department.
trust and interpret (see McPherson (2006) and Seiler et al. (1999) for examples of how seemingly irrelevant factors appear to influence student evaluations.) Thus, it is natural to include grade patterns, if possible, when evaluating such experiments, as these ceteris paribus should be the best indication of whether students have, in fact, learned more during the course.

Results

In this section, we will present the key findings from the three datasets. First, we consider the survey in which students specifically evaluate the flipped classroom approach.

Table 1: Student evaluations of the flipped classroom experience

<table>
<thead>
<tr>
<th>Statement</th>
<th>Average score</th>
<th>Std dev</th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Videos available online increase motivation to attend lectures</td>
<td>3.04</td>
<td>0.90</td>
</tr>
<tr>
<td>2) More discussion in lectures increases motivation to attend lectures</td>
<td>3.36</td>
<td>1.00</td>
</tr>
<tr>
<td>3) Videos available online make it easier to prepare for lectures</td>
<td>4.20</td>
<td>1.01</td>
</tr>
<tr>
<td>4) Videos available online made me prepare better for lectures</td>
<td>3.69</td>
<td>1.18</td>
</tr>
<tr>
<td>5) Videos available online make it easier to revise material after lectures</td>
<td>3.91</td>
<td>1.04</td>
</tr>
<tr>
<td>6) Videos available online made me revise material better after lectures</td>
<td>2.00</td>
<td>0.95</td>
</tr>
<tr>
<td>7) Videos available online make it easier to prepare for the final exam</td>
<td>4.22</td>
<td>1.08</td>
</tr>
<tr>
<td>8) Videos available online made me prepare better for the final exam</td>
<td>4.13</td>
<td>1.08</td>
</tr>
<tr>
<td>9) The flipped classroom method has changed my preparation for lectures</td>
<td>3.82</td>
<td>0.96</td>
</tr>
<tr>
<td>10) The flipped classroom method has changed how time is spent in lectures</td>
<td>3.45</td>
<td>0.82</td>
</tr>
<tr>
<td>11) The flipped classroom method has supported more effective learning</td>
<td>3.77</td>
<td>1.05</td>
</tr>
<tr>
<td>12) The flipped classroom method has increased my knowledge and understanding of the course area</td>
<td>3.76</td>
<td>0.74</td>
</tr>
</tbody>
</table>

*How much do you agree with the following statements on a scale from 1 to 5? 1 = completely disagree, 5 = completely agree. N=45.*

The table above presents the main results. Overall, the numbers suggest that the flipped classroom approach was relatively successful, with most averages between 3.2 and 4.2. Standard deviations are around 1 and are fairly similar for all statements. First, we would like to make a few comments...
regarding the low scores for statements 1) and 6). One “traditionalist” objection to such experiments is that online material takes away student incentives to attend lectures, and thus, it is not surprising that the online videos did not increase motivation to attend lectures. Nevertheless, the average score is 3.04, which suggests that the presence of online material neither increased nor decreased motivation. Motivation does not, of course, equal real participation. As class attendance is not compulsory in the course, there is no formal data about this, however, our subjective impression supports the results above; we did not notice any significant changes in attendance. This is also relevant in view of the fragmentation of the student environment as a potential problem – although attendance could be influenced by the fact that many students took classes with traditional lectures scheduled before or after meetings in this course. We base this on informal headcounts in class, and comparing these headcounts made by other teachers of the same class in other subjects: the turnout was much the same in our flipped class as in the regular classes.

We are positively surprised by this result, as some decline in motivation and attendance would be in line with conventional wisdom. One speculative reason could be that brief videos online, if made well, could work as “teasers” and could increase motivation to come to the lectures to learn more.

The other statement with a low average score was Statement 6. Although the students indicated in Statement 5 that the online videos made it easier to revise the material after lectures, they admit that they rarely actually do this. Although better revision after lectures would be a beneficial effect, it is again not very surprising that the students admit that little has changed in that regard.

The students also report that the online material made it much easier to prepare for both lectures and the final exam, and also, that they actually do this. Finally, the scores for Statements 11) and 12) are worth a closer look, as they give an overview of whether the general objectives have been achieved. The average scores of 3.76 and 3.77, respectively, for these two statements suggest that the flipped classroom approach was relatively successful in supporting more effective learning and increasing knowledge about the course area – at least from the student perspective. A potential pitfall, particularly when interpreting the results based on the last two questions, is that the questions, at least implicitly, ask the students to compare learning effectiveness and understanding with and without the flipped classroom, however, the students have never taken this course without the flipped classroom. Strictly speaking, they are thus not in a position to compare them, and it is possible that some students merely feel that the flipped classroom approach has increased their understanding compared to their own pre-course situation – which should be a minimum requirement for any teaching method! However, the students have completed a number of similar “traditional” courses (without flipped classroom features), so we believe that most have a fairly reasonable view of what this course would have been like with traditional lectures when making the comparison.

