The flipped classroom and cooperative learning: evidence from a randomised experiment

Njål Foldnes
BI Norwegian Business School

The flipped classroom and cooperative learning: Evidence from a randomised experiment

Njål Foldnes[*]

BI Norwegian Business School

Abstract

This article describes a study which compares the effectiveness of the flipped classroom relative to the traditional lecture-based classroom. We investigated two implementations of the flipped classroom. The first implementation did not actively encourage cooperative learning, with students progressing through the course at their own pace. With this implementation student examination scores did not differ between the lecture classes and the flipped classroom. The second implementation was organised with cooperative learning activities. In a randomised control-group pretest-posttest experiment student scores on a post-test and on the final examination were significantly higher for the flipped classroom group than for the control group receiving traditional lectures. This demonstrates that the classroom flip, if properly implemented with cooperative learning, can lead to increased academic performance.

Keywords

 flipped classroom; randomised controlled trial; collaboration; cooperative learning

1. Cooperative learning in higher education

Although the traditional lecture prevails in higher education as the most popular teaching method, recent technological advances have made possible blended or hybrid learning classrooms, in which students learn at least in part through online delivery of content. In a flipped classroom, also known as an inverted classroom, a type of blended learning approach is used whereby traditional lectures are moved outside the classroom through the use of online videos. Knowledge transmission

is delivered through videos that students view outside of class. Classroom time is then freed up so that the educator may challenge students using more active forms of learning than usually employed in lectures. Although the flipped classroom may seem particularly suited for active learning activities, it is acknowledged that more traditional lecture-based instruction may also contain some active learning components (Cavanagh, 2011)

Cooperative learning occurs when students work together in a group to reach their learning goals through discussion and peer feedback. The case for cooperative learning is theoretically well-founded in the theory of social interdependence (Johnson, Johnson, & Smith, 2007). Many studies indicate that persons involved in cooperative learning show a greater effort to achieve than persons involved in learning on their own, that is, when not learning in a group/as part of a team (Springer, Stanne, & Donovan, 1999; Johnson, Johnson, & Stanne, 2000; Hassanien, 2007; Roseth, Johnson, & Johnson, 2008). However, as noted by Herrmann (2013), the vast majority of these studies are with children in primary and secondary schools. These findings may not be expected to transfer to higher education. Little empirical evidence exists to support the use of cooperative learning to enhance academic performance in higher education (see Herrmann (2013) for a review). It is therefore of interest to further investigate ways to organise cooperative learning in higher education.

The literature concerning the effectiveness of the flipped classroom is diverse and growing (J. Bishop & Verleger, 2013; Zhao & Breslow, 2013). Outcome variables that have been studied are learning environment (J. Bishop & Verleger, 2013) and student perceptions and/or student performance (Lage, Platt, & Treglia, 2000; Davies, Dean, & Ball, 2013; Mason, Shuman, & Cook, 2013; Love, Hodge, Grandgenett, & Swift, 2013). Studies have been conducted in diverse courses such as engineering (J. L. Bishop & Verleger, 2013; Mason et al., 2013), information systems (Davies et al., 2013), economics (Lage et al., 2000), mathematics (McGivney-Burelle & Xue, 2013; Love et al., 2013) and statistics (Strayer, 2012). Studies also vary with respect to methodology, from single-group studies that do not control for confounding variables (for example,Lage et al. (2000)), via pre-post designs that investigate change occurring from one period to the next when a teaching approach is changed, for example (Lage et al., 2000; Davies et al., 2013), to studies with treatment (flipped classroom) and control (lecture) groups (Deslauriers, Schelew, & Wieman, 2011; Strayer, 2012; McGivney-Burelle & Xue, 2013; Love et al., 2013). However, in most treatment-control studies, the groups were not formed by randomisation, so it is possible that results are obscured by confounding variables. Studies also differ with respect to the duration of the flipped instruction period. Some studies (Deslauriers et al., 2011; Love et al., 2013) implement flipped instruction only in some units of the course, while in other studies flipped instruction was used for the full course length.

