International Journal of Engineering Education Vol. 36, No. 5, pp. 1615–1624, 2020
Printed in Great Britain

0949-149X/91 $3.00+0.00

Comparing Student Performance in Flipped and Non-Flipped Space Mechanics Classrooms*

KAELA M. MARTIN
Aerospace Engineering Department, Embry-Riddle Aeronautical University, 3700 Willow Creek Road, Prescott, AZ, USA.
E-mail: kaela.martin@erau.edu

JONATHAN M. GALLIMORE
University of Wisconsin – Madison, 1202 West Johnson Street, Madison, WI, USA. E-mail: jmgallimore@wisc.edu

Recent advances in active learning have shown that flipped learning is an effective pedagogical approach, but few studies have examined the differences between a flipped and non-flipped course across numerous semesters with the same instructor teaching both formats. This study examines the performance of students in an upper-level space mechanics course in flipped and non-flipped settings over seven semesters. Students in the flipped course performed better on peer evaluation, homework, and exams leading up to the final exam. The course is required for all aerospace engineering students who can decide to pursue the aeronautics (aircraft) track or the astronautics (spacecraft) track. After flipping the course, students in the astronautics track had larger gains in performance compared to in the aeronautic track implying that the flipped course helped students that could readily perceive how they would use the course concepts in the future (astronautics). When examining the viewing habits of students in the flipped courses, watching more videos was associated with higher grades on course elements. Students who watched videos for a longer time and viewed more unique videos received higher scores on graded course elements. This study implies that flipped learning is a successful pedagogy that improves student learning in an upper-level engineering course.

Keywords: flipped classroom; active learning; aerospace engineering

1. Introduction

Successful educators relentlessly pursue the most effective methods to teach students. Oftentimes "most effective" depends on the course learning objectives, the students, the course material, and the instructor. Recent pedagogical movements in the Scholarship of Teaching and Learning have suggested effective, evidence-based strategies to improve student learning. For example, active learning is a pedagogy that has been shown to produce better student learning and outcomes than traditional lecturing [1–4].

Active learning includes a broad range of classroom interventions and pedagogies such as the jigsaw classroom, think-pair-share, four corners, peer instruction, Team-Based Learning (TBL), Problem Oriented Guided Inquiry Learning (POGIL), gamification, and flipped or hybrid courses. Learner-Centered Teaching (LCT) also has active learning at its core. In LCT, the instructor becomes a facilitator and creates an environment that maximizes student engagement through "authentic, meaningful, and useful learning" [5].

Active learning and LCT can benefit engineering students and curriculum by providing realistic and significant learning experiences. Furthermore, these pedagogies require students to actively use the material they are learning. Learning and memory research has shown that actively using and recalling information lead to learning gains [6]. One such pedagogy that requires students to recall and actively use the course concepts during class is flipped learning.

1.1 What is Flipped Learning?

Many different terms have been used to describe flipped learning such as hybrid, inverted, or blended courses. The most clear-cut definition of flipped learning is that students do the work outside of class that they would normally do in class and vice-versa [7]. Flipped learning exposes students to new concepts in an individual space where students work alone or in small informal groups rather than a group space where students work in a formal group (which could be the entire class). The group space is "transformed into a dynamic, interactive learning environment" where the instructor helps students with the application and synthesis of the course concepts [8]. In flipped learning, the individual space can include watching videos to prepare for class but can include other types of preparation such as readings or discovery learning. The group space, often class time, is transformed into an engaging environment where students complete homework, work through example problems, or actively utilize the course concepts in a meaningful manner.

* Accepted 24 March 2020.

1.2 Why Flipped?

In a traditional classroom, students are introduced to the course material for the first time in class and then apply what they learned outside of the classroom with little instructor oversight. In this model, there is an inverse relationship between the cognitive difficulty of the work and access to support [8]. In other words, students need the instructor the most when they are individually working on problems. Furthermore, the traditional classroom fails to promote life-long learning by emphasizing the "instructor as the source and gatekeeper of knowledge" [8]. Life-long learning is a key part of Accreditation Board of Engineering and Technology's (ABET) student outcome criteria seven stating that students should have the "ability to acquire and apply new knowledge" [9].

Many engineering courses have a large volume of information that not only must be learned but used in a meaningful manner. Flipped learning often includes peer instruction which falls under the LCT umbrella and promotes the idea that the person who does the work learns the material [10]. In peer instruction students teach and learn from each other [11] which results in students actively recalling information leading to better long-term retention of the course concepts [6]. Peer instruction can also lead to improved team working skills [12], deep-level questioning [13], increased motivation [14], and better student attendance [15]. In a traditional classroom, students often comment that they understood a concept in class but struggled with the concept on a homework assignment. Peer instruction can reduce this problem because students must deeply process the course concepts to explain them to another student [16].

1.3 Previous Research on Flipped Courses

In the last 10 years, engineering instructors have been utilizing a flipped course design with mixed success. A great review of the engineering flipped learning literature can be found in Karabulut-Ilgu, Cherrez, and Jahren [17] which extended an earlier review article by Bishop and Verleger [18].

From its beginnings [7], flipped learning studies have seen improvement in student performance [2, 18, 19] resulting in positive student opinions about the class and increased homework and exam scores [20–22]. Unsurprisingly, flipped classrooms have also shown increased engagement and involvement due to the large amount of active learning [22, 23].

