A Retrospective Study on the Effects of Flipping a Calculus Course

Ng Wee Leng and Teo Kok Ming
weeleng.ng@nie.edu.sg
National Institute of Education
Nanyang Technological University
Singapore

Abstract
The purpose of this study was to examine the effects of a calculus course using the flipped classroom model on undergraduate students' achievement in mathematics which was measured by their scores on three quizzes, a test, and a final written examination, as well as their overall scores. The scores of a total of 58 second year students, comprising 17 students in the experimental group and 41 students in the control group, enrolled in a university degree programme in Singapore were analysed retrospectively using analysis of covariance (ANCOVA) so as to control for initial differences. The experimental group comprised students who took the flipped calculus course in the August 2016 semester while the control group comprised students who took the same calculus course taught using a lecture-tutorial approach in the August 2013 semester. Results of ANCOVA show that after controlling for initial differences the experimental group scored statistically significantly higher in the test but lower in the final examination than the control group.

1. Introduction
In the past few years, a new pedagogical approach called the flipped classroom, or flipped learning, has gained popularity in the United States. Under this model, the traditional approach of using class time to teach new academic content followed by doing homework outside class is “flipped” such that students learn the new content outside class followed by working on homework or extended tasks during class time. The claim is that through such a pedagogical approach students will develop higher-order skills and the skills to learn on their own using digital resources, and that essential 21st century competencies such as self-directed learning, critical thinking, and collaboration can be inculcated during class interactions with the teacher and peers (for a reference, see [9]). With the advent and availability of technology, the flipped classroom pedagogy has since been implemented in many schools and universities in the United States. Despite the popularity of the flipped classroom model, few studies have been done on the impact of this pedagogy on students’ learning outcomes, especially in higher education.

Though the flipped classroom is not a new pedagogy, in Singapore not many teachers in schools and instructors in universities have embraced this teaching approach. Part of the reason could be that not many research studies have been done to evaluate the effectiveness of this pedagogy on student learning. In view of the dearth of studies on flipped learning of undergraduate mathematics, in the past five semesters, we had experimented with flipped classrooms in two undergraduate mathematics courses, namely Calculus II and Algebra II for which students prepared for each lesson by viewing screencasts produced by the instructor, reading relevant sections of the textbook or course notes, and completing quizzes online. We last conducted the Calculus II course, offered once in every academic year, using the traditional lecture-tutorial mode in the August 2013 semester and delivered the flipped version of the Calculus II course for the first time in the August 2014 semester. A revised version of Calculus II was first conducted in the August 2015 semester and repeated in the August 2016 semester. Through conversations with past students, we found that many of them had positive attitude towards this learning mode, after they had overcome the initial
discomfort with this approach, which was different from the traditional lecture-tutorial style they were familiar with. With this prior experience, we have revised the existing video clips to incorporate stronger mathematics pedagogy and planned to conduct a rigorous study to evaluate the effects of the flipped learning experience for these two courses.

The present study served as a pre-study to gather preliminary insights that can be tested by the aforementioned study which we are currently planning. More specifically, we conducted retrospective analyses of data based on the flipped Calculus II course in the August 2016 semester vis-à-vis the non-flipped Calculus II course in the August 2013 semester to examine the effects of the flipped Calculus II course on undergraduate students’ achievement in mathematics which was measured by their scores on three quizzes, a test, and a final written examination, as well as their final overall scores. In this paper we report on the design of the flipped Calculus II course and the results of the retrospective analyses.

2. Literature Review

O’Flaherty and Phillips, in their scoping review of the use of flipped classrooms in higher education in [9], identified only twenty eight articles in peer-reviewed journals published from 1994 to 2014, after an extensive search of eight electronic databases. These studies span across different disciplines, but are predominantly (13 out of 28) in healthcare, medicine, nursing and pharmacy. In [5], Keengwe et al. included studies of flipped classrooms in American colleges, and the general findings were positive. These studies provide evidence that the flipped classroom can promote student-centred, active and meaningful learning, covering disciplines such as college algebra, history, psychology, statistics, and STEM. However, most of these studies did not report on achievement outcomes while some of them cited constructivist principles and included project-based learning. Our own literature search found a few other studies on flipped classrooms in higher education (e.g. [2], [4], [7], [8], [10], [13]).