The second dataset consists of general student evaluations, rating their learning outcomes from the Strategic Management lectures. These evaluations are part of the university’s formal student satisfaction surveys.
Table 2: Learning outcomes from participating in lectures

<table>
<thead>
<tr>
<th>Year</th>
<th>5</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>1</th>
<th>Avg.</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014</td>
<td>11%</td>
<td>57%</td>
<td>26%</td>
<td>4%</td>
<td>2%</td>
<td>3.69</td>
<td>46</td>
</tr>
<tr>
<td>2013</td>
<td>0%</td>
<td>46%</td>
<td>38%</td>
<td>10%</td>
<td>5%</td>
<td>3.26</td>
<td>38</td>
</tr>
<tr>
<td>2012</td>
<td>2%</td>
<td>22%</td>
<td>56%</td>
<td>16%</td>
<td>4%</td>
<td>3.02</td>
<td>50</td>
</tr>
</tbody>
</table>

“What was your learning outcome from the lectures in the Strategic Management course?”
(Scale from 1 to 5; 1=very low, 5=very high)²

We see that student satisfaction increased over time, from 3.02 to 3.26 to 3.69 in 2014. With only 38 respondents in 2013, it is difficult to obtain statistically significant results when comparing the averages in the two following years. The change in the average from 2013 to 2014 is, however, statistically significant (z-test; \( \alpha = 0.05 \)), whereas the change in the average from 2012 to 2013 is not. As mentioned above, all other parts of the course have remained more or less identical throughout the period. It seems reasonable to attribute at least some of the improved student satisfaction to the flipped classroom approach. One speculative explanation for the stepwise improvement is that although the flipped classroom was introduced in 2013 – and already improved the results somewhat that year – the teacher needed another year to improve the material used in class, resulting in a further – and greater – improvement from 2013 to 2014.

Finally, we have gathered the exam results from the course over the same three years. Again, the exams were made to be comparable, and they were graded by the same person in the same way; therefore, any differences could, at least to some extent, be attributed to the new pedagogical approach.

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

<table>
<thead>
<tr>
<th></th>
<th>2012</th>
<th>2013</th>
<th>2014</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>6%</td>
<td>4%</td>
<td>6%</td>
</tr>
<tr>
<td>B</td>
<td>26%</td>
<td>32%</td>
<td>24%</td>
</tr>
<tr>
<td>C</td>
<td>47%</td>
<td>53%</td>
<td>61%</td>
</tr>
<tr>
<td>D</td>
<td>18%</td>
<td>10%</td>
<td>8%</td>
</tr>
<tr>
<td>E</td>
<td>4%</td>
<td>1%</td>
<td>1%</td>
</tr>
<tr>
<td>F</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Average</td>
<td>3.11</td>
<td>3.29</td>
<td>3.26</td>
</tr>
<tr>
<td>N</td>
<td>105</td>
<td>91</td>
<td>101</td>
</tr>
</tbody>
</table>

Table 3: Exam results in the Strategic Management course, 2012-2014. The average scores are calculated by translating the grades to numbers. A=5 is the best grade, E=fail=1 is the worst passing grade, F(fail)=0.

The average grade was higher in 2013 and 2014 than in 2012, when traditional lectures were used. The differences here are much smaller than in the previous datasets, and none of the pairwise differences in average grades are statistically significant (z-test; α=0.05), however, when inspecting the material more closely, another interesting detail emerges. Whereas 22% of students got low grades (D/E/F) in 2012, only 11% and 9%, respectively, got low grades in 2013 and 2014. The low number of students at these grades again makes it difficult to obtain statistically significant results, but a speculative hypothesis could be that the flipped classroom proved to be particularly helpful for weaker students. One possible effect is that whereas before many of these students neither prepared for lectures nor revised after lectures, the online material made them watch short videos before the lectures, thus increasing their independent course work during the lecture period from “zero” to “small.” This, combined with more effective exam preparation, could be enough to raise the grades of many of these students’ from a D or an E to a C.

Conclusions and implications

Our analysis is based on three separate datasets:

1) A student evaluation of the “flipped classroom” approach, in particular, and its impact on various aspects of the learning situation.

2) Regular course evaluations over the past three years, with results before and after the flipped classroom was introduced for comparison.

3) Exam results for the past three years, again, with results before and after the flipped classroom was introduced for comparison.
Overall, the pedagogical method seems to have worked fairly well.

Dataset 1 suggests that the flipped classroom improved student preparation for lectures in particular. The students also reported that the flipped classroom made it easier to revise the material after lectures, but students reported that they still rarely do this. Students reported that the flipped classroom was particularly beneficial for revision for the final exam. Overall, the flipped classroom improved student evaluations of the lectures, compared to previous years.

Dataset 2 shows that students were significantly more satisfied with the course after the introduction of the flipped classroom than before. Although the teacher, the curriculum and the main structure of the course remained the same during the period investigated, we realize, of course, that it is difficult to conclude that this improvement is due to the new pedagogical approach alone.

In Dataset 3, we note 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 approach, with the opportunity to watch videos at home, makes it easier for weak students to grasp at least the most fundamental and important parts of the course.

We would like to emphasize that using as many available quality indicators as possible provides a broader understanding of how successful the approach has been. Although not available to us in this case, it would be interesting to use data on the results in following courses, time spent on studies or even more detailed user statistics for the short videos to further improve our understanding of the effectiveness of the pedagogical method.

Finally, it is worth pointing out that student satisfaction was much higher in the second year than in the first year. We believe this is important. Measuring a broad range of indicators makes it easier to improve a method in the right ways. The flipped classroom, like any other pedagogical method, is better viewed as a continuous process, than as an on/off switch. The approach was introduced in 2013, but both the videos and the lecture content were slightly modified in 2014, based on experience and feedback from the previous year.

Like any innovative approach, a flipped classroom approach is unlikely to work optimally at the first attempt. It thus seems unreasonable automatically to abandon a pedagogical experiment if the first results are disappointing – some adjustments over time should be allowed and seen as a natural part of the experiment.

References


Bergmann, J & Sams, A 2012. Flip your classroom: Reach every student in every class every day. Washington, DC: ISTE.