Flipped learning research has produced mixed results when comparing the effectiveness of flipped and traditional courses. Most studies reported positive improvements (e.g., higher final grades, increased classroom engagement, or better opinions about the class) [17], but of the studies that reported statistical tests, only seven found significant effects [24–30]. Alternatively, three studies directly compared a flipped course to a traditional course taught by the same instructor and found the flipped courses to be less effective in terms of final exam score or final grades than the traditional courses [31–33].

Of the seven papers that found significant benefits for flipped learning studies [24–30], six considered undergraduate courses. In a first-year engineering course, Ossman and Bucks found a significant decrease in the drop, failure, and withdraw rate for flipped courses compared to previous years when sections were taught by different instructors [26]. Likewise, Yelamarthi and Drake examined a first-year digital circuits course and found significant increases in course grades for a flipped course compared to a traditional course which led to higher retention and persistence [29]. Papadopoulos and Roman found increased pre- and post-test grades in a statics class for both traditional and flipped classrooms, but due to small sample sizes, the difference was not significant [27]. In a study by Schmidt, students who watched pencasts in a dynamics class had significant learning gains compared to a traditional course [28]. Mason, Shuman, and Cook flipped an undergraduate controls systems course which was able to cover more content and found a significant performance increase on three of five problem types in the class [25]. Finally, Fedesco and Troy found significant increases in quiz scores in a flipped upper-level civil engineering course [30].

Of these ten papers [24–33], three found positive effects and were published in journals [24, 25, 29]. All papers that found that flipped learning was less effective were published at conferences where most flipped learning papers are published [17]. Karabulut-Ilgu, Cherrez, and Jahren also looked at the mean scores of 25 studies that reported traditional and flipped course scores. When they controlled for the author, they found a significant difference in mean final score leading to the conclusion that learning in flipped classrooms is at least the same if not better than traditional classrooms [17].

1.4 Videos in Previous Flipped Courses

A common element in flipped learning is videos that students watch before attending class. Most of the research on student video watching comes from Massive Online Open Courses (MOOCs) [34–37] which typically have student completion rates of less than 10% [38]. Studies that consider video watching outside of MOOCs rely on two sources of data, student self-reported data and video management data [39]. The self-reported data showed that most students watched the videos [40] and their viewing habits changed over the semester [25]. Video management data showed that students actually watched the videos [41–43], but only select content for a homework solutions video [42–44]. Other studies using video management data found that video watching increased as the semester progressed and peaked directly before an exam [39, 45]. Students were also more likely to watch the videos if they were held accountable (e.g., a quiz on the material) [39].

Video length also varies among previous studies. Math and science MOOCs tend to have an average video length of 7.75 minutes [36]. Students were more likely to completely watch videos that were less than 3 minutes long, and the median watch time was 6 minutes regardless of total video length [35]. In the flipped learning studies that found significant effects, video length (when reported) ranged from 15 minutes or less [25, 28, 30] to 30 minutes [26]. When videos were more than 30 minutes long, students reported them as boring [46]. In the studies that found negative or no effects, video length was 20 minutes [32] or students viewed one video per week (a total of 12) for the entire course [33]. While video lengths for the latter study were not specified, it is probable that the videos were longer than 15 minutes to cover the course content for a week. It could be that the length of these videos hindered learning by exceeding student's attention capacity. The literature we reviewed (and personal experience) suggested that videos should be about 5 to 8 minutes.

To date, few studies have compared traditional and flipped classrooms that cover the same material for an entire semester [21, 47, 48] making it unclear how flipped learning influence students' learning relative to more "traditional" learning. Furthermore, video metrics are often absent when comparing traditional and flipped courses [39]. We examine student performance in a flipped upper-level space mechanics course and compare it to the same course taught in a non-flipped structure. All classes were taught by the same instructor. This paper extends our previous work [49, 50] and will provide evidence to support the effectiveness of the flipped classroom.

2. Methods

The considered course is an upper-level space mechanics course that is required for all aerospace engineering students to take. At the studied university, a private university in the southwestern United States, aerospace engineering students must choose one of two tracks, aeronautics (aircraft) or astronautics (spacecraft). The aeronautics students most often take space mechanics in their senior year and do not see the value in taking the course. The astronautics students take the course in their junior year as part of their core curriculum.

The course was originally taught using lecture with some active learning. Students would sit in groups and work though short example problems in groups during class. Each weekly homework assignment (a total of 12) was broken into three parts where the first two homework assignments (formative homework) were graded on completion and the third was graded on correction (summative homework). Students also had weekly in-class quizzes (11 or 12) and four exams. The exams included a group take-home exam and an individual exam. In fall 2016, some topics were introduced using videos and class time was spent going over content and working through the homework. Students were held accountable for watching the videos by answering questions using a video response tool called Recap. The student experience in fall 2016 was closer to the original, non-flipped course rather than the flipped course.

