Adv Physiol Educ 44: 131–137, 2020;
doi:10.1152/advan.00143.2019.

HOW WE TEACH | *Classroom and Laboratory Research Projects*

The refinement of flipped teaching implementation to include retrieval practice

Chaya Gopalan,[1,2] Andrea Fentem,[3] and Anna L. Rever[2]

[1]*Departments of Applied Health, Southern Illinois University, Edwardsville, Illinois;* [2]*Primary Care and Health Systems, Southern Illinois University, Edwardsville, Illinois; and* [3]*Department of Psychiatry, Washington University in St. Louis, St. Louis, Missouri*

Submitted 18 October 2019; accepted in final form 11 February 2020

Gopalan C, Fentem A, Rever AL. The refinement of flipped teaching implementation to include retrieval practice. *Adv Physiol Educ* 44: 131–137, 2020; doi:10.1152/advan.00143.2019.—There has been growing evidence that flipped teaching (FT) can increase student engagement. Traditional lecture-based teaching (TT) method was compared with FT and FT combined with retrieval practice (FTR) in a 400-level Exercise Physiology course over eight semesters. In the FT format, lecture content was assigned for students to prepare before class along with an online quiz. During class, the assigned content and quiz questions were reviewed, and a team-based learning (TBL) activity was conducted. Students found FT implementation three times a week (FT3) to be overwhelming, which led to reconfiguration of the FT design to minimize the quiz and TBL sessions to one per week. Subsequently, FT was combined with retrieval exercises (FTR), which involved recalling information, thus promoting retention. The students in the FTR format were given weekly quizzes in class, where no notes were allowed, which affected their quiz grade negatively compared with FT ($P < 0.0001$). Again, no resources were permitted during FTR's TBL sessions. When exam scores were compared with TT, student performance was significantly greater ($P < 0.001$) with the FT and FTR methods, suggesting these methods are superior to TT. While both male and female students benefited from FT and FTR methods compared with TT ($P = 0.0008$), male students benefited the most (($P = 0.0001$). Similarly, when the exam scores were organized into upper and lower halves, both groups benefited from FT and FTR ($P < 0.0001$) approaches. In conclusion, both FT and FTR methods benefit students more compared with TT, and male students are impacted the most.

flipped classroom; male and female students; retrieval practice; upper and lower half of class

INTRODUCTION

Until recently, the traditional lecture format (TT) has been the most common method of teaching in classrooms of higher education. Many educators continue to consider lecturing to be the best approach to teach students. Contradictory evidence, however, indicates that listening to a lecture is not an effective way to promote deep and lasting student learning (32). The average human attention span is no more than 20 min (3), and recall of information drops drastically after 20 min (37). Additionally, teaching in the form of lecture alone does not meet the needs of all learners and is not suited for teaching higher-order skills, such as critical thinking, synthesis, application, and analysis (4a, 34).

Evidence suggests that active-learning strategies promote student engagement and improve knowledge retention (4, 8, 13, 38). This has motivated many in higher education to change from teacher-centered to student-centered instruction. Learner-centered instruction puts more accountability on the students for their own learning, with the understanding that the teacher will maximize opportunities for students to learn (4, 16). A student-centered classroom, rich in collaborative and active learning, contributes to student success (14, 27).

Flipped teaching (FT) is an active-learning educational format that shifts lecture out of class, thus freeing up class time for student-centered learning (1). In its simplest terms, the lecture is shifted to an individual space via instructor-guided study materials. Problem solving and practice occurs in class under the guidance of the instructor in a group setting. This teaching strategy has gained immense attention lately, as it not only encourages active participation of students but also introduces access to help and opportunities to work with peers. While FT is shown to increase student preparedness, it also paves the way to use class time to engage students in higher levels of Bloom's taxonomy, such as application, analysis, synthesis, and evaluation (19, 20, 30). Team-based learning (TBL) combined with FT allows students to develop interdependence, accountability, autonomy, and skills in communication and collaboration (5, 11).

Although no two flipped classrooms look the same, they all share common characteristics (6, 31). This teaching method encompasses differentiated learning, where preclass assessments are meant to be of lower-order learning as per Bloom's taxonomy, while in-class assessments and discussions are focused on higher-order learning (12, 31). By allowing students to learn at their own pace, FT prevents cognitive overload of new information (22). Moreover, with FT, there is an intentional partial transfer of information delivery outside of the classroom to maximize face-to-face interaction in the classroom (18). A FT format also ensures unlimited access to class content, and students can rewatch lecture videos or utilize guided readings as needed. Additional class time provides opportunities to expand on higher-order thinking skills, collaboration, and enrichment. Thus FT is designed around engaging students deeply with content and providing immediate feedback following formative assessments (15, 31).

Address for reprint requests and other correspondence: C. Gopalan, Dept. of Applied Health, Primary Care & Health Systems, Southern Illinois University Edwardsville, Box 1126, Edwardsville, IL 62026 (e-mail: cgopala@siue.edu).

One crucial part of the FT model is in the careful design of class activities not only to engage students in learning, but also to apply knowledge and think critically (16, 27). It is not effective to "flip" a classroom by providing lecture videos and designating the face-to-face class time as only group office hours (29). Herreid and Schiller (17) reported student resistance to the use of assigned resources. Likely causes for such resistance could be lack of time, content that is too difficult for students to comprehend on their own, topics considered boring, or lack of reward (1). Based on a 5-yr study of FT, Senske (33) offers key lessons for building effective FT classes. These include the importance and accessibility of online resources, faculty's ability to respond when FT strategies fail, and the identification of instructors' efforts and misconceptions.

