

Chapter Eleven

Effect of Learning-by-Teaching in a Flipped Classroom

Victoria Li¹, Wendy Bailey², & Ray Littlejohn²

¹Department of Applied Mathematics, University of Colorado Boulder, USA, ²Lockheed Martin Engineering Management Program, University of Colorado Boulder, USA

Motivation

A growing US university student population will require adaptation of teaching styles in order to meet educational demands without sacrificing quality for quantity. According to the National Center for Education Statistics, enrollment in degree-granting institutions between 2002 and 2012 increased 24 percent, from 16.6 million to 20.6 million (U.S. Dept. of Education 2015). Specifically, undergraduate enrollment increased 24 percent and post baccalaureate enrollment increased 24 percent. Given constraints on university resources in terms of time and space, how can we maintain the quality of learning and teaching experiences? Research shows that large teaching loads correlate to decreased teacher satisfaction and likewise large lecture sizes correlate to decreased student satisfaction (Monks and Schmidt 2010). Low student satisfaction correlates to decreased student retention rates (Cuseo 2007). The reasons, simply explained, are that students need to be treated as individuals not as numbers, and education is meant to be an experience not an industry.

With the creation of Khan Academy in 2006 emerged the idea of the flipped classroom as a promising teaching style (Fink 2011). Rather than having students listen to lectures in class and do homework on their own, as in traditional teaching styles, flipped classrooms have students watch lectures on their own and work on problems in class. There are many variations of the flipped classroom, but it typically consists of readings/videos/reading questions before class followed by problem-solving/collaboration/hands-on activities in class (O’Flaherty and Phillips 2015; Winqvist and Carlson 2014).

Advantages of the flipped classroom include increased student responsibility for learning, higher scores, teamwork, development of interpersonal skills, student innovation, ability to learn at one's own pace, and ability to handle the increased student numbers and decreased state/institutional funding in today's education system (O'Flaherty and Phillips 2015). The success of the flipped classroom may be explained by the generation effect and the testing effect, both of which have been found to improve students' learning, and both of which are included in the flipped classroom (Winguist and Carlson 2014). The generation effect is the finding that knowledge generated by students is recalled at a higher rate than knowledge given to them from reading (Slamecka and Graf 1978). For example, having students generate the antonym to a word (i.e. given "hot", generate "cold") makes them remember the word pair better than having them simply read it (i.e. "hot-cold"). Meanwhile the testing effect shows that testing can also be used to improve students' long-term memory and learning, for example students who answer test questions prior to a final test outperform students who study for the same test (Bertsch et al. 2007). Indeed, a study comparing a traditional version of an introductory undergraduate statistics course to a flipped classroom version that used pre-reading questions (testing) and in-class exercises (generating knowledge) found that flipped classroom students did significantly better on the Psychology Area Concentration Achievement Test (ACAT) Statistics scale a year later, even accounting for student quality (ACAT non-Statistics scale scores).

There are also a number of disadvantages to flipped classrooms (O'Flaherty and Phillips 2015): The main complaint from students, when it did arise, was that there was too much student responsibility for learning. Instructor complaints mainly centered around the enormous amount of up-front preparation required. Use of technology for in-class activities is not necessarily a good thing: A flipped classroom based on an automated tutoring system had subsequent issues

with grasp of concepts (Strayer 2012), and a flipped classroom using apps to highlight teaching points showed no difference in effectiveness from a traditional classroom (Martin et al. 2013). However, the flipped classroom studies that used clickers or pre-lecture videos were met with improved grades and student satisfaction (Ferreri and O'Connor 2013; McLaughlin et al. 2013; Gilboy et al. 2015; Yeung and O'Malley 2014). Previous findings demonstrated the greater importance of content and student engagement over particular resources used, and showed a tendency of online resources to improve lower order rather than higher order cognitive skills. Several studies suggested a need for IT support personnel to prepare videos and content, and a need for "professional educators" to address issues of instructor ability and understanding regarding flipped classrooms. Lastly, while there is much qualitative data on the effectiveness of flipped classrooms, there are few quantitative studies of their effectiveness, and results tend to be limited to moderate improvements in exam scores (O'Flaherty and Phillips 2015; Winquist and Carlson 2014).