In [9], O’Flaherty and Phillips reported that the majority of articles evaluated student outcomes by comparing an existing course taught in a traditional manner with a course imbedding a flipped class, and that a large number of articles, using surveys with Likert scale and free text responses, reported an increased student satisfaction with the flipped approach and active learning methods used. In most studies the flipped approach either achieved increased academic performance as measured by improved examination results, overall improvement in pre-test to post-test scores, or course grades compared with historical controls. However, they also noted that a few studies had found that students were quite negative towards the introduction of flipped class despite an improvement in student grades. For example, Strayer, in his study comparing the learning environments of an inverted (flipped) introductory statistics class and a traditional introductory statistics class, reported that “students in the inverted classroom were less satisfied with how the classroom structure orientated them to the learning tasks in the course” [11, p. 171]. Strayer and Hanson described in [12] how this approach was used to understand mathematical discussion and strategies used by pre-service teachers when they worked on an algebraic task outside class time. They noted that it is important to align in-class and out-of-class tasks to minimize student frustration.

In the four papers, namely [4], [7], [8], and [10], which we found in the literature on flipped classrooms in undergraduate mathematics courses that were not reviewed in the study by O’Flaherty and Phillips in [9], the authors reported that students generally preferred flipped classrooms to traditional lectures, and that the students felt that they learned better in flipped classroom format compared with lecture format. In addition, Love et al. in [7] as well as McGivney-Burelle and Xue in [8] reported that students in flipped classrooms scored better in
examinations than those in traditional classes. Ogden et al. studied in [10] two implementations of flipped courses about college algebra but only managed to collect perceptions of learning from a handful of college students and there was no evidence about mathematics achievement. The foregoing findings collectively suggest that the flipped format can enhance both cognitive and affective outcomes.

Even though flipped classroom generally means that “events that have traditionally taken place inside the classroom now take place outside the classroom and vice-versa” [6, p. 32], there are different ways to implement the flipped classroom model, depending on what are being done before class and inside the classroom. In the studies by Jungić et al. in [4], Love et al. in [7], and McGivney-Burelle and Xue in [8], the pre-class activities included watching pre-recorded screencasts and/or reading relevant materials in the textbook. Students then did online quiz or assessment either before class or in class. In-class activities included discussing quiz questions, peer instruction (developed by Harvard University physicist Eric Mazur, as summarized in [1]), and/or solving mathematical problems in small groups. In [13], Talbert discussed three modes of inverted classroom design that he implemented in three linear algebra classes: “as a one-time class design to teach a single topic, as a way to design a recurring series of workshops, and as a way of designing an entire linear algebra course” (p. 361). However, he did not collect any formal data to support his ideas. On the other hand, anecdotal reports from US schools, in particular, a widely cited report about mathematics curriculum at Byron High School¹, suggest that flipping has benefitted some mathematics students. However, there is a need to evaluate this form of learning for university students outside of US.

In summary, the research supporting the flipped approach in tertiary mathematics is quite limited, and challenges abound. This is similar to the general sense of small research base to support this approach (see [3]). The aforementioned planned study aims to generate more locally relevant data on how identified challenges of this approach may be addressed, in order to support reforms towards innovative teaching practices at tertiary level and the pre-study on which this paper reports serves to provide us with useful preliminary data and insights on how the design of the planned study could be improved.

3. Method

The non-flipped Calculus II group, hereafter the control group, was comprised of 41 students (34 female and 7 male) who took Calculus II course in the August 2013 semester. These students were in the same lecture group, and thus attended mass lectures together, and assigned to 3 different tutorial groups randomly. On the other hand, the flipped Calculus II group, hereafter the experimental group, was comprised of 17 students (13 female and 4 male) who took the Calculus II course in the August 2016 semester and attended tutorial sessions together. Both groups were taught by the same instructor. Enrolments in our degree programmes typically comprise a higher proportion of female than male students.

The assessment for the Calculus II course for both the experimental and control groups consisted of three 15-minute written quizzes, one 50-minute written test, and a 2.5-hour written final examination. Each quiz consisted mainly of procedural questions, whereas the test consisted of some difficult questions, on top of routine ones. The final examination tested all the topics covered in the course, and it contained five questions, with each question having several parts. Though the questions for the quizzes, tests and final examinations for the two groups are different,

¹ https://sites.google.com/a/byron.k12.mn.us/byron-high-school-mathematics-department/flipped-classroom; also http://www.boston.com/bostonglobe/editorial_opinion/oped/articles/2011/09/18/flipping_for_math/a
they are mostly comparable in terms of their level of difficulty as we will explain in the section entitled Data Analysis and Discussion. As the 3 quizzes, the test, and the final examination are formal assessment components of the Calculus II course, the course instructor is required to ensure that questions set for these assessment instruments for each cohort are different from those set for previous batches.