The flipped classroom literature is still in an early phase, and no clear patterns have yet emerged that would indicate how to implement the flipped classroom effectively. Also properly designed studies are lacking. The study described in this article contributes to the literature firstly by investigating two different implementations of the flipped classroom. One implementation is where students in a flipped classroom work alone, that is, they do not work in groups/teams. In this more

"traditional" model, the students view videos for homework, and use class time to work, mostly individually, on assigned exercises. The other implementation encourages cooperative learning by organizing class time with group work. A second contribution is the use of a randomised controlled design in comparing the latter implementation to the traditional lecture format. To the best of our knowledge such a design has not yet been employed in the evaluation of the flipped classroom. The research aim was to investigate two ways of implementing the flipped classroom and to evaluate whether either of these may lead to increased student performance relative to the more traditional lecture-homework model.

2. Methods

This section comprises two parts. Called "Study 1", the flipped classroom was organised in such a way that students worked mostly alone during class time. All online videos were available at the beginning of the course, allowing students to progress at their own pace through the entire course. There were no organised cooperative learning activities, that is, the students were not asked to work in groups/teams. Called "Study 2", the flipped classroom was organised differently, with all students working on the same course unit at the same time. The students were required to work in groups/teams, that is, to work cooperatively in class. The sections below provide further detail on these two studies. In both studies, a particular implementation of the flipped classroom was compared to the traditional lecture class.

2.1. Case description

Both Study 1 and Study 2 were conducted in semester-long first year undergraduate courses at a business school that has campuses in the largest Norwegian cities. Study 1 was conducted in a statistics course and Study 2 in a mathematics course. In both studies a particular implementation of the flipped classroom was compared to a more traditional way of instruction, based on lectures and homework. In this traditional format, the courses were taught in a lecture hall with weekly lectures. Between lectures the students were given suggestions on which exercises to do for homework to further learn the material. These exercises were not mandatory. The lecture format employed in both studies contained elements of active learning in addition to traditional lecturing. That is, in each lecture session the lecturer regularly presented an exercise that the students were asked to work on for a few minutes, and after that the lecturer worked through the exercise in front of the class. The lecturer/instructor spent the same time with the students in the flipped classroom and in the traditional lecture format. Student assessment in both courses was done through an individual examination at the end of the semester.

In both studies the flipped classroom employed screencast videos produced by the instructor of the course. These videos captured the calculations of the instructor with synched audio commentary, and were used to transfer the same knowledge as lectures: demonstrating algorithms for problem solving and providing conceptual understanding by interpreting each calculation step. Videos were organised into learning modules, each module covering the same material as taught in a typical

lecture session of three 45-minute lecture hours. A learning module, in addition to two or three videos, also contained a description of learning goals, key words, a textbook reading list and a set of exercises. The modules were available for the students on the digital learning platform. The videos presented the material in a more condensed form than lectures. Material normally covered in a lecture session of three 45-minute lecture hours typically corresponded to two or three videos, each lasting about ten minutes. However, many students reported spending several hours watching these videos at home, sometimes pausing and replaying parts to gain better understanding.

Flipped instruction took place in a study area with a mix of tables and study cells, where the instructor could move about freely and engage with the students. The instruction schedule was identical for the flipped classroom and the lecture format, with three consecutive 45-minute sessions in both the flipped classroom and lecture format each week. Hence the instructor spent the same amount of time with the students in the flipped and in the lecture format. Between instruction sessions, students following lectures were asked to work on assigned textbook exercises, while students in the flipped classroom were asked to watch the videos. The exercises were given on a week-by-week basis, and an average student may have spent about three hours working through all exercises. During the three-hour instruction sessions, students in the flipped classroom were asked to work on the assigned exercises.

2.2. Study 1: The flipped classroom where students are not in teams/groups

A major design choice for implementing a flipped classroom is whether all content (for example, videos and exercises) should be released at the beginning of the semester, or in a week-by-week fashion so that students tend to work with the same material at the same time. If all course material is made fully available online at the beginning of the semester, students progress through the material at a reasonable pace. The student moves on to the next learning unit only after having mastered the content (Bloom, 1968). Bergmann and Sams (2012) refer to this as the flipped-mastery classroom model, and consider it best practice. In Study 1 this model was employed, with all learning modules (videos, exercises, etcetera) available to the student from the beginning of the semester. During class time the instructor helped students on a mostly individual basis, by answering questions the student might have concerning theory or conceptual issues, or helping with the exercises. This occurred with individual students or in small spontaneously formed groups. The instructor did not ask students to work in groups, but it happened that groups were improvised by the students when working on the same exercise. These groups dissolved when students had worked through the exercise. The instructor did not interfere with the forming and dissolving of groups. For such group work among students to occur, it was necessary that they had progressed to the same learning unit, which was not always the case. Hence group work in Study 1 was ad hoc and occurred only to a limited extent. The exercises that students in the flipped classroom worked with during class time were identical to those assigned as homework to students following lectures.

