CAN A SIMULATION-BASED INFERENCE COURSE BE FLIPPED?

Noelle Pablo and Beth Chance
Department of Statistics, Cal Poly – San Luis Obispo, CA 93407
bchance@calpoly.edu

"Simulation-based inference" has been advocated with the potential of improving student understanding of statistical inference, as well as the statistical investigative process as a whole. One justification is that the approach calls for improved pedagogy (i.e., more active learning and use of technology to explore statistical ideas). The "flipped classroom," where students spend class time working on explorations and out-of-class time reading the text and watching videos, has also been gaining popularity in recent years. But can a simulation-based inference (SBI) course be flipped? In this study, the same instructor taught an SBI course as a flipped class and in a more traditional format during the same term. We explore differences in student attitudes, conceptual understanding, and course performance between the two sections.

INTRODUCTION

Student attitudes, prior knowledge, delivery method, and classroom environment are all potentially critical components in understanding the varying levels of comprehension and success in an introductory statistics course. In an effort to positively influence student attitudes and performance in introductory statistics, many instructors have been implementing more active learning, student-focused technology, and simulation into their courses. In particular, "simulation-based inference" (SBI) has been advocated in the introductory course with the potential to improve students' conceptual understanding and attitudes towards statistics. Preliminary assessment results have supported this (e.g., Beckman, delMas, & Garfield, 2017; Chance, Wong, & Tintle, 2016; Hildreth, Robison-Cox, & Schmidt, 2018). Several studies have also examined the impact of classroom design and delivery method on students' success and attitudes in statistics. For example, Gundlach, Richards, Nelson, & Levesque-Bristol (2015), studied one instructor teaching the same content with three methods of delivery: traditional lecture, flipped, and fully-online, examining the impact of the method of delivery on students' attitudes, course evaluations, and conceptual understanding of statistics. The study found significant differences only in the attitudes categories of affect and perceived easiness, with the traditional format scoring higher on both. Traditional students also scored significantly higher on average on all three exams. A study by Hedges (2017) aiming to learn more about undergraduate students' performance and anxiety in an introductory statistics course found homework grades, student persistence, and test and class anxiety to be significantly different between traditional courses and an online course. Students in the online classroom had stronger homework scores but revealed higher anxiety levels and a higher withdrawal rate from the course. In this study, we seek to examine differences in undergraduate students' attitudes, performance, and course and instructor evaluation between two class sections taught by the same instructor in the same term (Fall, 2015). In particular, how do the novel course content (SBI) and the novel pedagogy (flipped) interact with each other? Can the focus on conceptual understanding and active learning also translate to the flipped classroom?

COURSE CONTENT

The textbook *Introduction to Statistical Investigations* (*ISI;* Tintle et al., 2016) was the primary text for both sections, and web-based applets were used for all analyses in the course. Distinguishing features of this text include a focus on the scope and logic of statistical inference. Students are introduced to the "statistical investigation process" in the first week of the course, and web-based applets are utilized to allow students to estimate p-values in those first investigations. They then spiral through the same ideas in the one mean, two proportion, two mean, multiple groups, and regression settings.

DELIVERY METHODS

Table 1 highlights distinctions between the flipped and lecture/lab deliveries. For both sections, during lab time both the instructor and an undergraduate teaching assistant were available

In M. A. Sorto, A. White, & L. Guyot (Eds.), *Looking back, looking forward. Proceedings of the Tenth International Conference on Teaching Statistics (ICOTS10, July, 2018), Kyoto, Japan.* Voorburg, The Netherlands: International Statistical Institute. iase-web.org

Table 1. Description of two delivery methods

	Flipped Classroom	Lecture/Lab
Number of students	34	34
Meeting times	4 days/week, 1-2pm	4 days/week 2-3pm
Classroom	Collaborative lab (34 computers, in rows of 3 computers with shared screen at end of row, lib.calpoly.edu/study-spaces-and-tech/library-spaces/all/.	Two days classroom (computer projection); two days library lab (34 computers)
Pre-class assignments	Reading (with optional videos) and reading quiz before each class period	Suggested reading
Graded written assignments	4 exploration reports/week, jointly written in on-line wiki (3-4 students), with formative assessment by instructor	2 lab reports/week, jointly written (2 students)

to answer questions. Though the instructor had taught the course in the lecture/lab format for many years, this was the first full flipped classroom for her. The main distinction between the two courses was the additional structure and the amount of class time spent discussing the material and demonstrating concepts as a whole class.

In addition to the written assignments, both classes completed weekly quizzes, two mid-term exams, and one final exam. The exams were identical, but the quizzes differed in nature. Both sections also completed the Survey of Students Attitudes Toward Statistics (SATS-36; http://evaluationandstatistics.com/) at the start of the course and during the last week of classes. Both classes also completed a concept inventory developed by the authors of the *ISI* text based on the multiple-choice CAOS instrument (https://apps3.cehd.umn.edu/artist/caos.html). These two instruments were to be completed outside of class at the beginning and the end of the course. Students were given homework points for submitting their name.

DEMOGRAPHICS

In the Lecture/Lab section, 55.9% completed the SATS-36 survey and opted in to allowing us to use their results for our research, compared to 64.7% for the Flipped section. These percentages were 76.5% and 73.5% for the concept inventory. Table 2 shows basic demographic information for the two sections; none of the differences were statistically significant.

Table 2. Comparison of demographic data for the two sections

	Lecture/Lab (n = 19)	Flipped (n = 22)
First generation	6/19 ≈ 32%	4/22 ≈ 18%
Female	14/19 ≈ 74%	19/22 ≈ 86%
Mean college GPA	3.14	3.15
Sophomore (2nd year)	18/19 ≈ 95%	21/22 ≈ 95%
Caucasian	12/19 ≈ 63%	14/22 ≈ 64%

CHANGES IN STUDENT ATTITUDES

Table 3 examines the pre/post/change in attitudes on the six SATS-36 subscales for the Lecture/Lab section and the Flipped section. One student was removed from the Flipped section because his responses were quite different from other students (perhaps reversing the scaling on the Likert items). Another who did not take the post test is included in the pretest summaries but not the change variable. Cronbach α values were also considered for the SATS-36 and results were very consistent with other published studies (e.g., Schau & Emmioğlu, 2012), with *Difficulty* the only subscale below 0.70.

Two-sample *t*-tests were used to compare the mean change from the pre to the posttests between the two sections for each attitude component. Though none of the differences were significant, students in the flipped classroom saw a larger increase in average competency (do they feel they can do Statistics, $p = 0.1383$) and in difficulty (higher scores indicate perception of a less difficult course, $p = 0.1431$). For difficulty, the flipped students thought the course would be more

difficult at the start of the course, but their perception was more similar at the end of the course. Though not significant, there was a higher increase in affect and a smaller decrease in value and in interest for the flipped students. (The negative changes are consistent with other published results.) Much of the lack of significance is due to large student to student variation in changes.

Table 3. Comparison of pre and post attitudes in the course for the two sections

Lecture/Lab (n = 19) Attitude Component	Pre		Post		Change (post-pre)	
	M	*SD*	*M*	*SD*	*M*	*SD*
Effort	6.342	0.501	5.303	1.098	-1.039	0.895
Affect	3.963	0.862	4.325	1.208	0.361	1.173
Competence	4.658	1.167	4.689	0.928	0.031	0. 980
Difficulty	3.699	0.561	3.791	0.724	0.092	0.617
Value	5.145	0.857	4.842	0.936	-0.303	0.844
Interest	4.803	0.729	3.987	1.300	-0.816	1.193
Flipped Attitude Component	Pre (n = 22)		Post (n = 21)		Change (post-pre)	
	M	*SD*	*M*	*SD*	*M*	*SD*
Effort	6.545	0.648	5.667	0.735	-0.857	0.903
Affect	3.656	0.750	4.048	0.711	0.448	0.821
Competence	4.197	0.654	4.698	0.747	0.52	1.070
Difficulty	3.247	0.448	3.646	0.640	0.395	0.658
Value	5.303	0.651	5.148	1.065	-0.143	0.751
Interest	4.864	0.738	4.440	1.275	-0.369	1.060

GAINS IN CONCEPTUAL UNDERSTANDING

Both sections showed higher scores on average on the post-test compared to the pre-test (paired *t*-test p-values below 0.00001). Figure 1 shows the similarities in these distributions. The main *gains* were 0.189 (lecture/lab) and 0.179 (flipped), with SD = 0.144 and 0.125. To guard against ceiling effects, we also examined *achievable gains* = *gain*/(1 – *pre*). Means and standard deviations were 0.334 (0.252) for lecture/lab, and 0.326 (0.212) flipped. Two-sample *t*-tests of mean gains (p = 0.817) and achievable gains (p = 0.919) were not statistically significant.

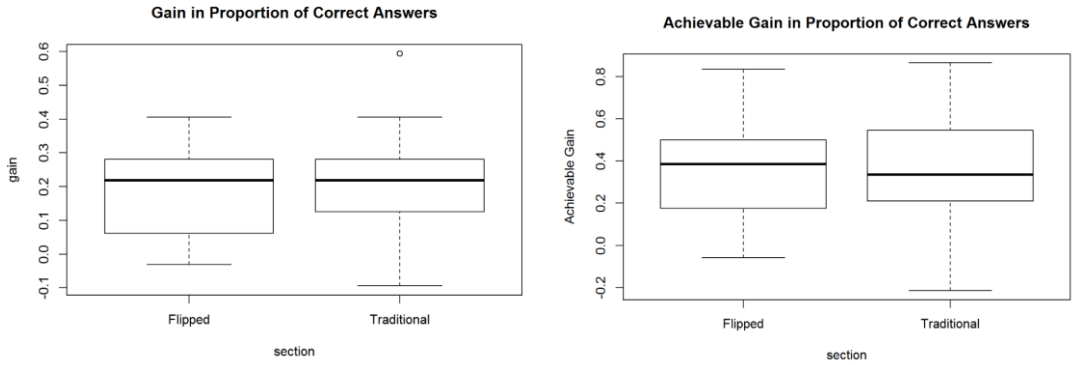

Figure 1. Comparison of gains and achievable gains on the concept inventory

STUDENT FEEDBACK

Student responses to the institution-administered course evaluations were also examined. Categories include teaching techniques, overall instructor rating, instructor created interest, and overall course rating. Chi-square tests of the ratings given across the two sections were not significant (p-values: 0.3587, 0.4701, 0.2219, 0.2689 respectively). From course evaluations in the flipped sections, most students indicated comfort with the daily format of reading quiz, brief discussion, group work on wiki; 20 of 31 agreed or strongly agreed that they liked the format. Of the available resources, students appeared to mostly value the on-line practice questions, the reading quizzes, and the wiki submissions with feedback. The perceived value of the online videos was more mixed. In open ended responses, there were a handful of negative comments about "having to teach ourselves" and too much group work. Another question asked them to mark which

statements about the course they most agreed with. The most common selections were "I liked working together in a group" (22), "I really liked the 'hands on' nature of the course" (21), and "I wish the instructor had done more lecturing on the readings" (22). Almost as popular were the statements "The instructor's feedback on the wikis before final submission was helpful" (17) and "the reading quizzes led me to complete more of the reading assignments than I probably would have otherwise" (15). Only 5 students reported difficulty with the technology or applets. When asked which components would be most helpful to add to the course, "more worked out examples" far outweighed other options including a data analysis project, homework submitted for a grade, and more online technology demos. In both sections, about two-thirds of students found the class somewhat interesting and enjoyable (61.3% flipped and 62% lecture/lab) or very interesting and enjoyable (32.3% flipped and 28% lecture/lab) and most were satisfied with their expected grade (64.5% flipped, 83% lecture/lab). In the flipped classroom, 13% were not happy and blamed the course/instructor, compared to 3.5% lecture/lab. Also in the flipped section, three students indicated they could not predict their grade, compared to zero in the lecture/lab section.

DISCUSSION

Similar to Gundlach et al. (2015), we did not find large differences in student performance or attitudes between the lab/lecture section and the flipped section. It is important to keep in mind special characteristics of the two implementations: The lab/lecture section also involved a large active learning component, some group work, heavy use of technology, and focus on conceptual understanding; the flipped section primarily relied on reading assignments outside of class, with less use made of videos accompanying the text. The distinction between the class formats was not large enough to demonstrate any loss of student performance or attitude with the flipped approach, even with the SBI curriculum. It is still quite plausible that more differences would be found with more distinct implementations, especially controlling for instructor effects. We are also exploring interesting interactions between prior attitudes, student gender, and GPA on post attitudes and gains in conceptual understanding.

In the end, most students in the flipped section were satisfied with their course performance, and a few specifically liked the format, but a handful of students did find the format detrimental to their learning. Most of the students felt there were sufficient resources, but the most commonly requested addition to the course materials was additional worked out examples. The effectiveness of the flipped format with the SBI course is quite similar to a course with more traditional content and sequencing, and no large differences were found between the two formats, though the flipped style is not for every student. Further research is needed on these factors.

REFERENCES

Beckman, M., delMas, R., & Garfield, J. (2017). Cognitive transfer outcomes for a simulation-based introductory statistics curriculum, *Statistics Education Research Journal*, *16*(2), 419-440.

Chance, B., Wong, J., & Tintle, N. (2016). Student performance in curricula centered on simulation-based inference: A preliminary report. *Journal of Statistics Education, 24*(3), 114-126.

Gundlach, E., Richards, K.A.R., Nelson, D., & Levesque-Bristol, C. (2015). A comparison of student attitudes, statistical reasoning, performance, and perceptions for web-augmented traditional, fully online, and flipped sections of a statistical literacy class. *Journal of Statistics Education*, *23*(1).

Hedges, S. (2017). Statistics student performance and anxiety: Comparisons in course delivery and student characteristics. *Statistics Education Research Journal, 16*(1), 320-336.

Hildreth, L., Robison-Cox, J., & Schmidt, J. (2018). Comparing student success and understanding in introductory statistics under consensus and simulation-based curricula. *Statistics Education Research Journal, 17*(1).

Schau, C. & Emmioğlu, E. (2012). Do introductory statistics courses in the United States improve students' attitudes? *Statistics Education Research Journal, 11*(2), 86-94.

Tintle, N. L., Chance, B., Cobb, G., Rossman, A., Roy, S., Swanson, T., & VanderStoep, J. (2016), *Introduction to Statistical Investigations*. New York: John Wiley and Sons.