To examine the effects of flipping the Calculus II course on student achievement in mathematics, analysis of covariance (ANCOVA) was employed to analyse retrospectively the scores on the 3 quizzes, the test, the final examination and the overall performance of the 17 students in the experimental group and 41 students in the control group. The overall score on Calculus I was used as a covariate in ANCOVA to control for initial differences. The Calculus I course is one which all students taking Calculus II are required to have taken and passed.

4. Design of the Flipped Calculus Course

Like students in the non-flipped Calculus II course, students in the flipped Calculus II course met three times a week for 12 weeks: two 50-minute lectures and one 50-minute tutorial session. Both the flipped and non-flipped courses were taught by the same instructor, who is the second author, and the contents covered for both courses were the same.

The weekly tutorial session for the flipped Calculus II course was conducted in the same way as the non-flipped Calculus II course. The students were given weekly problems that they were expected to solve before they came for each tutorial session. During every tutorial session, some students would be asked to present their solutions on the whiteboard, while the tutor would explain and guide students to arrive at solutions to problems they could not solve. On the other hand, the lectures for the two cohorts were conducted differently, as elaborated below.

For the flipped Calculus II class, screencasts prepared by the instructor for each lecture were uploaded into the Learning Management System (LMS) a week earlier, and students were required to view the screencasts before they came for the lecture. They were also encouraged to read the relevant pages of the course notes. After viewing the screencasts, students would need to complete an online quiz and a worksheet before lesson. The online quiz tested basic concepts covered in the screencasts, and it allowed the instructor to assess students’ understanding. The worksheet consisted of two to three procedural problems that were similar to the examples covered in the screencasts and notes. During the lecture, students sitting in groups of three to four were first asked to compare and look at each other’s solutions to the worksheet problems, while the instructor circulated the class to answer students’ questions. The instructor would then discuss and clarify concepts covered in the screencasts and notes that the students did not understand, or explain one or two online quiz questions that the majority of students answered wrongly. After that, the instructor would carry out one or more of the following activities: adopt peer instruction strategy; arrange for students to work collaboratively in groups to solve two to three higher-order thinking questions that aim to deepen their conceptual understanding of the topics covered. Towards the end of the lecture, the instructor would engage the class in discussion on the higher-order thinking questions that they had worked on.

For the non-flipped Calculus II class, every lecture was conducted in a traditional way. The instructor would present the content for that lecture using slides, explaining the concepts and working on examples on blank slides, and would occasionally pose some conceptual questions to the class. The instructor would also give students 10 to 15 minutes to work on some procedural problems that were similar to the examples that were presented (these problems were given as worksheet problems in flipped learning mentioned in the above paragraph), and would give them the answers after most of them had completed the problems.
In Figure 1 below, the screen capture on the left shows a slide of a screencast, with voiceover by the instructor and the options for students to stop or rewind the screencast at any juncture while viewing the screencast. The screen capture on the right shows a question from an online quiz, the answer chosen by the student, and the correct answer. There were options for students to view their answers and the feedback to each question after completing the online quiz.

Figure 1: Screen Captures of Screencast and Online Quiz

5. Data Analysis and Discussion

To examine the differences in scores between the experimental and control groups on the 3 quizzes, the test, the final examination, and the overall performance, ANCOVA was used where Calculus I overall scores served as a covariate. The Calculus I course was a prerequisite for all students taking the Calculus II course. It covers concepts pertaining to real-valued functions in one variable such as domain, codomain, range, composition of functions, graphs, limits and continuity, differentiation and applications of differentiation, as well as integration and applications of integration.

In order to compare the initial ability level of students in the control and experimental groups, an independent samples t-test was run on the Calculus I overall scores of the experimental and control groups. It was determined that the experimental group scored significantly higher than the control group at the 5% level.

Quiz 1
Quiz 1 for the flipped and non-flipped Calculus II course comprised five and four items respectively, where each item was on finding the limit of an infinite sequence. These items assessed students’ ability to apply the limit theorems to compute the limits of sequences, and the overall demands of the two quizzes were comparable. Table 1 below shows an item in Quiz 1 of the flipped class and a parallel item in Quiz 1 of the non-flipped class.