Purpose

The flipped classroom model implemented in our study could offer a solution to the above problems that also retains the advantages of flipped classrooms. EMEN 5005, Introduction to Applied Statistical Methods, is a flipped classroom course in the University of Colorado Boulder Engineering Management Program that uses pre-reading/questions before class, clicker questions (testing effect) during class, and homework after class. We proposed adding a learn-by-teaching activity (generation effect) to help the course take full advantage of what flipped classrooms have to offer. By providing students with a real-life data set (College Scorecard (U.S. Dept. of Education 2016)) and having them generate and answer their own clicker questions about the

lecture material (IU8: One-Sample Continuous and Discrete Hypothesis Testing), we engaged students' higher cognitive skills in a research-like activity. This could help increase students' responsibility for their own learning, balanced by instructor guidance and examples that were provided for them. In addition, the use of an online dataset integrated technology at reasonable levels. There was a low amount of preparation required, with no need for professional educator support. Quantitative analysis of results was based on grades (IU8 homework scores) and a control (non-IU8 homework scores) to account for quality of students, while qualitative results were obtained through a student survey. If successful, such a learn-by-teaching method could be considered as a solution to declining student retention amidst a growing student population.

Methods

We implemented a learn-by-teaching activity in the fall 2015 EMEN 5005 class taught by Dr. Ray Littlejohn, and used data from the spring 2015 and Fall 2014 EMEN 5005 classes, also taught by Dr. Ray Littlejohn, for control analyses. The fall 2015 class was comprised of 33 distance learner and on campus graduate students, while the spring 2015 and fall 2014 classes were comprised of a mix of distance/on campus and graduate/undergraduate students totaling 63 and 69 students respectively. Of the 63 students in spring 2015, there were 52 graduate students. While all fall 2015 on campus students participated in the activity, among them 15 agreed to participate in the study, resulting in a study sample size of 15.

The above three classes were all taught using a flipped classroom lecture style consisting of pre-readings/questions before class, clicker questions during class, and homework after class. Only the Fall 2015 IU8 (Instructional Unit 8) lectures received the learn-by-teaching activity: Students were given a subset (CollegeScorecard2013Data.xlsx,

CollegeScorecardDataDictionary.xlsx) of the College Scorecard dataset and given the entire class period (75 minutes) to come up with questions relating the lecture material to the dataset. In addition, example clicker questions were provided (examples.pptx). Students were free to collaborate, discuss, and receive guidance from the professor (Dr. Littlejohn), faculty advisor (Wendy Bailey), and PI (Victoria Li) throughout the activity. Once students had generated enough questions, they tried answering them using MVPStats software and their knowledge of hypothesis testing. Figure 1 shows the timeline of the study.