Table 1: Instructor, instruction method, and class size at six campuses. 2012 and 2013. Instructors: A, B and C. Instruction method: L = Lecture, F = Flipped. Campuses: 1 = Bergen, 2 = Oslo, 3 = Stavanger, 4 = Drammen, 5 = Kristiansand, 6 = Trondheim.

	Campus					
Year	1	2	3	4	5	6
2012	A-L-232	B-L-272	C-L-60	B-L-52	·	·
2013	A-L-226	B-L-441	C-F-51	B-L-40	B-L-40	A-L-143

Student performance in Study 1 was measured by the grade on the final examination. Performance was investigated at six campuses, in 2012 and 2013, in a total of ten classes, totalling 1,569 students as represented in Table 1. The classes were given by three instructors (A, B and C), all well-qualified and experienced. All students completed the same final examination. None of the instructors participated in writing the examinations. Examiners external to the universities were used to grade the examinations. These were senior members of staff from other universities. The flipped classroom was implemented in one class with 51 students, the nine others being taught with traditional lectures and homework. The goal was to investigate whether the introduction of flipped learning in one class in 2013 had a significant effect on student performance relative to the traditional lecture format.

Performance in each term was measured by the grade on the final exam. Each term all students took the same final exam. None of the instructors participated in writing exams. External examiners were used to grade the exams.

Since each student belongs to a larger group of students in the same class, we have a multilevel model of student performance. At the micro level, we have the response variable, namely student's grade on the final exam. To reduce unexplained variance we also include a micro-level covariate: aptitude, represented by the grade point average (GPA) obtained the preceding term. At the macro-level (i.e. class level), we have three categorical variables: instructor (A,B, or C), type of instruction (flipped or traditional), and term (2012 or 2013).

Data analysis used for Study 1 Since each student belongs to a larger group of students in the same class, a multilevel model of student performance is employed. At the micro (that is, student) level the response variable is the student's grade on the final examination. To reduce unexplained variance a micro-level covariate aptitude is included, represented by the grade point average (GPA) obtained the preceding semester. At the macro (that is, class) level there are three categorical variables: instructor (A, B, or C), type of instruction (flipped or traditional), and semester (2012 or 2013). The latter variable is included to account for the fact that the examinations given in 2012 and 2013 were different. Note however that in each year all classes were given the same examination. The nesting of students in classes requires a multilevel model with two levels: student, indexed by i, and class, indexed by j. The micro-level variables are g_{ij}, the grade on the final statistics examination for student i nested in class j, and the GPA obtained in the preceding semester, gpa_{ij}. At the class level, we

have the following categorical variables: instructor, represented by dummy variables A_j and B_j, for instructors A and B; instruction method M_j, with $M_j = 0$ and $M_j = 1$ for traditional and flipped, respectively; finally, Y_j for the year the examination was taken, with $Y_j = 0$ for 2012 and $Y_j = 1$ for 2013. The bi-level model with varying intercept is at the student level given by

$$g_{ij} \sim N(\alpha_j + \beta \overline{g}_{ij}, \sigma_g^2) \quad \text{for} \quad i = 1, \ldots, n_j, \quad j = 1, \ldots, 10 \tag{1}$$

where n_j denotes the number of students in class j. The intercept α_j is a random effect according to the class level model

$$\alpha_j \sim N(\gamma_0 + \gamma_A A_j + \gamma_B B_j + \gamma_M M_j + \gamma_Y Y_j, \sigma_\alpha^2), \quad j = 1, \ldots, 10. \tag{2}$$

Our main interest is the effect of the "treatment" predictor M_j, that is, whether γ_M is significantly different from zero.

2.3. Study 2: The flipped classroom where students work in teams

This study used a randomised control-group pretest-posttest design. From a large class, two groups were randomly formed at the beginning of the course. Of a total of 235 students, 93 students were randomly selected to the flipped group. The other 142 students received traditional lectures. The smaller group size for flipped instruction was due to restrictions on room availability. Forming the groups by a randomised procedure ensures that marked differences between the groups after the intervention are attributable to differences in received instruction. The intervention lasted over a period of 11 weeks, as depicted in Figure 1. Each week, on the same weekday, both groups received separate teaching sessions covering the same material. Both sessions were given by the same instructor (the author of this article). The traditional lectures were given before lunch, while the flipped instruction took place after lunch. The instructor spent the same number of hours with both groups. Hence, students in the two groups experienced fairly identical learning conditions, with the exception of received instruction method.