The use of memory retrieval is yet another important strategy that is argued to be highly beneficial in retaining information, and hence its application appears to be one of the most effective methods of long-term learning (21, 23, 24). Essentially, retrieval of information and learning is thought to occur during episodes of studying, whereas retrieval of information on testing serves merely to assess what was learned. The use of quizzes and exams to engage and enhance retrieval processes has been widely established as an effective strategy for facilitating learning (21, 23). Knowledge gained through retrieval practice is maximized when the practice is appropriately spaced out (21). Thus the integration of retrieval exercise into educational practices has the potential to boost performance in classrooms.

There is a great deal of evidence to suggest that the preferred learning styles vary among male and female students. Gross et. al. (12) examined sex differences in their FT and reported that female students benefited greatly compared with male students. Yan et. al. (36) showed that the FT strategy increased male student's motivation, whereas female students were more self-efficient in a traditional classroom setting compared with male students. There are no studies comparing the effect of retrieval practice combined with FT between male and female students thus far.

The learning styles could vary between high- and low-achieving students as well. Evidence suggests that high-performing students find online teaching as effective as traditional face-to-face teaching, whereas it is less effective for lower-performing students (28). Thus our study was aimed at identifying a teaching approach that is most effective for student learning, where we compared TT with that of two student-centered FT strategies, one with retrieval practice (FTR) and one without (FT). We predicted that students in both FT and FTR would demonstrate greater performance over students in the TT method. Whether the new teaching formats helped male students versus female students and the upper half of the class versus the lower half of the class were additional questions that were addressed in this study. In addition, qualitative data were collected to measure student perceptions of the TT, FT, and FTR teaching methods.

METHODS

This study was conducted over eight semesters using one or two sections of students per class in a 400-level course, Biology of Cardiovascular and Metabolic Diseases (KIN412), a required, undergraduate level course for Exercise Science majors. This study occurred during a time when the university's academic profile remained the same, that is, the instructor, the academic performance of the students, class meeting days, times, and duration, as well as their acceptance criteria. All of the classes were taught by the same instructor and were held during a regular semester of 16 wk, during which the class met for 50 min on Mondays, Wednesdays, and Fridays. The sample included 247 students, 114 male and 133 female students, with an age range of 20–23 yr, except for 1–2% of nontraditionally aged students. The class size ranged from 18 to 35 students.

Blackboard, the course management system, was utilized to post course content as well as online assessments. Table 1 summarizes the teaching approach used for this study in a chronological order. This study was approved by the university's Institutional Review Board (IRB no. 15–1016–2C). The required textbook, *Biology of Cardiovascular and Metabolic Diseases*, an in-house digital textbook, was developed explicitly for this course by the same instructor.

Lecture-based teaching or traditional classroom course design. The first semester of the study was designed to utilize lecture format for the entire semester. Students received reading assignments and PowerPoint slides at least 2 days before each class meeting. The lectures were given in the traditional podium style, but there were plenty of opportunities for students to ask questions. Thus this approach was considered to be an interactive lecture method. The students in this format, however, typically asked only one or two questions per class session. These students had one TBL session before each major exam, and there were four exams in total. Groups of four to five students were formed for the TBL session, as described by Gopalan and colleagues (9–11). Grade point average, sex, and ethnicity were some of the details of the students used to ensure diversity as well as a balance of academic ability in each group. The groups were created in the first week of the semester, and these groups

Table 1. *Study design of the teaching methods tested*

Preclass	In-class
Traditional teaching (TT): exams 1–3 of semester 1, exams 1–2 of semesters 2 and 3	
• Guided readings • PowerPoint slides	• Lecture • Random quizzes over previously covered content • One TBL activity before each exam
Flipped teaching, three sessions per week (FT3): exam 3 of semester 2	
• Guided readings • Lecture videos • PowerPoint slides • Individual quiz • Study guide	• Every session (3 sessions per week): review of lecture and quiz, TBL activity, and immediate feedback
Flipped teaching activities spread over 1 wk (FT): exam 3 of semester 3, exams 1–3 of semesters 4 and 5	
• Guided readings • Lecture videos • PowerPoint slides • Individual quiz • Study guide	• First session: question-and-answer session involving review of difficult topics • Online quiz before second session. Students may access resources • Second session: review of quiz questions, TBL activity. Students may access resources • Third session: review of TBL problem sets and study guide
Flipped teaching activities spread over 1 wk, combined with retrieval practice (FTR): exams 1–3 of semesters 6–8	
• Guided readings • Lecture videos • PowerPoint slides • Individual quiz • Study guide	• First session: question and answer session involving review of difficult topics. Individual quiz at the end of the first session without access to resources • Second session: review of quiz questions, TBL activity without access to resources • Third session: review of TBL problem sets and study guide

TBL, team-based learning.

remained permanent throughout the semester (9–11). In addition to lecture and TBL sessions, quizzes were given periodically over previously covered content, but not as frequently as in the other teaching formats tested in this study (Table 1).

