

Article

A Quasi Experiment to Determine the Effectiveness of a “Partially Flipped” versus “Fully Flipped” Undergraduate Class in Genetics and Evolution

Alison E. M. Adams,* Jocelyn Garcia, and Tinna Traustadóttir

Department of Biology, Northern Arizona University, Flagstaff, AZ 86011

Submitted July 23, 2015; Revised February 20, 2016; Accepted February 20, 2016
Monitoring Editor: Erin Dolan

Two sections of Genetics and Evolution were taught by one instructor. One group (the fully flipped section) had the entire class period devoted to active learning (with background material that had to be watched before class), and the other group (the partially flipped section) had just a portion of class time spent on active learning (with the background material presented during class time). The same materials and assessments were used for both sections. Analysis of objective measures revealed that there was no significant difference between the learning outcomes of students in the two sections. There was no main effect of gender, major, or ethnicity on success in the whole cohort or in either section. There appeared to be a significant main effect of class standing, with freshmen performing significantly less well than sophomores, juniors, or seniors (who all performed equally well) in both sections ($p < 0.01$); however, this was a very preliminary observation, as there were very few freshmen in either section. The only predictor of success in the two sections was prior grade point average. An anonymous end-of-semester survey showed no significant difference between the two sections in interest in the subject matter.

INTRODUCTION

Genetics, like other science, technology, engineering, and mathematics (STEM) fields, requires students to learn higher-order skills such as problem solving, and the importance of active learning in these fields is unequivocal (Freeman *et al.*, 2014; Wieman, 2014). In their meta-analysis of studies comparing active learning with the traditional lecture format, Freeman *et al.* (2014) point out there is no longer any question about the relative effectiveness of the active-learning

environment and research should now be focused on determining how best to teach in this way. For example, what types of active-learning exercises are optimal? What is the relationship between amount of time spent in active learning and student success?

In the traditional approach to teaching genetics, students are often exposed to the material and shown how to solve problems during an in-class lecture. The amount of classroom time remaining for students to engage in active learning, through practicing problem solving, is therefore very limited, and students are usually asked to do this for homework after class. Students frequently struggle with problem sets assigned for homework and would often benefit from having the instructor available to assist with questions and discussion during this process. The “flipped” classroom approach seems an ideal way to free up additional class time for active learning through problem solving. In the flipped approach, basic factual information can be provided before class time (e.g., through online lectures), and then classroom time can be used as an active-learning environment (e.g., involving group work on problem sets, case studies, and discussion; Lage *et al.*, 2000; Moravec *et al.*, 2010; Bergmann and

CBE Life Sci Educ June 1, 2016 15:ar11

DOI:10.1187/cbe.15-07-0157

*Address correspondence to: Alison E. M. Adams (Alison.adams@nau.edu).

© 2016 A.E.M. Adams *et al.* CBE—Life Sciences Education © 2016 The American Society for Cell Biology. This article is distributed by The American Society for Cell Biology under license from the author(s). It is available to the public under an Attribution–Non-commercial–Share Alike 3.0 Unported Creative Commons License (<http://creativecommons.org/licenses/by-nc-sa/3.0>).

“ASCB®” and “The American Society for Cell Biology®” are registered trademarks of The American Society for Cell Biology.

Sams, 2012; Strayer, 2012; Tucker, 2012; Gajjar, 2013; Hamdan *et al.*, 2013; Tune *et al.*, 2013; Love *et al.*, 2014; Kim *et al.*, 2014; Jensen *et al.*, 2015).

The flipped classroom has become increasingly popular in the past few years, and there is even an online resource—the Flipped Learning Network—whose mission is “to provide educators with the knowledge, skills and resources to successfully implement Flipped Learning.” The literature supporting the effectiveness of the flipped approach, however, mostly consists of surveys, case studies, and reports, and very few quantitative studies have been described that directly compare the flipped and nonflipped approaches when all other components (instructor, materials, assessments, etc.) are kept constant (Hamdan *et al.*, 2013; Jensen *et al.*, 2015). Thus, it has been difficult to discern the specific impact of flipping the classroom. Recently, Jensen *et al.* (2015) used a quasi-experimental approach to compare the flipped and nonflipped approaches, keeping all components (including the use of active learning) of the two sections constant except the way in which acquisition and application of concepts occurred: in the nonflipped section, instructor-facilitated acquisition of concepts occurred during class time and application occurred after class, whereas in the flipped section, acquisition of concepts occurred before class, and instructor-facilitated application occurred during class time. In that study, which was conducted with a nonmajors course in general education biology, at a private institution in the United States with highly select and culturally homogeneous students, Jensen *et al.* (2015) showed that there was no significant difference in learning and attitudes of students in the flipped versus the nonflipped sections.

In the present study, we varied the amount of time spent in active learning during class time, using a fully flipped approach for one section, and a partially flipped approach for the other section. The specific research goals were 1) to compare performance of the students in the two sections through analysis of exam and final scores; 2) to compare student withdrawal rates from the two sections; and 3) to compare student attitudes toward the subject matter and flipped approach, using an anonymous end-of-semester survey.

METHODS

IRB Approval

The study protocol was approved by the Northern Arizona University Institutional Review Board (project 681178-2), and students were asked to sign IRB-approved consent forms if they wished to participate in the study.

Design of the Fully Flipped and Partially Flipped Sections

The study was conducted with two sections of a sophomore-level class (BIO 240) in genetics and evolution, at a culturally diverse public institution in the United States. As students were expected to already know the fundamentals of Mendelian and population genetics, this was a somewhat intermediate-level class. The two sections were taught by the same instructor; were provided with the same materials, including lecture material and in-class problem sets; and had the same assessments. The main difference was in the way

in which the class was taught. For the fully flipped section, the two 75-min class periods each week were spent entirely in active learning (with the whole class period devoted to working on problem sets), with background information provided as an online lecture that had to be watched before class. For the partially flipped section, just the last 15–25 min of the two 75-min classes were spent in active learning (using the same problem sets), with background material provided during class time.

Several different outcomes of performance and student course-related attitudes were measured. To obtain objective measures of student performance in the two sections, we evaluated students with the same weekly quizzes, the Genetics Concepts Assessment (GCA; Smith *et al.*, 2008), and midterm and final exams. To obtain information about student attitudes toward the fully or partially flipped approach, we asked students in the two sections to complete an anonymous end-of-semester survey.

Both sections were taught in the same room, back-to-back (twice a week at 2:20–3:35 or 4:00–5:15 pm), and there were a total of 21 class periods (excluding exams and reviews for the exams) during the semester for each section. The two sections were designed to be taught identically, except that students in the fully flipped section were given an online lecture before coming to class, while students in the partially flipped section were given the same background information during class time. The primary consequence of this difference was that, for the fully flipped section, the entire class period (75 min) could be used for active learning (through working on problem sets), whereas for the partially flipped section, a minority (~15–25 min) of the classroom time was available for active learning (with the same problem sets). Because the time available for working on the problem sets was greatly reduced in the partially flipped section, students in this section were told to get as far as they could; they were not expected to complete the worksheets.

Background information provided during class or online was identical in content and was delivered in the same style—orally with whiteboard-style notes and diagrams. In the case of the fully flipped section, this material was provided as an online Livescribe PDF with audio and visual components and was not provided during class time, whereas for the partially flipped section, this material was provided in person during class time and was not made available online.

Both sections were given the same assigned pages of the text to read before coming to class and were given a pre-class online quiz to ensure the students had completed the required work. (In the case of the fully flipped section, this quiz had questions based on the online lecture material as well as the text. As these preclass quizzes were different between the two sections, they were not used in the analysis described in this study.)

During class time, students in both sections worked on the problem sets in groups of three to five, receiving help as needed from both the instructor and peer teaching assistants (TAs). Before the end of class, students handed in a group answer sheet with the names of those who contributed (they kept individual answer sheets for their own records), and the instructor went over the answers with the class as a whole. After each class period, an online postclass quiz was posted for both sections, based on the lecture material and material from

the in-class problem sets. The postclass quizzes were identical for both sections and closed before the next class period.

In addition to the same lecture material and weekly in-class problem sets, both sections had the same three midterm and final exams and the same opportunities for extra-credit points. Exam questions were designed to test background knowledge (mostly gained from the lecture material) and problem-solving skills (mostly gained from the in-class problem sets). Student success in the two sections over the course of the semester was compared using exam and total semester scores (excluding the preclass quiz scores, as described above) and the GCA (Smith *et al.*, 2008).

Thus, the way in which the two sections were taught differed only with respect to: 1) presentation of the background material (online vs. face to face); 2) in how class time was used: the fully flipped section had more opportunity to work on the problem sets, whereas the partially flipped section had more opportunity to discuss the material while it was being presented; and 3) the content of the online preclass quizzes (those for the fully flipped section contained questions on both the text and the online lecture material, whereas the preclass quizzes for the partially flipped section contained questions on the text only). An additional difference between the two sections, which could not be controlled, was time of day at which the classes met (2:20 pm for the partially flipped class and 4:00 pm for the fully flipped class), but this difference was minimized by having the two classes back-to-back in the same classroom. The similarities and differences between the two sections are summarized in Table 1.

The Online and In-Class Lectures

LIVESCRIBE (livescribe.com) was used to generate online lectures that contained both an audio and visual component, as described previously (Adams *et al.*, 2015). Online lectures

Table 1. Similarities and differences in design and assessments of the two sections

Element	Partially flipped ^a	Fully flipped ^a
Online lecture	N	Y
In-class lecture	Y	N
Preclass quiz ^b	Y	Y
Postclass quiz	Y	Y
In-class worksheet ^c	Y	Y
Midterm exams	Y	Y
Final exam	Y	Y
GCA (pre- and post-tests)	Y	Y
Anonymous survey	Y	Y
Time of day	2:20 pm	4:00 pm

^aElements used for each section indicated with Y; elements not used indicated with N.

^bFor the partially flipped section, this quiz had questions on pages of the text the students were required to read before class. For the fully flipped section, the preclass quiz had questions on both the online lecture and the required reading in the text.

^cThe same worksheet was given to both sections, but the partially flipped section had just 15–25 min of class time, whereas the fully flipped section had the entire class period (75 min) to work on and discuss the worksheets.

varied from 11 to 42 min, with a mean of 26 min if the lecture was not paused. In-class lectures covered the same material (each online lecture corresponding to one in-class lecture). In the partially flipped section, although the lecture material was the same as for the online lecture, it took longer to deliver (typically about twice as long) because of the time taken to answer questions and for students to take notes and so on, as noted previously (Day and Foley, 2006).

Students in both sections were strongly encouraged to take handwritten notes (and indeed, 96 and 84% of students in the partially and fully flipped sections, respectively, reported doing so; unpublished data), as we had previously noted that handwritten lecture notes improved performance (Adams *et al.*, 2015; see also Chi and Wylie, 2014). Students in both sections were welcome to ask questions during lecture (face to face for the partially flipped section and via email for the fully flipped section), but the vast majority of the questions in both sections occurred during the time spent on the problem sets.

Active Learning with the In-Class Problem Sets

For the fully flipped section, the entire 75-min class period was used for the students to work on the problem sets and subsequent discussion of the solutions. For the partially flipped section, only ~15–25 min of each class period was used for the problem sets and subsequent discussion of the answers.

Assessment of Student Learning

Student learning in both sections was assessed out of a maximum of 500 points, with 1) best 18 of 21 online preclass quizzes, each worth 5 points (90 points total); 2) best 18 of 21 online postclass quizzes, each worth 5 points (90 points total); 3) best two of three midterm exams (worth 200 points total) and a final exam (worth 100 points); 4) participation points, based on group problem sets, with participation points closely reflecting attendance (worth just 10 points total); and 5) the GCA (Smith *et al.*, 2008; worth 10 points if students completed both the pretest and the posttest). Six additional extra-credit points were made available to both sections during the semester for taking practice midterm exams.

Exams and quizzes were all developed by the instructor (A.A.) over the course of several years, and were designed to test course objectives. More important topics had more test questions, and confusing questions were clarified or eliminated over the years of teaching this class.

The GCA was developed and validated by Smith *et al.* (2008). For the present study, it was administered online both as the pretest and also as the posttest (in retrospect, it would have been better to have it be an in-class assessment, as described by Smith *et al.* [2008], to ensure that students attempted to do their best). Students earned 10 points (out of the 500 total points) for completing both the pre- and posttest. Improvement in students' GCA scores was obtained as described by Smith *et al.* (2008) by computing the average number of points improvement/average pre-GCA score.

Survey to Analyze Student Attitudes toward Genetics and Evolution and the Flipped Approach

In the last class period of the semester, students were given an anonymous survey to determine their attitudes toward

and experiences of the course. Students in both sections received the same survey (although some questions were applicable to just one of the two sections), which they completed during class time. Surveys included questions about grade before the final exam, interest in genetics and evolution, and views on flipped versus nonflipped formats. The response rate was 67% in the fully flipped section and 69% in the partially flipped section. The survey was developed by the instructor and was a modified form of one described previously (Adams *et al.*, 2015); it was validated by successively asking the undergraduate researcher (J.G.), four other undergraduates who had not taken the course, and finally, an entire undergraduate class in the humanities to read the survey and note any ambiguities in the wording. Any ambiguities identified by these students were used to revise the survey at each stage.

The Peer TAs

Because of the size of both sections (there were 84 students in the partially flipped section and 78 in the fully flipped section, including the postbaccalaureate and honors students, and excluding students who withdrew from the class; see *Participants* below), it was necessary to have help with answering students' questions as they worked on the in-class problem sets. Three peer TAs were recruited from students who had earned an "A" in BIO 240 in a previous semester, and all three assisted with both sections. Each week, before the first class of the week, all three peer TAs met with the instructor for a TA meeting to go over the worksheet. The peer TAs received independent study credit in pedagogy.

Participants

Class sizes for the two sections were similar. There were 77 students in the partially flipped section, and there were 73 students in the fully flipped section (these numbers do not include the 19 students who withdrew from the course [see *Results*], the one postbaccalaureate student, and the 11 honors students [seven in one section; four in the other] for whom we did not have demographic information). The demographics of the students in the two sections are described below and in Table 2. None of the students knew that there was a difference between the two sections when they registered for the class.

Data Analyses

Differences between the two sections were analyzed with independent *t* test or univariate general linear model when additional factors such as class year and major were analyzed. Post hoc testing was performed when appropriate, using least significant differences. Demographic differences between the sections were analyzed by a chi-square nonparametric test. Pearson correlation analysis was used to examine the relationship between prior grade point average (GPA) and class performance. All analyses were conducted using IBM SPSS statistical software, version 22.0 (SPSS, Chicago, IL). All comparisons were considered significant at $p < 0.05$. Data are presented as means \pm SD.

Table 2. Student demographics in the two sections^a

	Partially flipped section ($n = 77$)	Fully flipped section ($n = 73$)
Class standing		
Freshman	2	4
Sophomore	18	21
Junior	34	25
Senior	23	23
Gender ^b		
Male	26	36
Female	51	37
Major		
Biology	35	39
Biomedical sciences	15	17
Microbiology	8	3
Other (nine different majors)	19	14
Ethnicity		
White	51	54
Hispanic/Latino	20 ^c	9
American Indian/Alaska Native	3	1
Two or more	2	7
Asian	1	1
GPA ^d	3.03 \pm 0.61	2.99 \pm 0.59

^aOne postbaccalaureate and 11 honors students were not included in the analysis because of lack of demographic data on these students.

^bSignificantly higher number of females than males in the partially flipped section ($\chi^2(1, n = 77) = 8.12, p = 0.004$).

^cSignificantly higher number of Hispanic/Latino students in the partially flipped compared with the fully flipped section ($\chi^2(1, n = 29) = 4.17, p = 0.04$).

^dPrior GPA on a four-point scale; not significant.

RESULTS

Student Demographics in the Two Sections

Student demographics (major, student standing, gender, and ethnicity) for the two sections are shown in Table 2. There were no significant differences *between* sections in the number of students of each class standing, gender, major, or ethnicity, except that the partially flipped section had a significantly higher number of Hispanic/Latino students than the fully flipped section ($\chi^2(1, n = 29) = 4.17, p = 0.04$). Not surprisingly, *within* each section, there were significant differences in the number of students of different class standing (partially flipped: $\chi^2(3, n = 77) = 28.11, p < 0.001$; fully flipped: $\chi^2(3, n = 73) = 15.27, p = 0.002$), in different majors (partially flipped: $\chi^2(3, n = 77) = 20.40, p < 0.001$; fully flipped: $\chi^2(3, n = 73) = 37.41, p < 0.001$), and in the number of students within ethnic groups (partially flipped: $\chi^2(4, n = 77) = 118.78, p < 0.001$; fully flipped: $\chi^2(4, n = 73) = 139.67, p < 0.001$). The partially flipped section, but not the fully flipped section, had a significantly higher number of females versus males ($\chi^2(1, n = 77) = 8.12, p = 0.004$). Prior GPA of the students between the two sections did not differ (3.03 \pm 0.61 vs. 2.99 \pm 0.59, partially flipped vs. fully flipped, respectively; Table 2).

Class Performance

Comparison between the Two Sections. The overall performance in the class did not differ between the two sections (Table 3). Thus, there was no difference between the two sections in scores of any of the individual assessments, that is, exam scores, group work, extra-credit points, postquizzes, or improvement in the GCA (Table 3). Nor was there any significant difference between the two sections in various combined scores, that is, total of all four exams, best two midterm exams, or total points (as defined in Table 3). A linear regression analysis entering section, class standing, gender, major, and prior GPA revealed that a significant proportion of variance in total points was explained by

Table 3. Comparison of student scores between sections

Assessment	Partially flipped (<i>n</i> = 77)	Fully flipped (<i>n</i> = 73)
Exam 1 (out of 100 points)	77.7 ± 16.7	77.7 ± 17.2
Exam 2 (out of 100 points)	77.1 ± 15.8	78.3 ± 17.1
Exam 3 (out of 100 points)	73.6 ± 17.3 ^a	69.9 ± 22.3
Final exam (out of 100 points)	67.7 ± 12.2 ^b	67.5 ± 12.9 ^c
In-class group work (out of 10 points)	8.7 ± 1.7	8.6 ± 2.1
Extra-credit points (out of 6 points)	2.6 ± 2.3	2.4 ± 2.4
GCA improvement (percent) ^d	40 ± 50 ^e	33 ± 52 ^f
Best 18 postquizzes (out of 90 points)	58.6 ± 17.8	58.5 ± 18.6
Ex1+Ex2+Ex3+Final Ex (out of 400 points)	299.3 ± 44.1 ^g	299.0 ± 50.5 ^g
Best two of three midterm exams (out of 200 points)	160.9 ± 28.8	162.4 ± 27.8
Total points (percent of 320 points) ^h	76.7 ± 14.4	76.3 ± 14.9

Data are presented as means ± SD. None of the differences between the two sections are significant.

^a*n* = 76. The number is lower because students knew they could drop one exam, so not all students took the third exam.

^b*n* = 75 (partially flipped) and ^c*n* = 70 (fully flipped). These numbers are lower because some students who were failing the class before the final did not take this exam.

^dImprovement in GCA score is a relative score, obtained as the average number of points improvement/average pre-GCA score. No significant difference is seen between groups.

^e*n* = 57 (partially flipped) and ^f*n* = 44 (fully flipped). These values of *n* are lower because not all students completed both the pre- and post-GCA

^g*n* = 74 (partially flipped) and ^h*n* = 68 (fully flipped). These values of *n* are lower because some students did not take exam 3 and/or the final.

^hTotal points (out of 320) were obtained from the best two midterm exams (200 points), the final exam (100 points), in-class group work (10 points), extra-credit points for doing three practice exams (6 points not counted in the denominator), and for taking the pre- and post-GCA (10 points). Pre- and postclass quizzes were not included in the total points, because the content of the prequizzes was different between the two sections and because students were given the option of not including their quiz scores in their final grade calculation.

Table 4. Regression analysis of the effect of section, prior GPA, gender, major and class standing on total points

Model value	Unstandardized coefficients		Standardized coefficients		
	β	SE	β	<i>t</i>	<i>p</i> Value
(Constant)		93.665	14.70	6.371	<0.001
Section	0.786	4.06	0.011	0.194	0.847
Prior GPA	46.211	3.59	0.735	12.877	<0.001
Gender	0.199	4.18	0.003	0.048	0.962
Major	0.145	2.20	0.004	0.066	0.948
Class standing	4.609	2.43	0.108	1.895	0.060

the model ($R^2 = 0.589$, $F(5142) = 39.33$, $p < 0.001$). However, prior GPA was the only significant predictor of class performance (see Table 4). The same analyses with the change in GCA score as the dependent variable did not explain a significant proportion of the variance ($R^2 = 0.071$, $F(4100) = 1.83$, n.s.). However, the change in GCA score was significantly correlated with total points, in both sections (partially flipped: $r = 0.32$, $p < 0.05$, $n = 57$; fully flipped: $r = 0.31$, $p < 0.05$, $n = 44$).

Very similar numbers of students in each section chose not to include their online quiz scores for their final grades. In the case of the partially flipped section, 23% did better (mostly by a single letter grade) if their quiz scores were ignored, as did 28% in the fully flipped section. The vast majority of students in each section earned the same letter grade with and without their quiz scores, and just 7–8% in the two sections did better with their quiz scores than without.

We used a power analysis to determine how many students would be needed for the difference that we saw between the two sections to be significant. We used two different outcomes—the final exam scores and total points. To achieve a power of 0.8, the power analysis indicated we would need ~3000 students for the observed difference in the final exam scores to be significant and more than 32,000 students for the observed difference in total points to be significant.

Effect of Class Standing. Overall (i.e., combining students in both sections), freshmen had significantly lower total points than students of any other class standing ($p < 0.01$). This observation, however, is based on just six freshmen in the two sections, and any conclusion concerning it is therefore highly preliminary. There were no significant differences between any of the other class years (sophomores, juniors, or seniors). There was no main effect of section or an interaction between class year and section. In other words, when students of a particular class standing were compared between the two sections, there was no difference in class performance. When the effects of class standing were analyzed separately for each section, the freshman effect remained significant for the fully flipped section (just four students; $p < 0.05$) but was not significant for the partially flipped section (Table 2).

Effect of Gender, Major, or Ethnicity. There were no significant differences in total points when analyzed by gender, major, or ethnicity, whether the sections were analyzed together or separately.

Withdrawal Rates of Students from the Two Sections

In both sections, most students who dropped the course did so before the first exam: in the partially flipped section, six dropped in the first week of class, and three dropped just before or after the second exam. In the fully flipped section, seven dropped within the first 10 d of class, one dropped after the first exam, and two dropped just before or after the second exam. These students were not included in the analysis shown in Table 3. Comparison of withdrawal rates from the two sections indicated no significant difference in the number of students withdrawing from the two sections (nine from the partially flipped section, 10 from the fully flipped section), with no significant difference in gender, major, ethnicity, or class standing of students who dropped the class between the two sections. There were, however, disproportionate numbers of nonwhite students who dropped from both sections: 5/9 in the partially flipped section and 6/10 in the fully flipped section. This is consistent with data that show that a disproportionately large number of STEM-interested underrepresented minority students fail to complete their studies with a STEM degree (Freeman *et al.*, 2014).

Mastery of Students in the Two Sections of Factual Knowledge versus Application Skills

The effectiveness of active learning likely depends on how cognitively demanding the assessments are (i.e., there is likely to be little effect of active learning on exams that primarily assess factual recall). To address this issue, we categorized the exam questions we used (for all four exams) by Bloom's level; the percentages of each level of question for the different exams are shown in Table 5. In each exam, 50–73% of the questions were level 3 (“apply”) and only 17–24% were level 1 (“remember”).

A further consideration related to the difference in how class was spent by the two sections. In particular, students in the partially flipped section had a greater opportunity to ask questions during the lecture, when background material was being presented, than did students in the fully flipped section. It was therefore possible that the partially flipped section had a greater level of mastery of background material than did the fully flipped section. Conversely, students in the fully flipped classroom had more opportunity for application of concepts than did students in the partially flipped section. It was therefore possible the fully flipped

Table 5. Type of exam question (%) categorized using Bloom's taxonomy^a

	Bloom's level	Exam 1	Exam 2	Exam 3	Final
1	Remember	20	20	17	24
2	Understand	23	7	33	14
3	Apply	57	73	50	62
4	Analyze	0	0	0	0
5	Evaluate	0	0	0	0
6	Create	0	0	0	0

^aExam questions were categorized using definitions of Bloom's level listed in www.colorado.edu/sei/documents/Workshops/Handouts/Blooms_Taxonomy.pdf.

Table 6. Percentage incorrect answers per student per question for final exam questions categorized as more factual (background knowledge) versus more applied (requiring calculation, interpretation, etc.)

Type of question ^a	Fully flipped	Partially flipped	<i>p</i> Value ^b
Factual questions	30%	31%	0.85
Applied questions	33%	31%	0.44

^aThere were 13 factual and 37 applied questions.

^b*p* Values obtained using a *t* test comparing the two sections for number of incorrect answers per student per question for factual vs. applied questions.

section had a greater level of mastery of application skills than did the partially flipped section. Overall, the two sections might therefore do equally well, but for different reasons. For determination of whether there was any difference between the two sections in mastery of background material versus application skills, final exam questions were categorized as “more factual” versus “more applied.” The average percentage of incorrect answers of students for these two types of questions were then compared between sections. As shown in Table 6, there was no significant difference between the two sections for the two types of questions. Thus, the students in the two sections do not consistently appear to master different aspects of the course. In addition, regression analyses show that prior GPA was the only significant predictor of mastery of either type of question ($R^2 = 0.218$, $F(6, 96) = 6.23$, $p < 0.001$ for “more factual”; $R^2 = 0.427$, $F(6, 1755) = 16.64$, $p < 0.001$ for “more applied”). Section, gender, class year, and major did not account for a significant portion of the variance of either dependent variable.

Use of an Anonymous End-of-Semester Survey to Reveal Attitudes of Students in Both Sections toward Genetics and Evolution and the Flipped Approach

Attitudes toward Genetics and Evolution. In the anonymous survey, students in both sections were given a collection of five statements, to which they were asked to give a score of 1–5 (1 = strongly disagree; 5 = strongly agree) designed to address their attitudes toward Genetics and Evolution. Students in both sections were very positive about the course (Figure 1). A nonparametric Mann-Whitney *U*-test indicated that the distribution of level of agreement from 1 to 5 was not significantly different between the two sections for any of the six statements.

Preference for Partially Flipped versus Fully Flipped Approaches. When students in both sections were asked in the anonymous survey whether they preferred the flipped classroom (with online lecture instead of in-class lecture) or the nonflipped in-class lecture, a clear difference was seen: 89% of those who responded in the partially flipped section (who had not experienced the fully flipped approach for this class) said they preferred the nonflipped approach, whereas just 42% of those in the fully flipped section (who had experience with the flipped approach for this class) said that

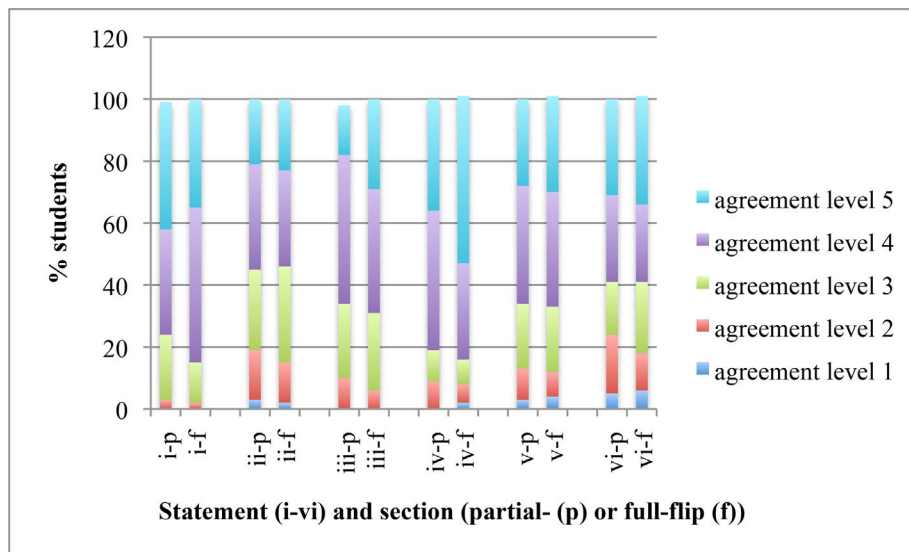


Figure 1. Attitudes of students in the fully and partially flipped sections. Agreement levels of 1–5 indicate student responses to statements i–vi, where 1 = strongly disagree, 3 = neutral, and 5 = strongly agree. The wording of the statements was: (i) “The material in this class is very interesting to me”; (ii) “Genetics and Evolution is one of my favorite subjects”; (iii) “The material presented in this class is relevant to my life”; (iv) “The material presented in this class is relevant to my major/career”; (v) “I am willing to spend additional time reading about genetics topics”; (vi) “If I had a chance, I would like to do a research project in genetics.”

they preferred the nonflipped approach. When students in the partially flipped section were asked to explain their preference, of the students who preferred the in-class lecture format, about one-third said it was more personal, engaging, and easier to ask questions; another third focused on the discipline provided by an in-class format, that is, that it was easier to come to class and to focus on the material in the classroom; and about one-third expressed miscellaneous reasons (e.g., they were paying to interact with the professor, they do better in traditional lecture formats, and they generally dislike online materials). When students in the fully flipped section were asked to explain their preference, of those who preferred the flipped approach, the most frequent reason (about half the comments) was that they liked using class time to apply what they had learned before class, and also to ask questions and get help rather than struggle with problems on their own. About one-fourth of the comments from these students also indicated that students liked being able to replay the lecture and go through it at their own pace. Of the students in the fully flipped section who would have preferred the in-class lecture format, the most frequent response (about half the comments) was that it would have been less time-consuming and easier to focus on material if it was presented in the classroom. Other comments included difficulty of getting the online lecture to work, finding in-class lectures more personal, and so on.

DISCUSSION

In this study, we compared the effect of using the fully flipped versus partially flipped approach in two sections of a course in genetics and evolution (BIO 240, a sophomore-level 16-wk course at Northern Arizona University, which is a large, culturally diverse, public institution in the United States). In both sections, which were taught in a single semester by the same instructor, identical lecture content, in-class assessments, postclass quizzes, and validated pre- and post-GCA (Smith *et al.* 2008) were used. Both sections used

active learning during class time, the main difference being the amount of time devoted to this: for the fully flipped section, all class time was available for active learning and discussion, as background material was presented before class time, whereas in the partially flipped section, there was much less time for active learning and discussion during class time, as background information was also presented during this time.

Overall Findings

Several key findings were made. First, there were no significant differences in objective measures of student success in the two sections (Table 3). Second, analysis of student success in the two sections revealed that students of non freshman standing did equally well both within and between sections. (The freshmen did significantly less well than did sophomores, juniors, and seniors in the fully flipped but not the partially flipped section, but this was based on just six freshmen [see *Limitations of the Study* below]; there was, however, no significant difference in freshman success between the two sections.) Third, prior GPA was the only significant predictor of success in the class. Fourth, rates and demographics of students withdrawing from the two sections were not significantly different, though a disproportionate number of nonwhite students withdrew from both sections. Fifth, analysis of the responses to the anonymous survey revealed positive attitudes toward the course in both sections and no significant differences between the two sections in students' attitudes toward the course and the topic of genetics and evolution (Figure 1). There was, however, a difference in student attitudes toward the flipped classroom format—89% of students in the partially flipped section said they preferred the nonflipped lecture format, whereas only 42% of students in the fully flipped section said they preferred the nonflipped approach. The latter finding suggests that, once students are exposed to more active forms of instruction, they may be less resistant than might be expected (Seidel and Tanner, 2013).

Our study shows that, when active-learning methods such as the ones described in this study are used, the fully flipped and partially flipped approaches have similar outcomes in a sophomore-level course in genetics and evolution at a large, culturally diverse public university in the United States.

Time Spent in Active Learning

In their meta-analysis of studies comparing active learning with the traditional lecture format, Freeman *et al.* (2014) suggest research now be focused on determining how best to teach with active learning. For example, what is the relationship between amount of time spent in active learning and student success? In the present study, we found that the use of active learning (through the use of problem sets) in the classroom was equally effective whether the entire class period or just a fraction of it was devoted to active learning. This was a somewhat surprising result, as we had expected that a higher proportion of class time spent in active learning would be beneficial and that rushing students through problem sets at the end of a lecture period would be less effective than allowing them to explore the problem sets at a more leisurely pace.

There are several possible explanations for this finding. First, it is possible that additional or other approaches to active learning would have been more effective than those used in this study; future research should be directed at determining what types of activities are most effective in eliciting active learning (Andrews *et al.*, 2011; Freeman *et al.*, 2014; Eddy *et al.*, 2015). Second, it is possible that, if we had measured differences in learning at Bloom's levels 4–6 on the exams, we would have seen a difference between the two sections, as active learning is likely to be most important for higher levels of learning. The GCA, however, contains questions that address higher Bloom's levels, and we saw no significant difference between the sections in improvement in the GCA (either overall or when questions were categorized by Bloom's level; unpublished data). Third, it is possible that students in the partially flipped section were more actively engaged with the background material than were students in the fully flipped section (e.g., the opportunity to listen to and respond to one another's questions during class, when this material was presented, may have increased engagement in the partially flipped class). This might have resulted in little overall difference between the two sections in time spent in active learning. Fourth, it is possible that, as the largest effect of active learning on student performance appears to be in smaller classes (Freeman *et al.*, 2014), we would have seen a relationship between time spent in active learning with problem sets and student success in a smaller class.

Limitations of the Study

In this study, we used a quasi-experimental approach to determine the impact of spending an entire, versus just a small proportion, of the class period on active learning. The study was designed to keep as many other elements of the class as possible the same between the two sections. There were, however, differences that were not possible to eliminate, and other limitations that arose due to the way in which the study was conducted. First, it was not possible to assign

students randomly to one of the two sections, as the sections were not concurrent (a physical impossibility, because the same instructor taught both sections); the design of the study is therefore "quasi-experimental" rather than "experimental." Thus, it is possible that students self-selected into one of the two sections in a way we were unable to detect. The demographics of the two sections, however, indicate no significant difference in prior GPA, gender, ethnicity, major, or class standing.

Second, although the data indicate a significant overall effect of class standing on performance, with freshmen performing significantly less well than students of higher standing, this conclusion is very limited by the number of freshmen (just two in the partially flipped section and four in the fully flipped section); it should therefore be considered no more than a preliminary finding that warrants further research to determine whether the result bears out with a larger sample.

Third, as described above, it is possible that the lack of observed difference between student outcomes in the two sections was because the data-collection tools (e.g., the exams) were not the correct ones for capturing this difference. Although the GCA measures learning at both higher and lower levels of Bloom's taxonomy, our finding that there was no significant difference between the two sections with this tool may be limited by the way in which the GCA was administered (online, instead of during class time). It is possible that it was not taken seriously by the students because of this, and in retrospect, it would have been better to embed the posttest questions in the final exam, as was done previously (Smith *et al.*, 2008). The very high SD that we see for the GCA improvement and the fact that fewer students completed both the pre- and post-GCA than completed the course (Table 3) may be indicative of this limitation. Future research should aim to determine whether students shift in their abilities to think at a higher level (e.g., analysis, evaluation).

Fourth, it is possible that there are differences in long-term appreciation of genetics and/or understanding or ability to apply the concepts from this class. We are currently conducting a longitudinal study to examine this possibility, by having students in courses for which BIO 240 is a prerequisite take the GCA before the beginning of those courses.

Fifth, the anonymous end-of-semester survey indicates that students had similar attitudes toward the subject matter. The results of this survey, however, are limited in their value, as they are based on just 67–69% of students responding to the survey. In future studies, we would provide a small incentive (e.g., a bonus point) for completing the survey to increase participation.

ACKNOWLEDGMENTS

We thank the BIO 240 students for their participation in this study, members of the NAU Scholarship of Teaching and Learning Community and the e-Learning Center for their support, Drs. John Pollard (University of Arizona) and Stephanie Winters and Yvonne Luna (NAU) for discussion and help with development of the survey, and Teresa Del Vecchio for help with obtaining data on student demographics. We also thank the editor and reviewers for their thoughtful and helpful comments.

REFERENCES

- Adams AEM, Randall S, Traustadóttir T (2015). A tale of two sections: an experiment to compare the effectiveness of a hybrid versus traditional lecture format in Introductory Microbiology. *CBE Life Sci Educ* 14, ar6.
- Andrews TM, Leonard MJ, Colgrove CA, Kalinowski ST (2011). Active learning not associated with student learning in a random sample of college biology courses. *CBE Life Sci Educ* 10, 394–405.
- Bergmann J, Sams A (2012). *Flip Your Classroom: Reach Every Student in Every Class Every Day*, Eugene, OR: International Society for Technology in Education.
- Chi MTH, Wylie R (2014). The ICAP framework: linking cognitive engagement to active learning outcomes. *Educ Psychol* 49, 219–243.
- Day JA, Foley DJ (2006). Evaluating a Web lecture intervention in a human-computer interaction course. *IEEE Trans Educ* 49, 420–431.
- Eddy SL, Converse M, Wenderoth MP (2015). PORTAAL: a classroom observation tool assessing evidence-based teaching practices for active learning in large science, technology, engineering, and mathematics classes. *CBE Life Sci Educ* 14, ar23.
- Freeman S, Eddy SL, McDonough M, Smith MK, Okoroafor N, Jordt H, Wenderoth PM (2014). Active learning increases student performance in science, engineering, and mathematics. *Proc Natl Acad Sci USA* 111, 8410–8415.
- Gajjar NB (2013). The role of technology in 21st century education. *Int J Res Educ* 2, 23–25.
- Hamdan N, McKnight PE, McKnight K, Arfstrom MK (2013). A Review of Flipped Learning. www.flippedlearning.org/cms/lib07/VA01923112/Centricity/Domain/41/LitReview_FlippedLearning.pdf?utm_source=hootsuite&utm_campaign=hootsuite (accessed 13 April 2016).
- Jensen JL, Kummer TA, Godoy PDdM (2015). Improvements from a flipped classroom may simply be the fruits of active learning. *CBE Life Sci Educ* 14, ar5.
- Kim MK, Kim SM, Khera O, Getman J (2014). The experience of three flipped classrooms in an urban university: an exploration of design principles. *Int High Educ* 22, 37–50.
- Lage MJ, Platt GJ, Treglia M (2000). Inverting the classroom: a gateway to creating an inclusive learning environment. *J Econ Educ* 31, 30–43.
- Love B, Hodge A, Grandgenett N, Swift AW (2014). Student learning and perceptions in a flipped linear algebra course. *Int J Math Educ Sci Technol* 45, 317–324.
- Moravec M, Williams A, Aguilar-Roca N, O'Dowd DK (2010). Learn before lecture: a strategy that improves learning outcomes in a large introductory biology class. *CBE Life Sci Educ* 9, 473–481.
- Seidel SB, Tanner DK (2013). “What if students revolt?”—considering student resistance: origins, options, and opportunities for investigation. *CBE Life Sci Educ* 12, 586–595.
- Smith MK, Wood WB, Knight KJ (2008). The Genetics Concepts Assessment: a new concept inventory for gauging student understanding of genetics. *CBE Life Sci Educ* 7, 422–430.
- Strayer JF (2012). How learning in an inverted classroom influences cooperation, innovation and task orientation. *Learn Environ Res* 15, 171–193.
- Tucker B (2012). The flipped classroom. *Educ Next* 12, 82–83.
- Tune JD, Sturek M, Basile PD (2013). Flipped classroom model improves graduate student performance in cardiovascular, respiratory and renal physiology. *Adv Physiol Educ* 37, 316–320.
- Wieman C (2014). Large-scale comparison of science teaching methods sends clear message. *Proc Natl Acad Sci USA* 111, 8319–8320.