In spring 2017, the course was transformed into a flipped learning course as the instructor wanted a more active classroom experience. Students in the flipped course watched a series of three to eight minute videos before coming to class with a maximum watching time of 25 minutes. They then completed an online quiz of three multiple-choice questions that were graded on completion before coming to class and working on worksheets in groups. The problems from the worksheets primarily came from the formative homework assignments, but students still had to complete one or two questions on the formative homework assignments. The non-flipped course taught during the same semester had the identical notes which they worked through in class and completed the online quizzes after the class period. The non-flipped course had longer formative homework assignments but identical exams, in-class quizzes, and summative homework assignments.

Only minor changes to the flipped learning structure have been made since spring 2017. In fall 2017, the homework was reduced to one assignment per week (the summative homework), and additional videos describing the learning objectives for each set of videos were created. No changes were made to the course in spring 2018, but in fall 2018, the online quizzes were removed in favor of 17 readiness assessments inspired by TBL. These readiness assessments were three to four multiple-choice question quizzes that students took individually and then immediately after as a group. No other changes were made in spring 2019.

The study considered 12 sections of space mechanics taught over eight semesters. Three sections were taught in a non-flipped manner spring 2016 (n = 10), fall 2016 (n = 31), and spring 2017 (n = 20) with a total of 61 participants. Nine sections of the flipped course were included in the study for spring 2017 (n = 20), fall 2017 (n = 26, n = 38), spring 2018 (n = 16, n = 31), fall 2018 (n = 41, n = 34), and spring 2019 (n = 17, n = 33) with a total of 256 participants. The participants do not include four non-aerospace engineering students or the seven students that stopped coming to class or turning in homework. The student body at the studied university is 75% male, 62% white, and primarily full-time, traditional students.

Data were collected from students after the courses were completed. Informed consent was obtained for students starting from spring 2017, and students that did not consent were omitted from the study. We obtained IRB approval to use grade data for spring 2016 and fall 2016. Grade data were collected after the semester ended for participating students. GPA was collected using our student management system and was reported for the incoming semester. Three transfer students did not have GPA in the student management system, so we used the mean GPA of 3.29 for these students.

The grade weightings for each semester are shown in Table 1 where NF refers to a non-flipped course and F refers to a flipped course. Exams included a take-home portion (group) worth 30% of the exam grade as well as an individual portion worth 70% of the exam grade with a total of four exams including the final. The lowest of the first three individual and group exams, not including the final, were dropped as well as the lowest formative homework score. When two sections were taught (spring 2017 to spring 2019), the first three individual exams had similar problems, and all group exams as well as the individual final were identical. Note that the quizzes in fall 2016 refer to video response quizzes and in fall 2018 and spring 2019 refer to TBL-style quizzes that were all graded on correctness. All other quizzes refer to online multiple-choice quizzes graded on completion. The lowest two quizzes across all semesters were excluded from the final grade. Class engagement for spring 2016 to spring 2018 includes in-class quizzes (graded on correctness) and worksheets (graded on completion) for the flipped sections. Starting in fall 2018, class engagement refers to only the in-class worksheets after group quizzes were dropped in fall 2018. The lowest two class engagement scores were excluded from the final grade. Even though fall 2016 had some videos, it is still considered a non-flipped course as the teaching style was closer to the non-flipped courses than the flipped courses.

Across all semesters, students worked in small groups in class as well as completed group (take-home) exams. Due to the group work, students completed peer evaluations at the end of the semester. From spring 2016 to fall 2018, students were told to give their peers and themselves a grade between 0 and 100. In spring 2019, students were given a set number of points that they had to divide among their group members. If everyone participated equally, each student's score would be 100; if not, students received more or less points. Starting in fall 2018, group exams were weighted by peer evaluations beginning with the second exam using CATME Peer Evaluation [51].

To remove the weighting differences across semesters, each element was instead given a set number of points. The total course points (3150) were allocated across the class elements as follows: homework (1200), class engagement (790), quizzes (660), peer evaluations (100), and exams (400).

Video data was collected from an online video management system, Kaltura. The system collected the number of unique video views, the total number of video views (could be larger than the number of available videos if a student watched a video more than once), the total video view time, the average drop (percent of a video completed), the number of loads, and the loads to plays ratio (total number of video views divided by the number of video loads) for each participant.

Table 1. Grade Weightings for Studied Semesters

	SP16 NF	**FA16 NF**	**SP17 NF**	**SP17 F**	**FA17 F**	**SP18 F**	**FA18 F**	**SP19 F**
Ind Exams	35%	31.5%	31.5%	31.5%	31.5%	31.5%	31.5%	31.5%
Group Exams	15%	13.5%	13.5%	13.5%	13.5%	13.5%	13.5%	13.5%
Formative Hw	21%	18%	20%	10%	–	–	–	–
Summative Hw	19%	17%	15%	15%	25%	25%	25%	25%
Class Engagement	5%	5%	5%	15%	15%	15%	15%	15%
Quizzes	–	10%	10%	10%	10%	10%	10%	10%
Peer Evaluation	5%	5%	5%	5%	5%	5%	5%	5%
Total	100%	100%	100%	100%	100%	100%	100%	100%