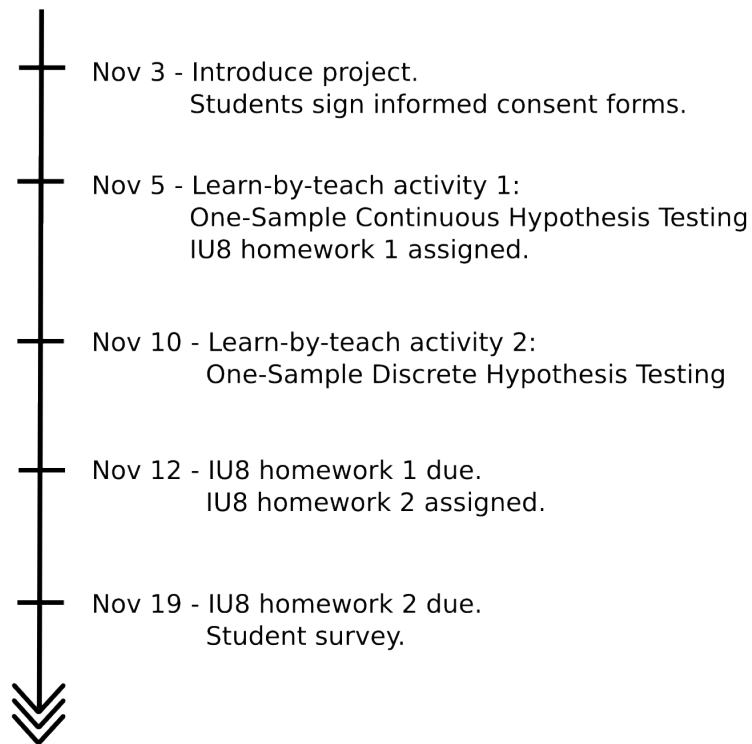


Figure 1: *Timeline of learn-by-teaching study.*

The learn-by-teaching activity was a rewarding experience for both students and teachers. Students were able to learn practical skills in MVPStats and Excel while solving problems, with the professor present to troubleshoot their work (Figure 2).

- Q:** What do you do when you encounter NULL and Private values in the data?
- A:** Set the value to -999 and change MVPStats settings to only include positive data values.
- Q:** What do you do when MVPStats says "Error: sample size must be $n > 3$ " but the sample size is > 3 ?
- A:** Use the comparison matrix to calculate correlation/p-value instead. We then verified the result using R and Python.
- Q:** Two groups both calculated the correlation between control of institution and tuition, and got different results. This happened using both in-state and out-of-state tuition.
- A:** This is because the point-biserial test can only be used for two-outcome nominal data, and control of institution is more than two-outcome.

The professor also showed students shortcuts in Excel.

Figure 2: Student questions during learn-by-teaching activity.

During the course of the activity, students got a sneak peek into a future lecture on regression from an interesting fact shared by the professor (Figure 3).

Create a scatterplot of the data, convert the x and y values to z-scores, and the slope will be the correlation coefficient (works for continuous data only).

Figure 3: Interesting fact shared by professor.

Some student-generated questions and answers are shown in Figure 4.

- Q:** What is the correlation between average SAT score and cost of tuition?
A: Higher SAT score correlates to higher cost of tuition with $r = 0.53$, $p = 0$.
- Q:** What is the correlation between female debt and male debt? Grad debt?
A: Female debt and grad debt are correlated with $r = 0.6$, female debt and male debt are correlated with $r = 0.5$, all with $p < 0.05$.
- Q:** What is the correlation between control of institution and tuition (in/out-of-state)?
A: They are correlated with $p = 0$.

Figure 4: Student-generated questions and answers.

As the survey results will show, these types of interactions were key reasons for student satisfaction with learning, while prolonged lack of interaction caused student complaints.

Quantitative Results

Anonymized homework scores were collected for all three classes: fall 2015, spring 2015, and fall 2014. We determined the IU8 and average non-IU8 homework score for each student, tested the samples for normality, compared average homework scores across the three classes (Figure 5), and conducted power analysis. For homework score comparison, we chose not to use ANOVA because: 1) IU8 and non-IU8 homework scores are repeated measures, 2) the sample sizes were unequal (different numbers of students in each class), and 3) due to low sample sizes we could not equalize the sample sizes by throwing out data. The techniques required to handle this type of ANOVA would have complicated the ANOVA analysis beyond our means. Instead, we used Welch's t test to compare fall 2015 (study participants) and spring 2015 (graduate students) IU8 scores, Welch's t test to compare fall 2015 (study participants) and spring 2015 (all students) non-IU8 scores, and z test to compare spring 2015 (all students) and fall 2014 (all students) non-IU8 scores.