In Study 2 the flipped classroom was implemented with more structure on class time than in the flipped classroom of Study 1. New online videos were released sequentially, on a week-by-week basis. In a given week, students prepared themselves for class by watching videos from the same learning module. Class time was then structured to allow for cooperative learning, by adopting some of the principles behind team-based learning (Michaelsen, Knight, & Fink, 2002), where active learning is implemented through cooperative teamwork. Team-based learning is based on relevant problem solving and group interaction, which makes it solidly grounded in constructivist learning theory (Hrynchak & Batty, 2012). The main learning objective in team-based learning is to provide students the opportunity to practise course concepts during class time. A key feature is multiple-choice quizzes that students take individually and then re-take as a team. In the weekly three-hour sessions, students first worked individually for one hour through the set of exercises in the current learning
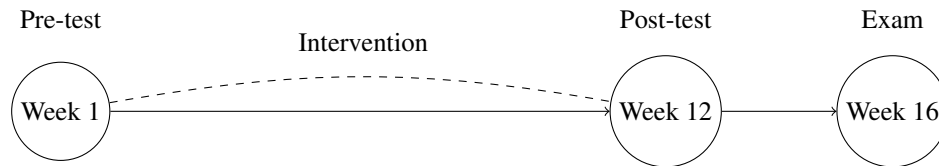
Fig. 1. Study 2: Longitudinal design. Performance measurements at three time points.

module, and then joined their designated team of five to seven students to discuss these exercises during the next hour. Importantly, these teams lasted the whole semester. Each team consisted of 7 or 8 students.

Teamwork consisted of discussing possible solutions to the exercises and agreeing upon common group answers (the exercises were given in multiple-choice format). Many of the exercises were procedural, testing the students for skills in calculation. However, there were also conceptual exercises, which involved no calculations, but required thinking clearly about concepts. These exercises were included to promote open discussions in the teams and help students understand the concepts involved in the calculations. The exercises were similar in nature to the post-test and examination questions. Students in the lecture group also had access to these exercises, but were given exercises from the textbook as (non-mandatory) homework assignments. The textbook exercises were similar to the exercises given in the flipped classroom, with a mix of procedural and conceptual questions. However, the textbook exercises were not in multiple-choice format, but had solutions at the back of the textbook.

In a flipped classroom session, after teamwork, the instructor went briefly through the exercises in plenum to give correct answers and discuss common pitfalls. Each team obtained a score that was kept in a record available for all to see on the online learning platform. These team scores did not influence grading, which was based solely on a final individual examination.

Measurements of performance (i.e., the pre- and post-tests in Figure 1), were administered at the beginning and at the end of the instruction period. Although pretesting is not necessary in randomised designs, they were included for the following reasons. By comparing pretest scores in the two groups we can confirm that the randomisation worked, with the scores being the same in both groups, up to sampling variability. Secondly, in team-based learning teams should be formed that are heterogeneous with respect to age, gender and academic performance. Team formation was partly guided by ensuring that all teams contained students with both low and high pretest scores. Also, pretesting allows testing for interaction between treatment and pretest scores, that is, whether the effect of the flipped classroom on performance might differ for weak versus strong students. Finally, by including pretest scores in the regression model statistical power is increased.

The pre-test measured the general mathematical ability of the students prior to instruction, by including topics and skills covered in secondary education. The post-test differed from the pre-test, covering course topics and skills, that is,

the central learning objectives of the course. Hence the post-test covered the same material (and had the approximately the same difficulty level) as the final examination. The examination, taken four weeks after the post-test, was a high-stakes assessment that determined the students grade. The multiple-choice format was employed for all tests. Finally, note that it does not make sense to directly compare scores from the pre-test, post-test and the final examination, since the tests had different levels of difficulty and were taken in different contexts at different time points.