Flipped classroom course design (three in-class sessions per week). During the second semester, a partial FT method was used where the first half of the semester used the TT approach as described above and the second half of the semester utilized the FT. It was divided in such a way that the first two exams were in the TT style, whereas the rest of the content was taught in the FT method. This particular FT design involved students experiencing FT during each of the three weekly class sessions (FT3). The resources that were made available for the homework portion of the FT included reading assignments, the PowerPoint slides, and instructor-recorded lecture videos (Table 1). The lecture videos were 20- to 40-min long and were prepared using Zoom software. Students were expected to complete an individual formative assessment in the form of a short, online quiz that consisted of five questions at the factual level of Bloom's taxonomy (26). During each class period, the instructor briefly reviewed key concepts of the lecture and addressed any questions during class. This was followed by a TBL session and instant feedback. The group work consisted of higher-order questions at comprehension, application, evaluation, analysis, and synthesis levels of Bloom's taxonomy. Although student participation and attendance showed significant increase in test scores compared with the TT format, students reported in their end-of-semester course evaluations that the FT3 style was overwhelming and time consuming. Moreover, the 50-min class period was too short for the review of lecture, review of the quiz that students had completed before their class period, a TBL activity, and its immediate feedback. Three quizzes and three problem sets per week and their grading was extremely busy work for the instructor as well. Therefore, this format was discontinued after one semester of testing.

Modified flipped classroom course design. The third semester of this study used the partial FT method once again, but here the FT method was revised. The first half of the semester, up to the completion of the second exam, was taught in the TT method, as described before. Based on the student comments on the FT3 teaching method, as well as from the instructor's experience with FT3, the course was redesigned so that the weekly content was consolidated into one package instead of three separate parts and was shared with students at least 5 days before their class meeting. For example, in the FT3 format, while teaching cardiac physiology, electrical activity within the heart would be covered as one lecture, cardiac cycle as another lecture, and the regulation of cardiac output as the third lecture for a total of three class meetings. In the revised FT format, all three of the lectures were combined into one and released at the same time as weekly content for the students to prepare. The week started off with the expectation that students had a chance to study the material that was made available during the previous week. The first session was dedicated to questions and answers and a review of difficult concepts. Students completed an online quiz before attending their second session. Thus, in this revised approach, students were able to learn the concepts on their own first, and the instructor clarified topics with which students needed help before their individual assessment. The second session started with the review of their online quiz and a TBL activity, as well as its review. In both their online quiz and their TBL session, students could access information, such as their class notes, slides, textbook, and online resources. The third session was utilized to revisit their weekly quiz and TBL activity as well as to continue the review of any difficult concepts (Table 1). Thus the content that was typically covered in three separate sessions in the FT3 format was combined into one without compromising content coverage, but the number of individual and group assessments was spaced from three to one per week. During the fourth and fifth semesters, this restructured FT format was spread throughout the entire semester, thus bypassing TT completely.

Flipped classroom course design with retrieval practice. From the sixth through eighth semesters, the fully flipped course design was modified to include retrieval exercise (FTR). In the FTR method, the course design was similar to FT, where the content was assigned as a weekly package several days before their class meeting. Students were enticed through a weekly quiz to come prepared for class. In this teaching method, the first session started off with questions and answers, as well as a review of difficult concepts as in the previous semesters. The students were given a quiz toward the end of the first session, and the quiz was reviewed soon after. Most importantly, students were not allowed to use any of their resources while taking their weekly quiz. During the second session, a TBL activity was conducted. Similar to the quiz, the TBL activity was completed without the use of any resources (Table 1). The third session was used to review problem sets used in the TBL session, as well as their practice questions. One important advantage with FT and FTR methods was that, if additional time was needed to review problem sets that were used as part of the TBL session, this was covered in *session 3* instead of rushing to finish it during *session 2*.

Peer evaluations. To promote every student participation in the TBL groups, two peer evaluations were carried out in the FT classes, one in the middle of the semester and one at the end of the semester. Here, students evaluated their team members with a score as well as an explanation of why a student received a higher or lower score.

Data collection. There were four exams in total, where the last exam included not only the new material that was introduced since the end of the third exam, but also all of the content that was covered throughout the semester (comprehensive exam). The fourth exam (final exam) was scheduled during finals week. All other exams were scheduled on Fridays, as it best fit the FT design. Exams were administered in a proctored computer laboratory, as they were created directly on BlackBoard. Students were allowed to review their exam during office hours or by one-on-one appointments. Exam questions were protected by not being released to the students to study from for their comprehensive final exam.

The first three exams typically covered 3–4 wk of course content. Each of the first three exams consisted of 50 questions and represented 100 points; the final exam included 100 questions for 100 points. Most items were multiple-choice questions at the knowledge, comprehension, and application level; there were very few true/false, matching, or essay questions. Eighty percent of the test questions were conserved across all three teaching approaches. Only the first three exam scores were used to study the effect of different modes of teaching, as well as student surveys. The comprehensive final exam was not included in our data analysis, considering the influence of the students' other exams during finals week. Yet another reason for not using the final exam was because our students typically predetermine the scores they would need to obtain a certain grade in the class, and such a practice may affect the study.

A comparison of male and female student grades was carried out to learn if one group benefited more from the structured FT and FTR approaches than the TT method. We also examined if FT would impact the performance of students in the lower half of the class differently from that of the students in the upper half of the class. This was achieved by determining the median of the three exam averages and then separating the class into the upper 50th and the lower 50th percentiles.