Pearson's chi-square tests for normality resulted in rejection of normality of the spring 2015 (graduate students), spring 2015 (all students), and fall 2014 (all students) non-IU8 homework data (all $p < 0.05$). In all cases we assumed that the distribution of IU8 homework scores was the same as that of non-IU8 homework scores. Since the sample sizes for the above samples were 52, 63, and 69 respectively, however, we assumed the samples were normal by the Central Limit Theorem. With a sample size of 15, the fall 2015 (study participants) data could not be tested for normality ($df = 0$). However, it is reasonable to assume that student grades have a normal distribution, and we made that assumption here.

Although fall 2015 (study participants) IU8 homework scores ($\bar{x} = 96.3$, $s = 4.7$) were higher on average than spring 2015 (graduate students) IU8 homework scores ($\bar{x} = 95.9$, $s = 7.3$), Welch's t test shows that they were not significantly different ($p < 0.05$). To account for differences in student quality, we compared fall 2015 (study participants) non-IU8 homework scores ($\bar{x} = 90.0$, $s = 4.2$) to spring 2015 (all students) non-IU8 homework scores ($\bar{x} = 88.7$, $s = 7.3$), and verified by Welch's t test that they were not significantly different ($p < 0.05$). Lastly, we were interested in whether fall and spring semester grades differed historically. Comparison of spring 2015 (all students) non-IU8 homework scores ($\bar{x} = 88.7$, $s = 7.3$) to fall 2014 (all students) non-IU8 homework scores ($\bar{x} = 84.1$, $s = 12.8$) by z test found that spring 2015 grades were significantly higher than fall 2014 grades ($p < 0.05$). Results are summarized in Figure 5.

| | Fall 2015 study participants | Spring 2015 graduate students all students |
|----------------|---|--|
| IU8 | $\bar{x} = 96.3$ $s = 4.7$ $n = 15$ | $\bar{x} = 95.9$ $s = 7.3$ $n = 52$ |
| | Welch's t test \longleftrightarrow | |
| non-IU8 | $\bar{x} = 90.0$ $s = 4.2$ $n = 15$ | $\bar{x} = 88.7$ $s = 7.3$ $n = 63$ |
| | Welch's t test \longleftrightarrow | |
| | Spring 2015 all students | Fall 2014 all students |
| non-IU8 | $\bar{x} = 88.7$ $s = 7.3$ $n = 63$ | $\bar{x} = 84.1$ $s = 12.8$ $n = 69$ |
| | z test \longleftrightarrow | |

Figure 5: Statistical test results.

The power of the test was computed for each of the above three statistical tests (Figure 6). The fall 2015 vs. spring 2015 IU8 comparison had effect size 0.063, power 0.055, and significance level 0.05. The low effect size and power was due to the small study sample size of 15 and a difference in sample mean homework scores of only .4 points, whereas the study was designed to detect homework differences of 5 points (which would yield effect size 7.3 and power 0.69 in this case). Thus there was a high probability of incorrectly failing to reject the null hypothesis. Similarly, the fall 2015 vs. spring 2015 non-IU8 comparison had effect size 0.17 and

power only 0.091. The spring 2015 vs. fall 2014 non-IU8 comparison, however, had large sample sizes 63 and 69, effect size 0.44, and power 0.81. This met the standard for acceptable power of a test.

| Comparison | Test | Effect Size | Power |
|-----------------------------------|-----------|-------------|-------|
| Fall 2015 vs. Spring 2015 IU8 | Welch's t | 0.063 | 0.055 |
| Fall 2015 vs. Spring 2015 non-IU8 | Welch's t | 0.17 | 0.091 |
| Spring 2015 vs. Fall 2014 non-IU8 | z | 0.44 | 0.81 |

Figure 6: Power analysis results.

Survey Results

Study participants were surveyed to qualitatively understand their attitudes about the learn-by-teaching experience. All study participants filled out the survey, and their responses are attached (survey responses.pdf). Overall, the result was that 9 of 15 students (60%) would prefer their classes be taught with the learn-by-teaching lecture style, if they had to choose yes or no.