Data analysis used for study 2 The average effect β of flipped instruction on post-test score $post_i$ may be estimated with the following regression model:

$$post_i = \alpha + \beta M_i + \gamma pre_i + \delta M_i \cdot pre_i + \epsilon_i. \tag{3}$$

Here M_i indicates whether the student received flipped instruction ($M_i = 1$) or traditional lecture ($M_i = 0$). To increase precision in the estimation of β we include the pre-test score pre_i as a predictor. Also, to allow for the possibility that the treatment effect varies with pre-test scores, we include the interaction term $M_i \cdot pre_i$. By substituting $post_i$ by the exam score $exam_i$ in the above model, we can similarly estimate the effect of flipped instruction on the final exam.

3. Results

3.1. Study 1

First, let us investigate visually whether the flipped class was different from the other nine classes that received traditional lectures. In Figure 2 we depict, for each of the ten classes, the mean difference between GPA and grade obtained, together with 95% confidence intervals. The y-axis measures GPA minus grade obtained on the statistics exam, for each class. These values are largely positive, indicating that it is more difficult to obtain a good grade in statistics than in courses included in the GPA. We also see that the exam in 2013 was somewhat easier than the one in 2012. However, of main interest to us is the flipped class, as indicated by dotted error bars. Students in the flipped instruction class appear to have equal performance to that of students in traditional classes. To test this statistically, the multilevel regression model given by equations 1 and 2 was estimated using the R package nlme (Pinheiro, Bates, DebRoy, Sarkar, & R Core Team, 2013). For the fixed-effects part of the model the estimated effects were $\gamma_0 : -2.15$ (p-value 0.000), $\beta : 1.45$ (p-value 0.000), $\gamma_M : 0.13$ (p-value 0.783) and $\gamma_Y : 0.36$ (p-value 0.104). The instructor effects γ_A and γ_B were estimated at -0.07 and 0.03, respectively. The likelihood ratio test associated with instructor effect yielded a p-value of 0.361. The lack of instructor effect on learning outcome in the traditional lecture format agrees with previous research in (Deslauriers et al., 2011, p. 863).
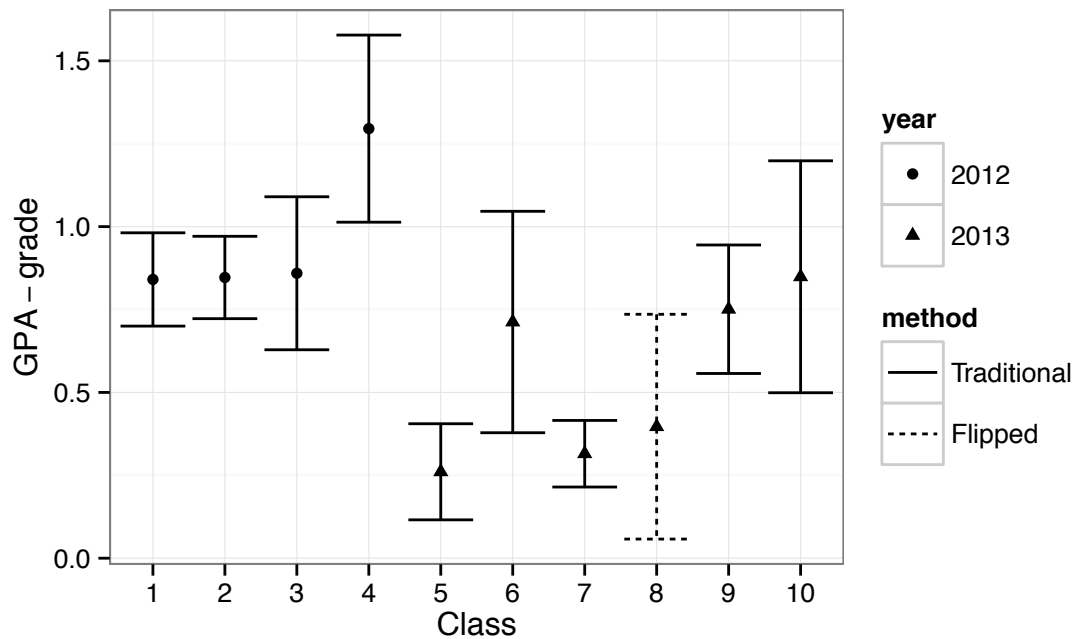
Fig. 2. $GradePointAverage(GPA) - grade$ obtained in statistics course for ten classes. Error bars depict 95% confidence intervals.