The exam scores from each teaching method were pooled between semesters and compared against other teaching methods. Weekly quiz and TBL activities served as formative assessments. Each of their scores added up to one exam grade. A comparison of quiz scores that were pooled from FT and FTR approaches was made to learn the impact of completing assignments without resources. An anonymous online survey was given during the middle of every semester to receive student feedback on the teaching format. The survey was intended to identify the teaching strategies students perceived as most

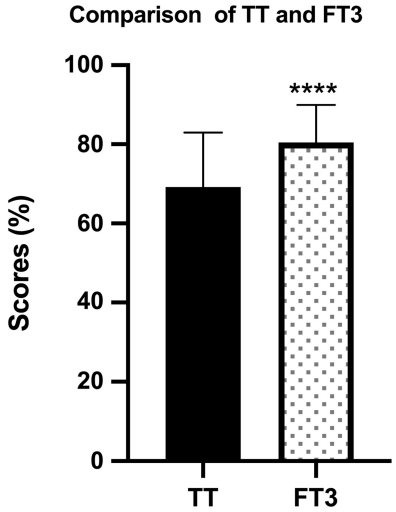

Fig. 1. Exam scores from the traditional teaching (TT) method were compared with the scores from flipped teaching three times per week (FT3). Results are as follows: FT3 (mean = 80.43, SD = 9.559, $n = 53$) and TT (mean = 69.24, SD = 13.74, $n = 220$). n, No. of students. ****$P < 0.0001$, Student's t test.

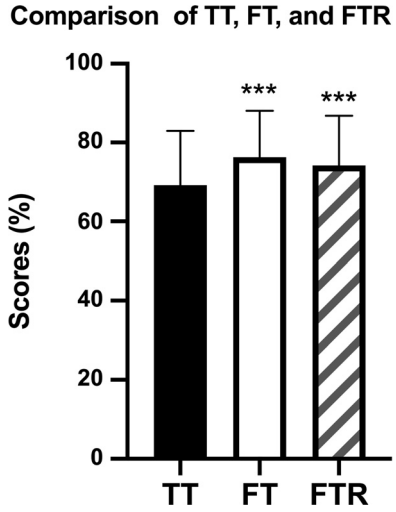

Fig. 2. Comparison of student performance using exam scores between traditional teaching (TT), flipped teaching (FT), and flipped teaching with retrieval practice (FTR) methods of teaching. Results are as follows: TT (mean = 69.24, SD = 13.74, $n = 220$), FT (mean = 76.29, SD = 11.76, $n = 162$), and FTR [mean = 74.2, SD = 12.63, $F(2,654) = 16.01$]. n, No. of students. ***$P < 0.001$, one-way ANOVA.

helpful to their learning. The summary of the survey results was discussed with the students soon after the survey was analyzed.

Statistical analysis. The exam scores were combined based on the teaching methods used, except for the grades from the final exam. A Student t test was carried out to compare the pooled FT and FTR quiz grades and the exam grades from the TT with FT3 methods. A one-way ANOVA was used to compare overall exam scores between TT, FT, and FTR strategies, as well as to compare exam scores of students above and below the 50th percentile between the three teaching approaches. The main effects were further assessed with post hoc Bartlett's tests. A two-way ANOVA was used to compare male versus female student performance, which was followed by Turkey's multiple-comparison test. All tests were conducted with an experiment-wise α-level of 0.05. The data were analyzed using Graph Pad Prism.

RESULTS

The results are summarized below, where both exam and quiz scores are expressed as means ± SD. Exam grades of the teaching styles TT and FT3 were compared using Student's t test to evaluate which teaching approach resulted in higher overall exam grades. The FT3 score [mean (M) = 80.43, SD = 9.559, $n = 53$] was significantly greater ($P < 0.0001$) compared with the TT score ($M = 69.24$, SD = 13.74, $n = 220$; Fig. 1). Due to the fact that students found FT3 to be intense, this method was restructured to lessen the workload outside of class, thus evolving into a new format of FT.

The average exam grade for FT ($M = 76.29$, SD = 11.76, $n = 162$) was significantly different from the TT method ($P < 0.001$). After using this teaching method for three semesters in a row, this method was refined further to include a retrieval strategy, referred to as FTR. Results of this teaching method were $M = 74.2$, SD = 12.63, $F(2,654) = 16.01$, $P < 0.001$ (Fig. 2).

Next, the exam grades were separated into male and female groups, and a two-way ANOVA was conducted. It was interesting to note that male students performed much better in the FT ($M = 76.11$, SD = 11.18, $P < 0.0001$) and FTR ($M = 76.65$, SD = 12.87) methods compared with TT ($M = 66.38$, SD = 13.94, $P < 0.0001$). Female students also scored higher with the FT ($M = 76.27$, SD = 12.29, $n = 91$; $P = 0.0294$) teaching mode compared with TT ($M = 70.85$, SD = 12.91, $n = 110$) and FTR ($M = 72.4$, SD = 11.85, $n = 146$; Fig. 3). There was a significant effect of sex on the teaching method ($F = 7.17$; DFn = 2, DFd = 639). Individual comparisons using Turkey's multiple test followed by two-way