Table 2. Descriptive Statistics for Variables Used

	n	Mean	Median	Std. Dev.	Min.	Max.
GPA	317	3.29	3.27	0.45	1.86	4.00
Homework	317	1032	1074	146	313	1235
Class Engagement	317	754	768	52.4	485	822
Quizzes	307	592	603	66.6	330	660
Peer Evaluations	317	95.1	96.0	7.3	48.0	140
Exams	317	346	347	27.6	257	401
Unique Views	256	124	131	36.8	5.0	174
Total Views	256	95.9	97.0	57.2	0.0	365
Total View Time (min)	256	408	414	241	0.0	1503
% Video Completed	256	88.4	90.0	14.2	0.0	122
Video Loads	256	174	165	73.8	5.0	416
Loads to Plays (%)	256	53.7	55.6	23.0	0.0	95.1

3. Results and Discussion

Before starting the analysis, variables were checked for normality. Seven students were removed as outliers because they stopped coming to class, turning in homework, and/or taking exams as well as four non-aerospace engineering students. After removing these students, the variables were normally distributed. Table 2 displays descriptive statistics for the variables used in this study. The sample size differs for quizzes (see Table 1) as the spring 2016 section did not have quizzes. The video metrics were only available for the flipped sections which reduced the sample size to 256.

Bivariate correlations were used to examine the relationship among the variables (see Table 3). Unsurprisingly, as cumulative GPA increased so did every other variable. Students with higher GPAs obtained higher scores on course elements and video metrics. Overall, the class elements (i.e., homework, class engagement, quizzes, peer evaluations, and exams) were positively correlated with each other, except peer evaluations and quizzes. That is, when students did better on one class element (e.g., homework), they also did better on other class elements (e.g., quizzes and exams).

In terms of video metrics, all video metrics showed significant positive correlations with one another. For example, students that viewed more unique videos also had a higher percentage of the videos that were completed. Some of the class elements showed positive correlations with video metrics but others showed no relationship. As homework points increased so did the total number of videos viewed and total amount of time spent viewing the videos. When students engaged in the class more, they also watched more videos, had longer total view times, completed a higher percentage of the videos, and played more of the videos that were loaded. As quiz points increased so did the number of video loads and the number of unique video views. As peer evaluations points increased, the total number of videos viewed, the total video view time, and the ratio of loaded videos to played videos also increased. Finally, as exam scores increased so did all video metrics. Put another way, students that used the videos more obtained more points on exams. No other correlations were significant.

Since all class elements were correlated, we utilized a 2 X 2 Multivariate Analysis of Covariance (MANCOVA) to compare the flipped and non-flipped courses with GPA as a covariate. Due to course format differences across semesters, quizzes were not included in the analysis because these elements were not present each semester (see Table 1). Class engagement was also not included in the analysis as elements that were included in class engagement were dissimilar between the flipped and non-flipped courses. Exams, peer evaluation, and homework, however, were similar across sections. Therefore, our Dependent Variables (DVs) were exam, peer evaluation, and homework points. Our factors were the course format (flipped or non-flipped) and student's aerospace engineering track (aeronautics or astronautics). GPA was a significant covariate, $F(3, 310) = 31.2$, $p < 0.05$, *Wilks'* $\Lambda = 0.628$, *partial* $\eta^2 = 0.37$. There was significant difference between the course format on a linear combination of the DVs, $F(3, 310) = 7.33$, $p < 0.001$, *Wilks'* $\Lambda = 0.934$, *partial* $\eta^2 = 0.07$. There was not a significant difference between student's aerospace engineering track on a linear combination of the DVs, $F(3, 310) = 1.41$, $p = 0.24$, *Wilks'* $\Lambda = 0.986$, *partial* $\eta^2 = 0.01$. However, there was significant interaction between the course format and the student's aerospace engineering track, $F(3, 310) = 3.05$, $p < 0.05$, *Wilks'* $\Lambda = 0.971$, *partial* $\eta^2 = 0.03$. For significant effects (see Table 4), we examined the estimated marginal means for each DV with pairwise comparisons. Exam points were not different between flipped and non-flipped

Table 3. Correlations among Class Elements and Video Metrics

	GPA	Hw	Class Engage.	Quizzes	Peer Evals	Exams	*Unique Views*	*Total Views*	*Total View Time*	*% Video Completed*	*Video Loads*
Homework	0.41****										
Class Engagement	0.29****	0.53****									
Quizzes	0.24****	0.39****	0.29****								
Peer Evaluations	0.33****	0.24****	0.25****	0.03							
Exams	0.55****	0.44****	0.30****	0.40****	0.28****						
Unique Views	0.15*	0.07	0.11	0.21***	0.01	0.26****					
Total Views	0.30****	0.14*	0.21***	0.09	0.16*	0.33****	0.68****				
Total View Time	0.29****	0.15*	0.24***	0.008	0.14*	0.33****	0.64****	0.98****			
% Video Completed	0.17**	0.11	0.17**	–0.01	0.03	0.14*	0.23***	0.35****	0.43****		
Video Loads	0.13*	0.06	0.11	0.16*	0.03	0.23***	0.80****	0.73****	0.69****	0.15*	
Loads to Plays	0.32****	0.12	0.22***	–0.01	0.25****	0.27****	0.24****	0.73****	0.74****	0.44****	0.15*

* p < 0.05; ** p < 0.01; *** p < 0.001; **** p < 0.0001.
Note: The video variables (in italics) only examined the flipped students (n = 256). The other variables include all students (n = 317) except quizzes which include all but spring 2016 (n = 307).

courses, but the flipped courses had more peer evaluation (*mean difference* = 2.1, $p < 0.05$) and homework (*mean difference* = 72.1, $p < 0.001$) points than the non-flipped courses. For the interaction, the astronautic students scored more points on homework in the flipped course compared to the non-flipped course (*mean difference* = 124, $p < 0.001$). No other differences were significant.