3.2. Study 2

Test scores on the pre- and post-test and on the exam were measured on a 0-100% scale. First, we depict aggregated scores in the flipped and traditional groups, on the three tests. Figure 3 presents mean values and associated 95% confidence intervals, with missing data excluded. What is of interest is the difference between the two groups in each of the three panels. It does not make sense to compare the scores across panels, since these scores come from three different tests. The pre-test, post-test and exam have different test difficulty, and were taken in different contexts at different time points. In the pre-test the flipped group had a mean score of 62.5% compared to 60.7% in the traditional group. This confirms that the forming of groups was truly random. Statistically, the difference is non-significant, the p-value of the t-test being 0.4574. The difference on the post-test is however significant, with respective mean scores of 63.2% and 50.1% for the flipped and traditional group, with p-value equal to 0.0001. On the examination, the respective scores for the flipped and traditional groups were 64.8% and 54.0%, still significant with a p-value of 0.003.

Figure 3 clearly indicates higher performance in the flipped group relative to the lecture group, and the effect may be estimated with eq. (3). In this study 235 students were involved, with performance measured on three occasions; so with no missing test scores, the cross-sectional time-series data would consist of 705 test scores. Unsurprisingly, not all students participated in all tests, with 119 of the test scores missing. List-wise deletion would result in loss of a substantial amount of data. So, to reduce bias and increase efficiency, multiple imputation with the R package Amelia (Honaker, King, & Blackwell, 2011) was used to handle missing data. Multiple imputations were combined according to standard rules.
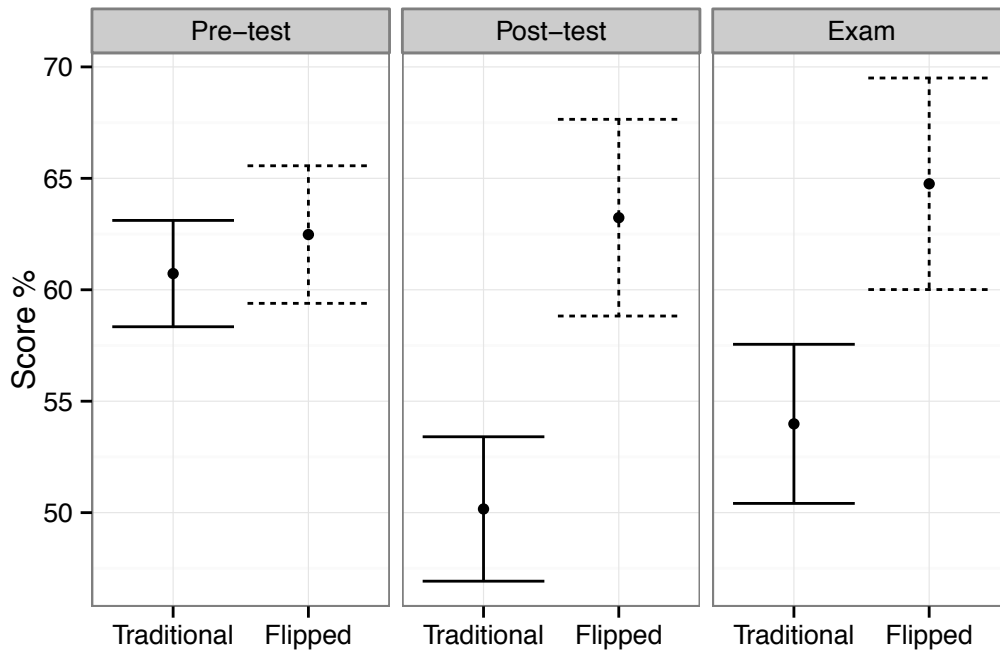
Fig. 3. Percentage score on three performance tests for flipped and traditional groups. Error bars depict 95% confidence intervals.

In estimating the effect of flipped instruction on the post-test scores, we found the interaction parameter δ in eq. (3) to be highly insignificant ($p > 0.7$). That is, the treatment effect did not vary across pre-test scores. Hence, the interaction term was removed to yield a simplified model:

$$post_i = \alpha + \beta M_i + \gamma pre_i + \epsilon_i. \tag{4}$$

The estimates, with standard errors in parentheses, were as follows: $\hat{\alpha} = 28.3(4.8)$, $\hat{\beta} = 12.1(3.0)$ and $\hat{\gamma} = 0.36(0.08)$. The 95% confidence interval for the treatment effect β on post-test scores was $(5.7, 18.3)$.