Whatever their preference, though, almost all students agreed that the time given was more than what was necessary for the activity. The other complaint from students was the lack of structure. Suggestions for improving this were:

- Go over the example questions and their solutions first (we chose to display them but not go over them).
- Make the activity more challenging.
- Make use of the product (for example have students present their work in front of the class).
- Conclude with a summary and discussion.

Interestingly, the amount of interaction with instructors and peers seemed to be the major determinant of student satisfaction/dissatisfaction. Students enjoyed working in a stress-free and flexible environment where they could practice on their own, discuss with peers, and have quick access to instructor guidance when needed. Students did not enjoy lack of feedback and not knowing whether they answered the question right or not. However, only two students encountered this situation, indicating that the majority of students benefitted from increased interaction in the learning activity. Regarding learning outcomes, students generally wrote that the activity taught them to 1) apply MVPStats and 2) make and test statistical hypotheses. In addition, three students wrote that they learned to use the appropriate correlation test for the appropriate data type involved.

Discussion

Quantitatively, the learn-by-teaching activity did not significantly change students' homework scores either for better or for worse. This is probably due to the fact that the class was a flipped classroom to begin with, and the learn-by-teaching activity only made it slightly more so.

Qualitatively, the learn-by-teaching activity primarily affected student-student and student-teacher interactions, as well as students' abilities to apply MVPStats software for statistical testing. The general consensus from survey results was that the learn-by-teaching activity taught hands-on skills, balanced individual work with peer and instructor guidance, and could benefit from more structure. Student survey results indicate that success of the flipped classroom is contingent upon the amount of interaction achieved.

Our findings indicate that learning-by-teaching in flipped classrooms would be suitable for a growing university population. This combination of teaching styles requires low

preparation time, allows students to benefit from interactions with each other and instructors, emphasizes application and troubleshooting of knowledge, and requires less classroom time to achieve the same learning outcomes. By providing students with guidance where they need it most, while decreasing in-class lecture load of instructors, learning-by-teaching could potentially maintain student satisfaction and quality of learning in large classrooms. However, our quantitative study was unable to obtain a large enough study sample size to achieve adequate statistical power, nor to reasonably comment on the effectiveness of this teaching style for large classroom sizes. Future work should also study instructors' perspectives on this teaching style. Thus future studies in large classrooms and diverse fields will be needed to support our findings.

TAR Experience

My TAR experience has been eye opening and very positive from beginning to end, partly because of the people, partly because of the research. In terms of people, Laura Border and Adam Blanford have a contagious enthusiasm for teaching that truly comes from the heart, and really know what they're doing. From summer workshop to project mentorship to meeting old and new TAR students to post-project follow-up, this has been a well-organized and highly educational year and a half experience. I love how the focus is on research (hence, Teaching As Research) because the best way to learn teaching well is to apply it.

I have always been fascinated by the flipped classroom, but never encountered it at CU until TAR gave me the chance to do research on it. One of the workshop speakers, Wendy Bailey, teaches flipped classrooms at FRCC, and her perspectives on engineering management connected with me on a deep level. She introduced me to Dr. Ray Littlejohn, who kindly agreed to help us conduct a study on EMEN 5005 Intro to Applied Statistical Methods. From planning

to IRB approval to classroom implementation to results, I benefited greatly from their statistics and teaching expertise, and gained experience being a project PI.