Interestingly, exam points were not different for both pedagogical approaches (flipped and non-flipped) suggesting that students' exam performance did not depend on the pedagogy. Perhaps this result is due to students' intense focus on exams and exam performance, and many students are determined to succeed regardless of what happens in class. Looking back, we think that there was no effect because the exam scores included the final exam and students' best two out of three "regular" exams. The studied university only assigns flat letter grades (A, B, C, D, F) creating large grade steps (~10% of points possible in class). When students take the final exam (which is 12.5% of a student's grade), many realize they either need a few points to keep their current grade or reach the next grade step or almost all the available points because they are at the bottom of a grade step. These factors influence many students to be strategic with the final exam and try for a set number of points or not try at all. Therefore, as a follow-up analysis, we examined students' best two exams without the final exam and compared flipped and non-flipped courses while covarying GPA. The flipped students had higher exam scores (M = 181) than the non-flipped students (M = 177) on their best two exams, $F(1, 314) = 5.83$, $p < 0.05$, *partial* $\eta^2 = 0.02$. This finding suggests that the flipped courses produce better exam scores than the non-flipped courses when controlling for GPA, but this effect was muted by the final exam. This gain in exam scores may be explained by the increase in homework scores in the flipped courses because the homework prepared students for the exams. If we had only examined students' final or total exam scores, like previous research [31–33], we might have missed evidence supporting the effectiveness of the flipped courses on exam scores.

Next, we analyzed the video metrics from the flipped courses to determine whether student's video use could predict any aspect of classroom performance, filling a gap identified by Ahn and Bir [39]. We utilized a series of multiple linear regressions with backwards selection to remove video metrics that did not predict course elements. For these analyses, the dependent variables were the course elements (i.e., homework, class engagement, quizzes, peer evaluations, and exams), and the independent variables were the video metrics (i.e.,

Table 4. Course Elements by Track and Course Type

Dependent Variable	Track	Flipped	Not Flipped
Exam Points		345 (1.45)	346 (2.95)
	Aeronautics	342 (2.16)	349 (4.27)
	Astronautics	348 (1.93)	344 (4.07)
Peer Evaluation Points		**95.5 (0.43)**	**93.4 (0.89)**
	Aeronautics	95.0 (0.65)	92.6 (1.28)
	Astronautics	96.0 (0.58)	94.2 (1.22)
Homework Points		1045 (8.14)	973 (16.6)
	Aeronautics	1035 (12.16)	1015 (24.02)
	Astronautics	**1055 (10.82)**	**931 (22.87)**

Note: Estimated marginal means are displayed with GPA covaried and standard error of the mean is in parentheses. Bold numbers indicate a significant mean difference.

unique views, total views, total view time, percent of video completed, video loads, and loads to plays). Class engagement and quizzes were included because these elements were similar across the flipped sections. Recall that the maximum number of points for homework was 1200, class engagement 790, quizzes 660, peer evaluations 100, and exams 400.

We only reported the last step of each regression model (see Table 5) after the nonsignificant predictors were omitted from the regression equation. Across all regression analyses, three video metrics (total view time, unique views, and loads to plays) predicted the points earned on course elements. When total view time increased by 1 minute, homework points increased by 0.08 and class engagement increased by 0.16. When unique views increased by one video, quiz points increased by 0.35 and exam points increased by 0.15. When loads to plays increased by 1%, exam points increased by 0.27 points and peer evaluations increased by 0.07 points. While these point increases may not seem large, the course included around 170 videos. For example, students watching all videos would have an increase of 58.3 points on quizzes and 26.1 points on exams. Taken together, student performance on graded course elements increased when students viewed more unique videos, spent more time viewing videos, and played a higher percent of the videos than were loaded. We believe that this finding indicates that students' increased knowledge of the course concepts was most likely due to their increased exposure to the course concepts via videos.

Over numerous semesters of data, we found that video use metrics were related to student performance and predicted a student's score on graded course elements. Taken together, the videos helped students perform better on all course elements. The more unique videos that students viewed and the longer they viewed those videos, the higher the score they obtained on graded course elements. We believe that the key to students watching the videos is the short video length (usually less than 7 minutes).

Broadly speaking our overall goal with this paper was to evaluate the effectiveness of a flipped course and explore how students' use of videos could predict their scores on graded course elements. Overall, we found that students in flipped courses earned more peer evaluation and homework points than non-flipped courses. This finding implies that flipped learning helps students with their homework because students work on problems similar

Table 5. Video metrics predicting performance on course elements

Dependent Variable		*t*	*p*	β	*F*	*df*	*p* <	*adj.* R^2
Homework	**Overall Model**				5.60	1,254	0.05	0.02
	Total View Time	2.37	<0.05	0.15				
Class Engagement	**Overall Model**				9.33	1,254	0.001	0.06
	Total View Time	2.63	<0.01	0.82				
Quizzes	**Overall Model**				11.1	1,254	0.001	0.04
	Unique Views	3.33	<0.001	0.20				
Peer Evaluations	**Overall Model**				16.5	1,254	0.001	0.06
	Loads to Plays	4.07	<0.001	0.25				
Exams	**Overall Model**				16.2	2,253	0.001	0.11
	Unique Views	3.32	0.001	0.20				
	Loads to plays	3.69	0.001	0.23				

to their homework during class and receive immediate feedback from an expert (i.e., their instructor). However, the biggest gain in homework points was for students that easily perceived how they would use the course concepts in their future career (astronautics) versus students that might not have perceived how they would use these concepts in their future careers (aeronautics).