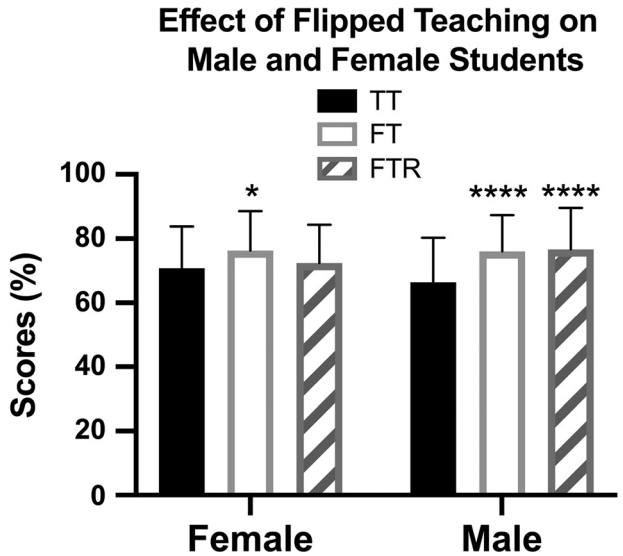

Fig. 3. Comparison of student performance using exam scores between male and female students in the three methods of teaching: traditional teaching (TT), flipped teaching (FT), and flipped teaching with retrieval practice (FTR). Results are as follows: male students, TT (mean = 66.38, SD = 13.94), FT (mean = 76.11, SD = 11.18), and FTR (mean = 76.65, SD = 12.87); female students, TT (mean = 70.85, SD = 12.91, $n = 110$), FT (mean = 76.27, SD = 12.29, $n = 91$), FTR (mean = 72.4, SD = 11.85, $n = 146$); $F = 7.17$; DFn = 2, DFd = 639. n, No. of students. Two-way ANOVA. Turkey's multiple test revealed a significant difference between female TT and FT (*$P < 0.05$), male TT and FT (****$P < 0.0001$), and male TT and FTR (****$P < 0.0001$) groups.

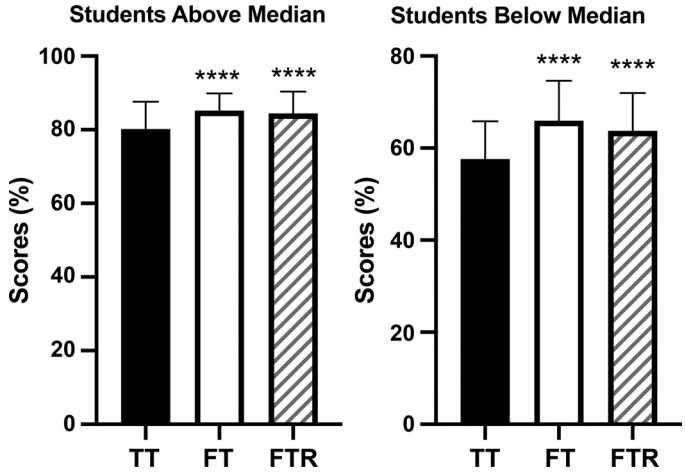

Fig. 4. Comparison of student performance using exam scores in the above 50th percentile and below 50th percentile groups among the three teaching methods: traditional teaching (TT), flipped teaching (FT), and flipped teaching with retrieval practice (FTR). In the above-median group, results are as follows: TT (mean = 80.24, SD = 7.346, n = 113), FT (mean = 85.19, SD = 4.684, n = 87), and FTR (mean = 84.43, SD = 5.957, n = 140); $F(2,337) = 20.26$. ****$P < 0.0001$. In the below-median group, results are as follows: TT (mean = 57.63, SD = 8.232, n = 107), FT (mean = 65.97, SD = 8.633 n = 75), and FTR (mean = 63.75, SD = 8.214, n = 137; $F(2,316) = 26.26$. n, No. of students. ****$P < 0.0001$, one-way ANOVA.

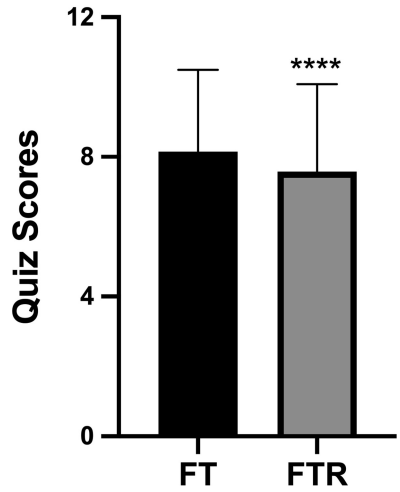

Fig. 5. Comparison of quiz grades between the flipped teaching (FT) and flipped teaching with retrieval practice (FTR) teaching methods. Results are as follows: FT (mean = 8.154, SD = 2.339, n = 1,273) and FTR (mean = 7.574, SD = 2.513, n = 868). n, No. of students. ****$P < 0.0001$, Student's t test.

ANOVA analysis suggested that the level of significance was much greater between the study methods ($P < 0.0001$) for the male students. The level of significance was much lower between the TT and FT methods for the female students ($P < 0.05$). These results suggest that the FT and FTR methods help male students achieve higher scores more than the female students.

The median score was determined for each method of teaching and was used to separate students into above-median and below-median groups. Similarly, When the above-median group was compared, TT score (M = 80.24, SD = 7.346, n = 113) was significantly lower than FT (M = 85.19, SD = 4.684, n = 87) and FTR scores (M = 84.43, SD = 5.957, n = 140). When the below-median group was compared between the three teaching methods, the TT score (M = 57.63, SD = 8.232, n = 107) was again significantly lower than FT (M = 65.97, SD = 8.633, n = 75) and FTR (M = 63.75, SD = 8.214, n = 137; Fig. 4). There was a significant effect in both the upper [$F(2,337) = 20.26$ $P < 0.0001$] and lower [$F(2,316) = 26.26$ and $P < 0.0001$] halves of the class (Fig. 4).

All of the quiz scores from the FT and FTR teaching methods were pooled and compared using Student's t test. The average quiz score for FT (M = 8.154, SD = 2.339, n = 1,273) was significantly higher ($P < 0.0001$) than the FTR quiz score (M = 7.574, SD = 2.513, n = 868; Fig. 5).