I learned many things from my TAR project, including the result of the study itself. They are:

- How to write a protocol
- How to conduct power tests and tests for normality
- When ANOVA should and should not be used
- Application of statistics vs. theory of statistics
- Engineering Management vs. Applied mathematics approaches to statistics
- How to calculate a correlation coefficient from the slope of a line
- Students liked having on-demand guidance and problem-solving in class
- Students requested more instruction at the beginning of the class and a chance to present/make use of their work at the end of the class
- Quantitatively, the learning-by-teaching class style did not significantly improve or worsen student homework scores
- Qualitatively, the effectiveness of the learning-by-teaching class style increased with amount of interaction
- The flipped classroom can benefit students, teachers, and administrators by making education enjoyable and convenient while effectively using class time and space

Looking back, it's pretty cool to have contributed to cutting edge education research, and to be publishing my results in the 2017 TAR Edited Volume.

Special Thanks

We give special thanks to Laura Border, director of CU Boulder CIRTl and Co-PI of CU Boulder TIGER, and to Adam Blanford, TIGER TAR project advisor, for their mentorship.

Funding: TIGER TAR is funded by CU Boulder CIRTl (NSF grant #1231286). The funders had no role in data collection and analysis, decision to publish, or preparation of this manuscript.

Appendix

All attachments available at:

<https://drive.google.com/open?id=0BxLlpYqj2joZWfY4b2FxZEIJMVk>

References Cited

- Bertsch, S., B. J. Pesta, R. Wiscott, and A. McDaniel. 2007. The generation effect: a meta-analytic review. *Memory & Cognition* **35**: 201-210.
- Cuseo, J. 2007. The empirical case against large class size: adverse effects on the teaching, learning, and retention of first-year students. *The Journal of Faculty Development* **21**(1): 5-21.
- Ferreri, S. and S. O'Connor. 2013. Redesign of a large lecture course into a small-group learning course. *American Journal of Pharmaceutical Education* **77**(1): 1-9.
- Fink, Z. (Producer). 2011. Salman Khan on Liberating the Classroom for Creativity. *Big Thinkers Series*. Video retrieved from <http://www.edutopia.org>
- Gilboy, M., S. Heinerichs, and G. Pazzaglia. 2015. Enhancing the student engagement using flipped class. *Journal of Nutrition Education and Behaviour* **47**(1): 109-114.
- Martin, S., J. Farnan, and V. Arora. 2013. FUTURE: New strategies for hospitalists to overcome challenges in teaching and today's wards. *Journal of Hospital Medicine* **8**(7): 409-413.
- McLaughlin, J. E., L. M. Griffin, D. A. Esserman, C. A. Davidson, D. M. Glatt, M. T. Roth, N. Gharkholonarehe, and R. J. Mumper. 2013. Instructional design and assessment: Pharmacy student engagement, performance, and perception in a flipped satellite classroom. *American Journal of Pharmaceutical Education* **77**(9): 1-8.

- Monks, J. and R. Schmidt. 2010. The impact of class size and number of students on outcomes in higher education. *Cornell University ILR School Working Papers*. digitalcommons.ilr.cornell.edu.
- O’Flaherty, J. and C. Phillips. 2015. The use of flipped classrooms in higher education: a scoping review. *Internet and Higher Education* **25**: 85-95.
- Slamecka, N. J. and P. Graf. 1978. The generation effect: delineation of a phenomenon. *Journal of Experimental Psychology: Human Learning & Memory* **4**: 592-604.
- Strayer, J. 2012. How learning in an inverted classroom influences cooperation, innovation and task orientation. *Learning Environments Research* **15**: 171-193.
- U.S. Department of Education, National Center for Education Statistics. 2015. *Digest of Education Statistics, 2013* (NCES 2015-011), [Chapter 3](#).
- U.S. Department of Education. 2016. *College scorecard data*. Retrieved from <https://collegescorecard.ed.gov/data>.
- Winqvist, J. and K. A. Carlson. 2014. Flipped statistics class results: better performance than lecture over one year later. *Journal of Statistics Education* **22**(3): 1-10.
- Yeung, K. and P. O’Malley. 2014. Making “The Flip” work: Barriers to and implementation strategies for introducing flipped teaching methods into traditional higher education courses. *New Directions for Institutional Research* **10**(1): 59-63.