For the effect of flipped instruction on the final-exam score, the interaction term was again highly insignificant. The estimates for the examination response were $\hat{\alpha} = 20.4(6.3)$, $\hat{\beta} = 8.9(3.5)$, and $\hat{\gamma} = 0.56(0.09)$. The 95% confidence interval for the treatment effect β on exam scores was $(1.5, 16.3)$.

4. Discussion and conclusion

The flipped classroom was compared to the traditional lecture with regard to test and examination scores. Conducted in entry-level undergraduate courses, the traditional lecture format was the same in both studies. It consisted of traditional lecturing, that is, the transmission of conceptual knowledge and procedural skills. In addition, lectures were also interspersed with active learning elements, where students were asked to spend a few minutes solving exercises to consolidate the knowledge transmitted. The students were asked to solve a set of exercises between lectures as non-mandatory homework.

How the flipped classroom was implemented differed. In our first study all online material was made available from the beginning of the course. Students worked as individuals (rather than in a group or team) throughout the course. There was a limited amount of peer interaction during class time, with each student getting help and feedback from the instructor mainly on an individual basis. This way of organising the flipped classroom is referred to as the flipped-mastery model (Sams and Bergmann, 2012). Students in the flipped classroom performed as well, but not significantly better, than students following traditional lectures. It was hypothesised that the inclusion of more cooperative learning elements into the flipped classroom might increase learning. Hence, in the second study the flipped classroom was organised with students being asked to work in groups/teams that lasted the whole semester. This was facilitated by sequential delivery of material, week by week. A typical learning session in the flipped classroom first consisted of individual work on a set of exercises, then teamwork on the same set of exercises, with discussions leading to a common team answer to each exercise. At the end of the session the instructor then briefly demonstrated how to solve the exercises. It might be argued that this particular kind of cooperative group work, based on actively answering questions with almost instant feedback, is a powerful learning tool in itself, much more so that the cooperative nature of the work. However, team-based learning (Michaelsen et al., 2002) advocates a particular kind of collaborative work, based on answering questions both individually and in group. The feedback a student receives is first and foremost from fellow students in a group, so learning through instant feedback should be seen as part of a cooperative learning experience.

Study 2 was conducted as a randomised controlled trial in order to detect a potential causal effect of the cooperative flipped classroom relative to the lecture format. A highly significant increase in performance was found, 12 percentage points, in the flipped group relative to the lecture group. This effect was the same for both weak and strong students.

There are some limitations to the findings. The use of experiments to evaluate teaching interventions has limitations, with several potential pitfalls that might threat the value of an experiment (Kember, 2003). There are many ways to implement an innovative teaching method, and many factors might influence its impact. For instance, some instructors might function better with the traditional lecture format than in a flipped classroom. It is therefore not straightforward to generalise from the experiment in Study 2. Another pitfall is the ethical issue of possibly disadvantaging one group of students by letting them be the subject of a different teaching method. The control group received lectures, which is the business-as-usual method of teaching and were not put at a disadvantage relative to previous cohorts or students at other campuses. Another limitation is its restricted operationalisation of performance, by using examination-like tests to measure knowledge of the subject matter. The pre-test, post-test and examination were multiple choice tests composed of items that measured to a large extent procedural knowledge, and to some extent conceptual knowledge. It is clear that academic performance is a multidimensional concept, measured only to some extent by test scores. Other limitations include that the study is discipline

specific, limited to undergraduates and that data were gathered from only one institution in only one country. Future work will need to explore this with different levels of students, in different disciplines and in different cultural contexts.

The main conclusion drawn from this study is that the flipped classroom has potential to help students learn more than they do with traditional lectures. In the traditional lecture there is often little room for collaboration between students. The flipped classroom can accommodate a wide variety of active learning elements, including collaboration between students, whether in teams or not. The transmission of knowledge is outsourced to video tutorials, so that hours in the classroom may be organised to encourage and/or facilitate teamwork. In Study 1 the flipped classroom was implemented with little in the way of cooperative learning experiences (students had little interaction with their fellow students), and this did not result in better examination scores than the traditional lecture classes. In Study 2 the way that the flipped classroom was organised was changed, with cooperative learning as a central element. Given the highly significant increase in test and examination score performance over the traditional lecture group, the conclusion is that the flipped classroom implemented with cooperative learning is a more effective teaching method than the traditional lecture-homework format. Further studies are needed, over a variety of flipped classroom implementations, in order to establish best practices for this promising, but still understudied, teaching method.

References

Bergmann, J., & Sams, A. (2012). *Flip Your Classroom: Reach Every Student in Every Class Every Day*. International Society for Technology in Education.

Bishop, J., & Verleger, M. (2013). Testing the flipped classroom with model-eliciting activities and video lectures in a mid-level undergraduate engineering course. In *Frontiers in Education Conference, 2013 IEEE* (pp. 161–163).

Bishop, J. L., & Verleger, M. (2013). The flipped classroom: A survey of the research. In *ASEE National Conference Proceedings.* Atlanta, GA.

Bloom, B. S. (1968). *Learning for mastery.* Regional Education Laboratory for the Carolinas and Virginia.

Cavanagh, M. (2011). Students experiences of active engagement through cooperative learning activities in lectures. *Active Learning in Higher Education*, *12*(1), 23–33.

Davies, R. S., Dean, D. L., & Ball, N. (2013, June). Flipping the classroom and instructional technology integration in a college-level information systems spreadsheet course. *Educational Technology Research and Development*, *61*(4), 563–580.

Deslauriers, L., Schelew, E., & Wieman, C. (2011, May). Improved Learning in a Large-Enrollment Physics Class. *Science*, *332*(6031), 862–864.

Hassanien, A. (2007). A qualitative student evaluation of group learning in higher education. *Higher Education in Europe*, *32*(2-3), 135–150.

Herrmann, K. J. (2013). The impact of cooperative learning on student engagement: Results from an intervention. *Active Learning in Higher Education*, *14*(3), 175–187.

Honaker, J., King, G., & Blackwell, M. (2011). Amelia II: A program for missing data. *Journal of Statistical Software*, *45*(7), 1–47. Retrieved from http://www.jstatsoft.org/v45/i07/

Hrynchak, P., & Batty, H. (2012). The educational theory basis of team-based learning. *Medical teacher*, *34*(10), 796–801.

Johnson, D. W., Johnson, R. T., & Smith, K. (2007). The state of cooperative learning in postsecondary and professional settings. *Educational Psychology Review*, *19*(1), 15–29.

Johnson, D. W., Johnson, R. T., & Stanne, M. B. (2000). *Cooperative learning methods: A meta-analysis.*

Kember, D. (2003). To control or not to control: the question of whether experimental designs are appropriate for evaluating teaching innovations in higher education. *Assessment & Evaluation in Higher Education*, *28*(1), 89–101.

Lage, M. J., Platt, G. J., & Treglia, M. (2000). Inverting the classroom: A gateway to creating an inclusive learning environment. *The Journal of Economic Education*, *31*(1), 30–43.

Love, B., Hodge, A., Grandgenett, N., & Swift, A. W. (2013). Student learning and perceptions in a flipped linear algebra course. *International Journal of Mathematical Education in Science and Technology*, *45*(3), 317–324.

Mason, G. S., Shuman, T. R., & Cook, K. E. (2013). Comparing the effectiveness of an inverted classroom to a traditional classroom in an upper-division engineering course. *IEEE Transactions on Education*, *56*(4).

McGivney-Burelle, J., & Xue, F. (2013). Flipping Calculus. *PRIMUS*, *23*(5), 477–486.

Michaelsen, L. K., Knight, A. B., & Fink, L. D. (2002). *Team-based Learning: A Transformative Use of Small Groups*. Praeger.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R Core Team. (2013). nlme: Linear and nonlinear mixed effects models [Computer software manual]. (R package version 3.1-113)

Roseth, C. J., Johnson, D. W., & Johnson, R. T. (2008). Promoting early adolescents' achievement and peer relationships: the effects of cooperative, competitive, and individualistic goal structures. *Psychological bulletin*, *134*(2), 223.

Springer, L., Stanne, M. E., & Donovan, S. S. (1999). Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: A meta-analysis. *Review of educational research*, *69*(1), 21–51.

Strayer, J. F. (2012, July). How learning in an inverted classroom influences cooperation, innovation and task orientation. *Learning Environments Research*, *15*(2), 171–193.

Zhao, Y., & Breslow, L. (2013). *Literature review on hybrid/blended learning* (Working paper). MIT Teaching & Learning Laboratory.