<table>
<thead>
<tr>
<th>Flipped Class</th>
<th>Non-Flipped Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Find ( \lim_{n \to \infty} \ln(1 + \frac{(-3)^n}{n!}) )</td>
<td>Find ( \lim_{n \to \infty} \cos(2\pi + \frac{1}{n}) )</td>
</tr>
</tbody>
</table>

In answering either item, students were expected to first note that logarithm, or respectively cosine, is a continuous function in its domain, and recall knowledge of the limit of the sequence
\(\frac{(-3)^n}{n!}\), or respectively \(\frac{1}{n}\), both of which were explicitly discussed in the course, then apply a limit theorem to arrive at the answer \(\ln 1\), or respectively \(\cos 2\pi\).

Results of data analysis reveal that although the mean score for the experimental group was higher than that of the control group, after controlling for the Calculus I scores, the adjusted mean score of the control group was higher than that of the experimental group. This could be due to the fact that the mean Calculus I score for the experimental group is statistically significantly higher than that for the control group. The difference, however, was not statistically significant at the 5% level. The mean scores and adjusted mean scores on Quiz 1, and the F values are presented in Table 2 below.

### Table 2 – Comparison of Mean Scores and Adjusted Mean Scores on Quiz 1

<table>
<thead>
<tr>
<th>Quiz 1</th>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>Adjusted Mean</th>
<th>Difference</th>
<th>F</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Experimental</td>
<td>17</td>
<td>9.0882</td>
<td>8.711</td>
<td>-0.141</td>
<td>.066</td>
<td>.798</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>41</td>
<td>8.6951</td>
<td>8.852</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Quiz 2

Quiz 2 for the flipped class comprised four items: one item covered the definition of convergence of series, and the other three items covered \(n\)th-term test and comparison tests; while Quiz 2 for the non-flipped class had three items that covered \(n\)th-term test and comparison tests. Table 3 below shows two items each in Quiz 1 of the flipped and non-flipped classes.

### Table 3: Items in Quiz 2 of the Flipped and Non-Flipped Classes

<table>
<thead>
<tr>
<th>Flipped Class</th>
<th>Non-Flipped Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Determine the convergence of each of the following series:</td>
<td></td>
</tr>
<tr>
<td>((a) \sum_{n=2}^{\infty} \frac{n}{\ln n}) \qquad ((b) \sum_{n=1}^{\infty} \frac{\cos^2 n}{2^n})</td>
<td>Determine the convergence of each of the following series:</td>
</tr>
<tr>
<td>((a) \sum_{n=1}^{\infty} \ln \left(2 + \frac{1}{n}\right)) \qquad ((b) \sum_{n=1}^{\infty} \frac{\sin^2 n}{n^2 + n})</td>
<td></td>
</tr>
</tbody>
</table>

To answer each item correctly, students needed to determine the appropriate test to use, apply the test, and then arrive at the correct conclusion. As mentioned above, Quiz 2 in the flipped class had one additional item, reproduced below, that tested students’ concept of convergence of series.

\[\text{Suppose } \sum_{n=1}^{\infty} a_n \text{ is an infinite series such that the } n\text{th partial sum } S_n \text{ is given by} \]
\[S_n = a_1 + a_2 + \cdots + a_n = \frac{2n}{6n+9}\]

\[\text{Show that the series is convergent, and find its sum.}\]

Because of this additional item, the overall demand for the flipped class was slightly more than that of the non-flipped class as the duration for each quiz was 15 minutes for both the flipped and non-flipped classes.

Similar to the results pertaining to Quiz 1, even though the Quiz 2 mean score for the experimental group was higher than that of the control group, after controlling for the Calculus I scores, the adjusted mean score of the control group was higher than that of the experimental group. The difference was again not statistically significant at the 5% level. The mean scores and adjusted mean scores on Quiz 2 and the F values are presented in Table 4 below.
Quiz 2
Comparison of Mean Scores and Adjusted Mean Scores on Quiz 2

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>Adjusted Mean</th>
<th>Difference</th>
<th>F</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental</td>
<td>17</td>
<td>8.6471</td>
<td>8.207</td>
<td>-0.256</td>
<td>.157</td>
<td>.694</td>
</tr>
<tr>
<td>Control</td>
<td>41</td>
<td>8.2805</td>
<td>8.463</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Quiz 3
There were three items in Quiz 3 for each of the flipped and non-flipped classes. One item in both classes involved finding the partial derivatives of a given real-valued function of two variables. For the other two items in the flipped class, one was to show that the limit of a real-valued function of two variables does not exist as \((x, y)\) tends to a point, while another was to use the squeeze theorem to find the limit of a real-valued function of two variables as \((x, y)\) tends to a point. On the other hand, for the other two items in the non-flipped class, each involved determining whether the limit of a real-valued function of two variables as \((x, y)\) tends to a point exists, and if it does, find its limit. These two items for the non-flipped class were therefore more demanding than the corresponding two items for the flipped class. Table 5 below shows one item each from the two classes.