4. Limitations and Future Directions

This study considered a single course with a single instructor at a single institution with a homogenous population, which limits the generalizability of our results. Just like other pedagogies, flipped learning is highly dependent on the instructor, and it is easy to apply flipped learning in a non-effective manner. Therefore, we recommend that instructors read the literature on flipped learning and try flipping elements or parts of the course before flipping the entire course.

Another limitation to the study was the growth of the instructor over the study. At the beginning of the study (spring 2016), the instructor had been teaching full-time for a single semester but by the end of the study (spring 2019) had four years of experience. During the study, the instructor discovered and implemented new teaching techniques that slightly changed the course from semester to semester, and these changes could have led to improved student learning. The flipped course design was implemented after the non-flipped design was used for three semesters. Therefore, one possible explanation for our results is instructor maturation leading to better student performance. To rule out this hypothesis, we conducted a one-way ANCOVA with the overall points that students earned as the dependent variable, the year the course was taken as the factor, and GPA as a covariate. Students earned a similar amount of points in the course regardless of the year that the course was taken, $p > 0.05$, implying that overall student performance cannot be explained by instructor maturation.

The course changed each semester via feedback from students and peers. Assignments were not always identical across semesters, but some course elements (homework, exams, and peer evaluations) were similar throughout the study. Our analyses standardized the differences in course element weightings across the study, but students may have performed differently based on the weightings.

Even with these limitations, we believe that this study fills gaps in the flipped learning literature. Future studies should continue to compare flipped learning to more traditional learning (or lecture only) because the non-flipped courses in this study included some active learning concepts such as solving problems in teams. Additionally, future research should examine all graded courses elements similar to the current study instead of primarily focusing on exams. It would also be interesting to know when students watched the videos. It is possible that student performance on graded course elements could be influenced by the time of day that the videos are viewed.

5. Conclusion

Implementing flipped learning in the classroom led to better student performance. We found that students in an upper-level aerospace engineering space mechanics course performed better on peer evaluations and homework in the flipped course. This performance increase was especially significant for students in the astronautics (or spacecraft) track. While total exam differences across the flipped and non-flipped sections were not different, students in the flipped courses performed better on the exams leading up to the final primarily due to better performance on homework.

Students that watched more videos performed better in class. Interestingly, the only video metrics that predicted points on course elements were unique video views, total view time, and loads to plays. The duration of the videos that students watched did not predict students' performance on graded course elements but starting the video did.

This study adds to the growing literature on flipped learning and suggests that implementing a flipped course leads to increased student performance compared to a non-flipped course.

Acknowledgements – This research was partially supported by the Embry-Riddle Aeronautical University (ERAU) Faculty Research Development Program. The authors want to extend their sincerest gratitude to Dr. Dina Battaglia and ERAU Prescott's Center for Teaching and Learning Excellence (CTLE) for the guidance with flipped course design and active learning strategies.

References

1. R. M. Felder and R. Brent, *Teaching and Learning STEM: A Practical Guide*, Jossey-Bass, San Francisco, CA, 2016.
2. S. Freeman, S. L. Eddy, M. McDonough, M. K. Smith, N. Okoroafor, H. Jordt and M. P. Wenderoth, Active learning increases student performance in science, engineering, and mathematics, *Proceedings of the National Academy of Science*, **111**(23), pp. 8410–8415, 2014.
3. M. J. Prince, Does active learning work? A review of the research, *Journal of Engineering Education*, **93**(3), pp. 223–231, 2004.

4. C. E. Wieman, Large-scale comparison of science teaching methods sends clear message, *Proceedings of the National Academy of Sciences*, **111**(23), pp. 8319–8320, 2014.
5. T. Doyle, *Learner-Centered Teaching*, Stylus, Sterling, VA, 2011.
6. H. L. Roediger and J. D. Karpicke, Test-Enhanced Learning: Taking Memory Tests Improves Long-Term Retention, *Psychological Science*, **17**(3), pp. 249–255, 2006.
7. M. J. Lage, G. J. Platt and M. Treglia, Inverting the classroom: A gateway to creating an inclusive learning environment, *Journal of Economic Education*, **31**(1), pp. 30–43, 2000.
8. R. Talbert, *Flipped Learning: A Guide for Higher Education Faculty*, Stylus, Sterling, VA, 2017.
9. ABET criteria, https://www.abet.org/accreditation/accreditation-criteria/criteria-for-accrediting-engineering-programs-2019-2020/#GC3, Accessed 12 December 2019.
10. T. Doyle, *Helping Students Learn in a Learner Centered Environment: A Guide to Teaching in Higher Education*, Stylus, Sterling, VA, 2008.
11. D. W. Johnson, G. Maruyama, R. Johnson, D. Nelson and L. Skon, The Effects of Cooperative, Competitive, and Individualistic Goal Structures on Achievement: A Meta-Analysis, *Psychological Bulletin*, **89**(1), pp. 47–62, 1981.
12. J. E. Miller and J. E. Groccia, Are four heads better than one? A comparison of cooperative and traditional teaching formats in an introductory biology class, *Innovative Higher Education*, **21**(4), pp. 253–273, 1997.
13. G. Marbach-Ad and P. G. Sokolove, Can undergraduate biology students learn to ask higher level questions?, *Journal of Research in Science Teaching*, **37**(8), pp. 854–870, 2000.
14. A. O'Donnell, The role of peers and group learning, in P. A. Alexander and P. H. Winne (eds), *Handbook of Educational Psychology*, 2nd edn, Lawrence Earlbaum, Mahwah, NJ, pp. 781–802, 2006.
15. S. F. Schomberg, 1986, Involving High Ability Students in Learning Groups, *American Educational Research Association Annual Meeting*, San Francisco, CA, April 16–20, 1986.
16. M. D. Svinicki and W. J. McKeachie (eds.), *McKeachie's Teaching Tips*, 14th edn, Wadsworth, Belmont, CA, 2014.
17. A. Karabulut-Ilgu, N. J. Cherrez and C. T. Jahren, A systematic review of research on the flipped learning method in engineering education, *British Journal of Engineering Technology*, **49**(3), pp. 398–411, 2018.
18. J. Bishop and M. A. Verleger, The Flipped Classroom: A Survey of the Research, *ASEE Annual Conference & Exposition*, Atlanta, GA, June 23–26, 2013.
19. B. H. Ferri, A. A. Ferri, D. M. Majerich and A. G. Madden, Effects of an In-class Hands-On Laboratories in a Large Enrollment, Multiple Section Blended Linear Circuits Course, *Advances in Engineering Education*, **5**(3), 2016.
20. J. O'Flaherty and C. Phillips, The use of flipped classrooms in higher education: A scoping review, *The Internet and Higher Education*, **25**, pp. 85–95, 2015.
21. C. G. Merrett, Applying the IMPULSE Method to Second and Fourth Year Aerospace Engineering Courses, *AIAA SciTech Forum*, Grapevine, TX, January 8–13, 2017.
22. R. M. Clark, M. Besterfield-Sacre, D. Bundy, K. M. Bursic, W. W. Clark, B. A. Norman, R. S. Parker, J. F. Patzer II and W. S. Slaughter, Flipping Engineering Courses: A School Wide Initiative, *Advances in Engineering Education*, **5**(3), 2016.
23. H. Ritz and L. Schneider-Bentley, Collaborative Problem Solving at Chalkboard vs. On Paper For First-Year Calculus, *Advances in Engineering Education*, **6**(3), 2018.
24. Y. Chiang and H. C. Wang, Effects of the in-flipped classroom on the learning environment of database engineering, *International Journal of Engineering Education*, **31**(2), pp. 454–460, 2015.
25. G. S. Mason, T. R. Shuman and K. E. Cook, Comparing the Effectiveness of an Inverted Classroom to a Traditional Classroom in an Upper-Division Engineering Course, *IEEE Transactions on Education*, **56**(4), pp. 430–435, 2013.
26. K. A. Ossman and G. W. Bucks, Effect of Flipping the Classroom on Student Performance in First-Year Engineering Courses, *ASEE Annual Conference & Exposition*, Indianapolis, IN, June 15–18, 2014.
27. C. Papadopoulos and A. S. Roman. Implementing an inverted classroom model in engineering statics: Initial results, *ASEE Annual Conference & Exposition*, Louisville, KY, June 20–23, 2010.
28. B. Schmidt, Improving motivation and learning outcome in a flipped classroom environment, *2014 International Conference on Interactive Collaborative Learning*, Dubai, UAE, December 3–6, 2014.
29. K. Yelamarthi and E. Drake, A Flipped First-Year Digital Circuits Course for Engineering and Technology Students, *IEEE Transactions on Education*, **58**(3), pp. 179–186, 2015.
30. H. N. Fedesco and C. Troy, Why this Flip Wasn't a Flop: What the Numbers Don't Tell You About Flipped Classes, *ASEE Annual Conference & Exposition*, New Orleans, LA, June 26–29, 2016.
31. C. J. McClelland, Flipping a Large-enrollment Fluid Mechanics Course – Is it Effective?, *ASEE Annual Conference & Exposition*, Atlanta, GA, June 23–26, 2013.
32. E. J. Hagen and D. Fratta, Hybrid Learning in Geological Engineering: Why, How, and to What End?, *Geo-Congress 2014 Technical Papers: Geocharacterization and Modeling for Sustainability*, Atlanta, GA, February 23–26, pp. 3920–3929, 2014.
33. J. P. Lavelle, M. T. Stimpson and E. D. Brill, Flipped out engineering economy: Converting a traditional class to an inverted model, *Industrial Systems Engineering Research Conference*, San Juan, Puerto Rico, May 18–22, pp. 397–407, 2013.
34. N. Li, L. Kidziński, P. Jermann and P. Dillenbourg, MOOC Video Interaction Patterns: What Do They Tell Us?, *Design for Teaching and Learning in a Networked World*, Toledo, Spain, September 15–18, pp. 197–210, 2015.
35. P. J, Guo, J. Kim and R. Rubin, How Video Production Affects Student Engagement: An Empirical Study of MOOC Videos, *ACM Conference on Learning @ Scale Conference*, Atlanta, GA, March 4–5, pp. 41–50, 2014.
36. J. Kim, P. J. Guo, D. T. Seaton, P. Mitros, K. Z. Gajos and R. C. Miller, Understanding In-Video Dropouts and Interaction Peaks in Online Lecture Videos, *ACM Conference on Learning @ Scale Conference*, Atlanta, GA, March 4–5, pp. 31–40, 2014.
37. L. Breslow, D. E. Pritchard, J. DeBoer, G. S. Stump, A. D. Ho and D. T. Seaton, Studying Learning in the Worldwide Classroom: Research into edX's First MOOC, *Research & Practice in Assessment*, **8**, pp. 13–25, 2013.
38. K. F. Hew and W. S. Cheung, Students' and Instructors' Use of Massive Open Online Courses (MOOCs): Motivations and Challenges, *Educational Research Review*, **12**, pp. 45–58, 2014.