The pass/fail rates were compared among the teaching methods and are summarized in Table 2. In terms of the highest grades among the various teaching methods tested, the best grades (A and B combined) were within the FT method (73%), followed by FTR (71%), TT (60%), TT/FT (56%), and TT/FT3 (47%). Attrition rate was very stable among the eight semesters tested. Zero to one student dropped out of the course per semester.

An anonymous student survey collected each semester suggested that students liked the TT method the most (82%) and FTR the least (50%). All groups liked working in groups. However, the FTR method was where group work was most favorable (88%). While students in the TT method liked lectures (60%), students in the FT and FTR methods liked the review of content in the form of mini-lectures the most (87% in FT and 88% in FTR). Similarly, students in the TT method were less appreciative of reading resource compared with both FT and FTR (TT 41%, FT 53%, and FTR 75%). At least 80% of all students in a FT format would recommend FT, and 75% would recommend FTR (Table 3). Additionally, based on the student comments, they were able to prepare for class and not

Table 2. *Grade distribution among teaching methods*

		Grade				
Semester	Teaching Method	A	B	C	D	F
1	*Exams 1–3*: traditional teaching method	25	35	35	5	0
2	*Exams 1–2*: traditional teaching method *Exam 3*: flipped teaching three sessions per week	17	30	45	7	0
3	*Exams 1–2*: traditional teaching method *Exam 3*: flipped teaching spread over 1 wk	15	41	37	7	0
4	*Exams 1–3*: flipped teaching spread over 1 wk	22	56	17	6	0
5	*Exams 1–3*: flipped teaching spread over 1 wk	30	37	30	0	4
6	*Exams 1–3*: flipped teaching spread over 1 wk combined with retrieval practice	26	42	26	0	1
7	*Exams 1–3*: flipped teaching spread over 1 wk combined with retrieval practice	37	37	20	6	0
8	*Exams 1–3*: flipped teaching spread over 1 wk combined with retrieval practice	26	44	26	4	0

Table 3. *Anonymous student survey data*

Questions	TT	FT	FTR
Does the style of teaching support the way you prefer to learn?	Yes (82)	Yes (60)	Yes (50)
I like the fact that we have group work in this course.	Yes (76)	Yes (80)	Yes (88)
I enjoy learning from lecture or mini-lecture.	Yes (65)	Yes (87)	Yes (88)
Reading assignments for this course are appropriate.	Yes (41)	Yes (53)	Yes (75)
Would you recommend flipped teaching in the future?		Yes (80)	Yes (75)

Values in parentheses are percentages of responses.

feel as dependent on the instructor during lectures. Students also reported that they were able to get more out of the course as well as retain the information that they learned.

DISCUSSION

Our results suggest that incorporating the FT method increased overall exam scores compared with TT. The best approach among all of the FT methods tested was FT3, as it resulted in the highest average test scores (Fig. 1), although students as well as the instructor found this method to be the most challenging. Students in the FT3 format reported that the number of preclass assignments and assessments was far too many to complete in the short time they had. Deslauriers et al. (7) reported the actual learning versus feeling of learning in response to being actively engaged in the classroom was inconsistent. The management of large number of quizzes and TBL sessions and the time constraints were equally challenging for the instructor. However, once the design of FT was shifted to reduce the number of TBL sessions and quizzes, students' scores continued to stay stronger in the FT method compared with the TT method, but they also reported more positive perceptions of FT. Thus an optimum method of FT was chosen that was manageable to students and the professor alike.

Many studies have reported success with FT similar to our findings (1, 25, 31). One reason for this success in our study could be regular attendance that this active-learning method requires (2). Our student attendance exceeded 90% each semester with the FT design, whether it was FT3, FT, or FTR, compared with student attendance in the TT method. We also believe that the repeated exposure to the course content and frequent assessments and immediate feedback that the FT method allows in the form of lecture video, lecture slides, readings, quizzes, in-class discussions with the instructor, and peer-teaching in the TBL sessions was a reason why students received higher exam grades in FT. Moreover, the FT method involves frequent formative assessments, where students have opportunities to build their grades to a higher level by participating on a regular basis instead of cramming for the exam the night before. Active participation in the assessments helps students practice and raise their scores while building collaborative skills and confidence. It is also likely that FT instilled a sense of importance in students to come to class to complete in-class activities and assessments, whereas TT did not (17).

Since both FT and FTR were more effective when compared with TT, the specific role of retrieval practice used in the FTR is unclear. It is possible that students are at their best performance with the FT method and no further gains could be accomplished by adding the retrieval strategy. However, retrieval practice involving recalling information is still a very important skill to develop as a student. It reinforces them to depend on themselves rather than accessing resources constantly. Others have reported similar findings where students appear to lack the ability to develop metacognitive awareness (24).

All students, whether they were in the upper or lower 50th percentile, achieved higher scores in the FT and FTR teaching methods compared with the TT method. Similarly, when we shifted our focus to male and female groups, we found that male students' exam scores were significantly lower than those of the female students in the TT method suggesting that male students underperform in the TT method compared with female students. On the other hand, the average scores in the FT and FTR groups were higher for both sexes. In particular, male students' scores were higher with FT and FTR compared with the changes in the scores of female students. The active-learning strategy utilizing lecture videos and TBL sessions in the FT method, as well as the repeated exposure to content, appears to appeal to male students more. It must be noted that female student scores were significantly greater in the TT method compared with those of the male students ($P < 0.05$), suggesting that female students learn well in a unimodal setting (35).

The higher quiz grade in the FT method compared with FTR suggests student dependence on resources in completing quizzes, because quizzes were taken online outside of class in the FT style. The in-class quizzes in the FTR method limited student access to content and hence the significantly lower quiz grades. Similarly, student response to group work was most favorable in the FTR group, suggesting that students interacted more when the access to content was limited. Regular recall through formative assessment and immediate feedback had a positive influence on student performance on their exams, although it was unable to raise the scores above the FT exam averages.

When the pass/fail grades were examined between eight semesters of the course, the highest performance was in the FT method, and a very close second was FTR, just like in the exam performance (Fig. 1). The partially flipped semesters (*semesters 2* and *3*) had the least success rate (Table 2). The attrition rate between semesters was very stable, with zero to one student dropping out of the course. Since this course is a required course and also taken by the students in their senior year, sometimes in their very last semester, there is an urge to complete this course.

In an anonymous survey given to students of the FT class, the majority of the students reported how much they enjoyed having the online resources. "I learned better hearing and seeing the material more than once and being able to re-watch when needed while following along with my PowerPoint slides." The survey data, however, suggest that more students preferred TT compared with FT or FTR. Thus the perception of the teaching methods is inversely related to the student performance and is similar to the findings of Deslauriers et al. (7). Students in the TT method appeared to have underutilized the reading assignment and may have mostly depended on in-class lecture, unlike the students from the FT and FTR groups. This may have contributed to the overall student grades.

A limitation of this study was that the participants in the TT, FT, and FTR methods of teaching were not the same, but they

were students in the classes that were taught using different strategies. Since they were not the exact same students experiencing all of the teaching strategies that were tested, the comparison may have not accounted for unknown factors. Therefore, it is suggested that future studies consider using the same students across different teaching methods to account for those unknown factors.

In conclusion, the FT model does enhance overall student performance across the different populations. Retrieval practice simply echoes the FT approach. Although both male and female students benefit from the FT and FTR teaching methods, male students appear to have a greater degree of benefit.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

AUTHOR CONTRIBUTIONS

C.G. conceived and designed research; C.G. performed experiments; C.G. and A.F. analyzed data; C.G. and A.R. interpreted results of experiments; C.G. prepared figures; C.G. and A.F. drafted manuscript; C.G. and A.R. edited and revised manuscript; C.G. and A.F. approved final version of manuscript.

REFERENCES

1. **Akçayır G, Akçayır M.** The flipped classroom: a review of its advantages and challenges. *Comput Educ* 126: 334–345, 2018. doi:10.1016/j.compedu.2018.07.021.
2. **Bijsmans P, Schakel AH.** The impact of attendance on first-year study success in problem-based learning. *High Educ* 76: 865–881, 2018. doi:10.1007/s10734-018-0243-4.
3. **Bradbury NA.** Attention span during lectures: 8 seconds, 10 minutes, or more? *Adv Physiol Educ* 40: 509–513, 2016. doi:10.1152/advan.00109.2016.
4. **Cavanagh AJ, Aragón OR, Chen X, Couch A, Durham F, Bobrownicki A, Hanauer DI, Graham MJ.** Student buy-in to active learning in a college science course. *CBE Life Sci Educ* 15: ar76, 2016. doi:10.1187/cbe.16-07-0212.

4a. **Daouk Z, Bahous R, Bacha NN.** Perceptions on the effectiveness of active learning strategies. *J Appl Res High Educ* 8: 360–375, 2016. doi:10.1108/JARHE-05-2015-0037.

5. **Della Ratta CB.** Flipping the classroom with team-based learning in undergraduate nursing education. *Nurse Educ* 40: 71–74, 2015. doi:10.1097/NNE.0000000000000112.
6. **DeLozier SJ, Rhodes MG.** Flipped classrooms: a review of key ideas and recommendations for practice. *Educ Psychol Rev* 29: 141–151, 2017. doi:10.1007/s10648-015-9356-9.
7. **Deslauriers L, McCarty LS, Miller K, Callaghan K, Kestin G.** Measuring actual learning versus feeling of learning in response to being actively engaged in the classroom. *Proc Natl Acad Sci USA* 116: 19251–19257, 2019. doi:10.1073/pnas.1821936116.
8. **Freeman S, Eddy SL, McDonough M, Smith MK, Okoroafor N, Jordt H, Wenderoth MP.** Active learning increases student performance in science, engineering, and mathematics. *Proc Natl Acad Sci USA* 111: 8410–8415, 2014. doi:10.1073/pnas.1319030111.
9. **Gopalan C.** Effect of flipped teaching on student performance and perceptions in an Introductory Physiology course. *Adv Physiol Educ* 43: 28–33, 2019. doi:10.1152/advan.00051.2018.
10. **Gopalan C, Fox DJ, Gaebelein CJ.** Effect of an individual readiness assurance test on a team readiness assurance test in the team-based learning of physiology. *Adv Physiol Educ* 37: 61–64, 2013. doi:10.1152/advan.00095.2012.
11. **Gopalan C, Klann MC.** The effect of flipped teaching combined with modified team-based learning on student performance in physiology. *Adv Physiol Educ* 41: 363–367, 2017. doi:10.1152/advan.00179.2016.
12. **Gross D, Pietri ES, Anderson G, Moyano-Camihort K, Graham MJ.** Increased preclass preparation underlies student outcome improvement in the flipped classroom. *CBE Life Sci Educ* 14: ar36, 2015. doi:10.1187/cbe.15-02-0040.
13. **Haak DC, HilleRisLambers J, Pitre E, Freeman S.** Increased structure and active learning reduce the achievement gap in introductory biology. *Science* 332: 1213–1216, 2011. doi:10.1126/science.1204820.
14. **Hacisalihoglu G, Stephens D, Johnson L, Edington M.** The use of an active learning approach in a SCALE-UP learning space improves academic performance in undergraduate General Biology. *PLoS One* 13: e0197916, 2018. doi:10.1371/journal.pone.0197916.
15. **Hamdan N, McKnight P, McKnight K, Arfstrom KM.** *A Review of Flipped Learning*. New York: Flipped Learning Network, Pearson, 2013.
16. **Hanewicz C, Platt A, Arendt A.** Creating a learner-centered teaching environment using student choice in assignments. *Distance Educ* 38: 273–287, 2017. doi:10.1080/01587919.2017.1369349.
17. **Herreid CF, Schiller NA.** Case studies and the flipped classroom. *J Coll Sci Teach* 42: 62–66, 2013.
18. **Hussey HD, Richmond AS, Fleck B.** A primer for creating a flipped psychology course. *Psychol Learn Teach* 14: 169–185, 2015. doi:10.1177/1475725715592830.
19. **Ihm J, Choi H, Roh S.** Flipped-learning course design and evaluation through student self-assessment in a predental science class. *Korean J Med Educ* 29: 93–100, 2017. doi:10.3946/kjme.2017.56.
20. **Jdaitawi M.** The effect of flipped classroom strategy on students learning outcomes. *Int J Instr* 12: 665–680, 2019. doi:10.29333/iji.2019.12340a.
21. **Kapler IV, Weston T, Wiseheart M.** Spacing in a simulated undergraduate classroom: Long-term benefits for factual and higher-level learning. *Learn Instr* 36: 38–45, 2015. doi:10.1016/j.learninstruc.2014.11.001.
22. **Karaca C, Ocak MA.** Effect of flipped learning on cognitive load: a higher education research. *J Learn Teach Digit Age* 2: 20–27, 2017.
23. **Karpicke JD, Blunt JR.** Retrieval practice produces more learning than elaborative studying with concept mapping. *Science* 331: 772–775, 2011. doi:10.1126/science.1199327.
24. **Karpicke JD, Grimaldi PJ.** Retrieval-based learning: a perspective for enhancing meaningful learning. *Educ Psychol Rev* 24: 401–418, 2012. doi:10.1007/s10648-012-9202-2.
25. **Koo CL, Demps EL, Farris C, Bowman JD, Panahi L, Boyle P.** Impact of flipped classroom design on student performance and perceptions in a pharmacotherapy course. *Am J Pharm Educ* 80: 33, 2016. doi:10.5688/ajpe80233.
26. **Krathwohl DR.** A revision of Bloom's Taxonomy: an overview. *Theory Pract* 41: 212–218, 2002. doi:10.1207/s15430421tip4104_2.
27. **Lang JM.** Learning on the edge: classroom activities to promote deep learning. *Faculty* 28: 11–13, 2014.
28. **Lu F, Lemonde M.** A comparison of online versus face-to-face teaching delivery in statistics instruction for undergraduate health science students. *Adv Health Sci Educ Theory Pract* 18: 963–973, 2013. doi:10.1007/s10459-012-9435-3.
29. **Makice K.** *Flipping the Classroom Requires More Than Video* (Online). Wired. https://www.wired.com/2012/04/flipping-the-classroom/ [15 Jan 2019].
30. **O'Flaherty J, Phillips C.** The use of flipped classrooms in higher education: a scoping review. *Internet High Educ* 25: 85–95, 2015. [Erratum in *Internet High Educ* 27: 90, 2015.] doi:10.1016/j.iheduc.2015.02.002.
31. **Roehl A, Reddy SL, Shannon GJ.** The flipped classroom: an opportunity to engage millennial students through active learning strategies. *J Fam Consum Sci* 105: 44–49, 2013. doi:10.14307/JFCS105.2.12.
32. **Schmidt HG, Wagener SL, Smeets GACM, Keemink LM, van der Molen HT.** On the Use and Misuse of Lectures in Higher Education. *Health Prof Educ* 1: 12–18, 2015. doi:10.1016/j.hpe.2015.11.010.
33. **Senske N.** Five years of flipped classrooms: lessons learned. The 33rd National Conference on the Beginning Design Student. Salt Lake City, UT, March 10–11, 2017.
34. **Skiba DJ, Barton AJ.** Adapting your teaching to accommodate the net generation of learners. *Online J Issues Nurs* 11: 5, 2006.
35. **Wehrwein EA, Lujan HL, DiCarlo SE.** Gender differences in learning style preferences among undergraduate physiology students. *Adv Physiol Educ* 31: 153–157, 2007. doi:10.1152/advan.00060.2006.
36. **Yan J, Li L, Yan J, Niu Y.** A comparison of flipped and traditional classroom learning: a case study in mechanical engineering. *Int J Eng Educ* 34: 1876–1887, 2018.
37. **Young MS, Robinson S, Alberts P.** Students pay attention!: Combating the vigilance decrement to improve learning during lectures. *Active Learn High Educ* 10: 41–55, 2009. doi:10.1177/1469787408100194.
38. **Zepke N, Leach L.** Improving student engagement: ten proposals for action. *Active Learn High Educ* 11: 167–177, 2010. doi:10.1177/1469787410379680.