Table 5: Items in Quiz 3 of the Flipped and Non-Flipped Classes

<table>
<thead>
<tr>
<th>Flipped Class</th>
<th>Non-Flipped Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Let (f(x, y) = \sin(x^2 + 2xy)). Find (f_x(x, y), f_y(x, y)) and (f_{xy}(x, y)).</td>
<td>If (f(x, y) = e^{xy} \ln y), find (\frac{\partial f}{\partial x}) and (\frac{\partial f}{\partial y}).</td>
</tr>
</tbody>
</table>

The demands of the two items in Table 5 were comparable to each other. However, the overall level of difficulty for the non-flipped class was higher than that of the flipped class because of the other two items.

Although after controlling for the Calculus I scores, the adjusted mean score of the experimental group was still higher than that of the control group, the difference was not statistically significant at the 5% level. The mean scores and adjusted mean scores on Quiz 3 and the F values are presented in Table 6 below.

Table 6: Comparison of Mean Scores and Adjusted Mean Scores on Quiz 3

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>Adjusted Mean</th>
<th>Difference</th>
<th>F</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Experimental</td>
<td>17</td>
<td>9.5294</td>
<td>9.009</td>
<td>0.647</td>
<td>1.863</td>
<td>.178</td>
</tr>
<tr>
<td>Control</td>
<td>41</td>
<td>8.1463</td>
<td>8.362</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Test
The test for both flipped and non-flipped classes each contained five questions, with at most two parts to each question. Both tests assessed students on essentially the same topics: finding limits of infinite sequences, using the comparison tests, ratio test, and root test to determine the convergence (including absolute convergence and conditional convergence) of series, power series and their radii and intervals of convergence, term-by-term differentiation and integration of power series and their applications to finding sums of some series. The following questions were the most difficult ones in the respective tests for both classes:
### Flipped Class

Suppose $\sum_{n=1}^{\infty} a_n$ is an infinite series that is convergent.

(a) If $a_n > 0$ for all $n \geq 1$, prove that $\sum_{n=1}^{\infty} a_n^2$ is convergent. [Hint: Since $\sum_{n=1}^{\infty} a_n$ is convergent, $\lim_{n \to \infty} a_n = 0$. Now use the $\varepsilon-N$ definition of limit of sequence with appropriate value for $\varepsilon$.]

(b) If infinitely many $a_n$'s are negative, is it still true that $\sum_{n=1}^{\infty} a_n^2$ is convergent? If it is true, prove it; if it is false, give a counterexample.

### Non-Flipped Class

Determine whether each of the following statements is true or false. If it is true, prove it; if it is false, give a counterexample.

(a) If $\sum_{n=1}^{\infty} a_n$ is a series with positive terms such that $\sum_{n=1}^{\infty} (-1)^n a_n$ is divergent, then $\sum_{n=1}^{\infty} a_n$ is divergent.

(b) If $\{a_n\}_{n=1}^{\infty}$ is a sequence such that the series $\sum_{n=1}^{\infty} (-1)^n a_n$ is convergent, then the series $\sum_{n=1}^{\infty} 2^n a_n$ is convergent.

The demands of the above two questions were comparable, and the overall levels of difficulty of the two test papers were about the same.

Even after controlling for the Calculus I scores, the adjusted mean score of the experimental group was still higher than that of the control group, and the difference was statistically significant at the 5% level. Indeed the experimental group had performed very much better in this test than the control group. The mean scores and adjusted mean scores on the test and the $F$ values are presented in Table 7 below.