39. B. Ahn and D. D. Bir, Student Interactions with Online Videos in a Large Hybrid Mechanics of Materials Course, *Advances in Engineering Education*, 2018.
40. S. Zappe, R. Leicht, J. Messner, T. Litzinger and H. W. Lee, Flipping the Classroom to Explore Active Learning in a Large Undergraduate Course, *ASEE Annual Conference & Exposition*, Austin, TX, June 14–17, 2009.
41. S. L. Dazo, N. R. Stepanek, R. Fulkerson and B. Dorn, An Empirical Analysis of Video Viewing Behaviors in Flipped CS1 Courses, *ACM Conference on Innovation and Technology in Computer Science Education*, Arequipa, Peru, July 11–13, 2016.
42. T. Pinder-Grover, K. R. Green and J. M. Millunchick, The Efficacy of Screencasts to Address the Diverse Academic Needs of Students in a Large Lecture Course, *Advances in Engineering Education*, **2**(3), 2011.
43. K. R. Green, T. Pinder-Grover and J. M. Millunchick, Impact of Screencast Technology: Connecting the Perception of Usefulness and the Reality of Performance, *Journal of Engineering Education*, **101**(4), pp. 717–737, 2012.
44. S. P. Gross and D. W. Dinehart, Pre- and Post-class Student Viewing Behaviors for Recorded Videos in an Inverted Sophomore Mechanics Course, *ASEE Annual Conference & Exposition, New Orleans*, LA, June 26–29, 2016.
45. G. K. Kinsella, C. Mahon and S. Lillis, Using Pre-Lecture Activities to Enhance Learner Engagement in a Large Group Setting, *Active Learning in Higher Education*, **18**(3), pp. 231–242, 2017.
46. R. Olson, Flipping Engineering Probability and Statistics – Lessons Learned for Faculty Considering the Switch, *ASEE Annual Conference & Exposition*, Indianapolis, IN, June 15-18, 2014.
47. J. A. Day and J. D. Foley, Evaluating a Web Lecture Intervention in Human-Computer Interaction Course, *IEEE Transactions on Education*, **49**(4), pp. 420–431, 2006.
48. J. Russell, M. S. Andersland, S. van Horne, J. Gikonyo and L. Sloan, Large Lecture Transformation: Improving Student Engagement and Performance through In-Class Practice in an Electrical Circuits Course, *Advances in Engineering Education*, **6**(3), 2018.
49. K. M. Martin and J. M. Gallimore, It's Not Rocket Science: The Flipped Classroom in Space Mechanics, *ASEE Annual Conference & Exposition*, Salt Lake City, UT, June 24–27, 2018.
50. K. M. Martin and J. M. Gallimore, Flipping Space Mechanics: Comparing a Traditional and Less Traditional Classroom, *SciTech 2018*, Kissimmee, FL, January 8–12, 2018.
51. M. L. Loughry, M. W. Ohland and D. D. Moore, Development of a theory-based assessment of team member effectiveness, *Educational and Psychological Measurement*, **67**(3), pp. 505–524, 2007.

Kaela M. Martin is an Associate Professor in the Department of Aerospace Engineering at Embry-Riddle Aeronautical University in Prescott, Arizona. She earned a bachelor's degree in Aerospace Engineering and Mathematics at Iowa State University as well as a master's and PhD in Aeronautical and Astronautical Engineering from Purdue University. Her research interests in engineering education include developing classroom interventions that improve student learning, designing experiences to further the development of students from novices to experts, and creating engaging classroom experiences.

Jonathan M. Gallimore is a Senior Lecturer in the Department of Psychology at the University of Wisconsin-Madison in Madison, Wisconsin. He earned a bachelor's degree in Psychology and master's degree in Human Development and Social Policy from the University of Utah. He earned a PhD in Experimental Psychology from Washington State University. His research interests in education include developing gamified classroom activities, the influence of distributed practice on long-term memory, and creating class tools that reduce student resistance and support student success.