Table 7: Comparison of Mean Scores and Adjusted Mean Scores on Test

<table>
<thead>
<tr>
<th>Test</th>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>Adjusted Mean</th>
<th>Difference</th>
<th>$F$</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Experimental</td>
<td>17</td>
<td>69.5588</td>
<td>64.505</td>
<td>18.702</td>
<td>16.831</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td>Control</td>
<td>41</td>
<td>43.7073</td>
<td>45.803</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Final Examination

The final examination of Calculus II for both cohorts comprised five questions, with each question having several parts. As it was not possible to cover all the topics taught in the course in a 2.5-hour written examination, some topics were tested in flipped class and not the non-flipped class, and vice-versa. For example, the flipped class had a moderately difficult question on showing that a given real-valued function of two variables is continuous but not differentiable at the origin, whereas the non-flipped class did not have question involving differentiability. Another challenging question that appeared in the flipped class but not the non-flipped class was on term-by-term differentiation and integration of power series, and using the resulting power series to find the sums of an infinite series and a power series. On the other hand, the non-flipped class had a question on showing that a sequence defined by recurrence relation is convergent and finding its limit, but the flipped class did not have question involving such a sequence. Some topics were tested in examination papers of both flipped and non-flipped classes, such as finding the limits of two infinite sequences, and finding the radius of convergence of a power series. For questions on topics common to both examination papers, the ones in flipped class were generally more difficult than those in the non-flipped class in most cases. Overall, the examination paper for the flipped class was more demanding than that of the non-flipped class.
Consequently, results of an application of ANCOVA revealed that the adjusted mean score on the final examination of the control group was higher than that of the experimental group but the difference was not statistically significant at the 5% level. The mean scores and adjusted mean scores on the final examination and the F values are presented in Table 8 below.

Table 8: Comparison of Mean Scores and Adjusted Mean Scores on Final Exam

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>Adjusted Mean</th>
<th>Difference</th>
<th>F</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exam</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experimental</td>
<td>17</td>
<td>63.8529</td>
<td>58.940</td>
<td>-8.829</td>
<td>12.126</td>
<td>.001</td>
</tr>
<tr>
<td>Control</td>
<td>41</td>
<td>65.7317</td>
<td>67.769</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Calculus II Overall

The assessment for both the flipped and non-flipped Calculus II courses consisted of continual assessment and the 2.5-hour final examination, with weighting of 40% and 60% respectively. The continual assessment comprised the 15-minute Quizzes 1 through 3, the 55-minute test and class participation.

Although the overall Calculus II mean score for the experimental group was higher than that of the control group, after controlling for the Calculus I scores, the adjusted mean score of the control group was higher than that of the experimental group. The difference, however, was not statistically significant at the 5% level. The overall mean scores and adjusted overall mean scores on the Calculus II courses and the F values are presented in Table 9 below.

Table 9: Comparison of Overall Mean Scores and Adjusted Overall Mean Scores

<table>
<thead>
<tr>
<th>Group</th>
<th>N</th>
<th>Mean</th>
<th>Adjusted Mean</th>
<th>Difference</th>
<th>F</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Experimental</td>
<td>17</td>
<td>71.0471</td>
<td>66.371</td>
<td>-1.907</td>
<td>.635</td>
<td>.429</td>
</tr>
<tr>
<td>Control</td>
<td>41</td>
<td>66.3390</td>
<td>68.278</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

6. Conclusion

This paper has reported on the design of a flipped calculus course for second year undergraduates in a university in Singapore and results of retrospective analyses of students’ achievement in mathematics, measured by scores on 3 quizzes, a test and the final examination, between a flipped calculus class and a non-flipped one. Results of applications of ANCOVA show that the flipped Calculus II class scored statistically significantly higher in the 55-minute written test but lower in the 2.5-hour final examination than the non-flipped Calculus II class, and that the differences in performance between the two classes in the three 15-minute quizzes were not statistically significant.

We are aware that most sources of extraneous variables are usually controlled by random assignment of participants but this could not be arranged given the timetabling constraints we faced. In this study, that both the experimental and control groups were taught by the same instructor had helped eliminate any variability due to different teaching styles. However, we were unable to compare a class of students taking a flipped Calculus II course with another class taking a non-flipped Calculus II course in the same semester because in recent academic years, cohort sizes for mathematics majors have been small (not more than 20 for second year programme in particular). The small number of participants also limits the application of this study to the population of undergraduates in Singapore. Furthermore, in spite of our effort in ensuring that the cognitive
demands of the quizzes, the test, and the final examination for the experimental group were comparable to those for the control group, that the questions for the experimental group were not identical to those for the control group might still be a factor which could confound the results of our analyses.

Nevertheless, as this study was intended to generate preliminary insights for a further study which we have planned, we believe that much could still be gleaned from this work even though we fully appreciate and acknowledge the difficulty which limitations of the present study could cause in establishing statistically valid results. In the planned larger study, we will have an opportunity to reach more definitive conclusions about the effectiveness of the flipped classroom model in the teaching and learning of undergraduate mathematics in the Singapore context.

References: