



A partial flip classroom exercise in a large introductory general biology course increases performance at multiple levels

Neil Lax, James Morris & Benedict J. Kolber

To cite this article: Neil Lax, James Morris & Benedict J. Kolber (2016): A partial flip classroom exercise in a large introductory general biology course increases performance at multiple levels, Journal of Biological Education, DOI: [10.1080/00219266.2016.1257503](https://doi.org/10.1080/00219266.2016.1257503)

To link to this article: <http://dx.doi.org/10.1080/00219266.2016.1257503>

 View supplementary material [↗](#)

 Published online: 17 Nov 2016.

 Submit your article to this journal [↗](#)

 View related articles [↗](#)

 View Crossmark data [↗](#)

A partial flip classroom exercise in a large introductory general biology course increases performance at multiple levels

Neil Lax^{a,b}, James Morris^c and Benedict J. Kolber^{a,b}

^aDepartment of Biological Sciences, Duquesne University, Pittsburgh, PA, USA; ^bChronic Pain Research Consortium, Duquesne University, Pittsburgh, PA, USA; ^cDepartment of Biology, Brandeis University, Waltham, MA, USA

ABSTRACT

Incorporation of active learning into large lecture classes is gaining popularity as a pedagogical method due to its known benefits in helping learning outcomes. A more recent active learning technique that has emerged is the flipped classroom. In this study, we investigated the effects of incorporating a 'partial-flip' into an introductory general biology course, where only a portion of the class time was spent in a flipped classroom format. This partial flip presented the classic biology experiment of Meselson and Stahl that discovered the mechanism of DNA replication. Performance on in-class formative assessments and subsequent summative assessments (eg out-of-class assignments and exams relating to this material) was compared between a class that had the partial flip and a control class that had a traditional lecture. The partial flip students scored higher on in-class formative questions, specific exam questions and final exam essay questions. We found that the partial flip had different effects in males versus females depending on the assessment. The partial flip manipulation appeared equally effective in aiding both below and above average students in formative assessments. Overall, this study shows that the partial flip classroom can be an effective technique to incorporate into existing courses and that it does provide some benefits compared to traditional lecture.


KEYWORDS

Introductory biology; active learning; flipped classroom; partial flip

Introduction

Traditionally, lecturing has been the standard form of teaching in large introductory science classes. Lectures allow instructors to cover large amounts of material, including much of the fundamental information that students need for advanced biology courses. One problem with lectures is that most students fail to retain the more challenging topics and information is typically presented at a low Bloom's cognitive level (Bloom et al. 1956; Handelsman et al. 2004). Moreover, since students in these introductory classes come from diverse backgrounds with different learning styles, utilising only lectures may be grossly ineffective. Recently, instructional methods have started to include more active learning strategies. In general, these strategies generate greater retention and deeper understanding of material (Huang and Carroll 1997; Pelaez 2002; Plendl et al. 2009).

Active learning methods in the classroom have been shown to be more effective than traditional lectures in science, technology, engineering and mathematics courses. In these courses, students are 1.5 times more likely to fail in a lecture class compared to an active learning class (Freeman et al. 2014).

CONTACT Benedict J. Kolber  kolberb@duq.edu

 The supplementary data for this article is available online at <http://dx.doi.org/10.1080/00219266.2016.1257503>

While active learning strategies, such as problem-based learning and learner-centred approaches, have been shown to have minimal effects on surface learning of basic concepts, they tend to more significantly enhance deep learning of the material (Dangel and Wang 2008; Dolmans et al. 2015; Spencer and Jordan 1999). Incorporating active learning techniques has been shown to be effective for deep learning in larger introductory science courses (Bevan, Chan, and Tanner 2014). Moreover, students that are taught in a classroom that utilises active learning have shown growth in their problem-solving abilities over the course of a term (Klegeris, Bahniwal, and Hurren 2013). This implies that incorporating active learning into classrooms can lead to a stronger understanding of material and improve critical thinking skills.

One active learning instructional method that has not been extensively explored in large introductory biology classes is the use of flipped classrooms, which are also known as inverted, hybrid or blended classrooms (Margulieux et al. 2014). This pedagogy normally incorporates multimedia technology into the curriculum, but students can do a variety of activities including watching a pre-recorded lecture, watching an interactive video, listening to a podcast or reading an assigned textbook chapter (Bergmann and Sams 2012; Johnson et al. 2015; Margulieux et al. 2014; O'Flaherty and Phillips 2015). This model of teaching requires students to invest their own time in preparatory work before class and learn the material at their own pace. Then, in class, students spend time in groups, working together, discussing and answering questions while working on questions or activities related to the pre-class exercise. The instructor, instead of lecturing on the class material, acts as a guide for students, answering questions and helping students with the in-class activity. By working through basic concepts on their own during the pre-class activity, students spend class time in a greater cognitive engagement with the material, thereby facilitating deeper learning and better mastery of concepts. Studies have shown that flipped classroom students perform better on exams and have a more enjoyable classroom experience than traditional lecture students (McLean et al. 2016; Mortensen and Nicholson 2015) and also are more likely to attend class (Prober and Khan 2013). Additionally, flipped classrooms have been used in large introductory sciences classes and have led to learning outcomes that were the same or better than traditional lecture classes (Baepler, Walker, and Driessen 2014). Ideally however, flipped classrooms involve smaller numbers of students in an environment where students can work in small groups (Berrett 2012). We were interested in exploring how this pedagogy could be incorporated into a large course in an auditorium-style room.

We presented one experiment in a flipped classroom format while leaving the rest of the course material in traditional lecture style. Because the flipped exercise covered only a small portion of time (one-third of the class period), this is classified as a 'partial-flip' exercise (van Vliet, Winnips, and Brouwer 2015). In our partial flip, we presented the Meselson and Stahl DNA replication experiment (Meselson and Stahl 1958). This is a classic experiment that is typically covered in general biology courses and is conceptually complex for many students. Performed in 1958 by Matthew Meselson and Franklin Stahl, this experiment provided substantial evidence for the semi-conservative basis of DNA replication. Due to the elegant experimental design and clearly testable hypotheses, it is often called 'the most beautiful experiment in biology' (Holmes 2001). Despite this, the principle behind the experiment and the meaning of the results are often difficult for students to grasp.

In 2013, the instructor (BJK) identified the Meselson and Stahl experiment as an area of student confusion. In response to students' struggles with higher order Bloom questions about the experiment in 2013, we were interested in determining whether or not a partial flip would allow students to more fully understand and retain the information. In 2014, the instructor implemented a modified partial flip exercise for this material. Comparing 2014 to 2013, there was a qualitatively small improvement in the Meselson and Stahl multiple-choice question correct response rate (data not shown). The lack of a large difference between the years could be due to a number of variables beyond the implementation of the partial flip exercise. This illustrates one of the challenges of evaluating new teaching strategies where a comparison between years introduces uncontrolled variables (eg time, cohort differences, instructor experience). To better evaluate this question about whether an active learning partial flipped exercise improves student outcomes, we performed a case-control study between two general biology sections in 2015, presented here.

One section had the Meselson and Stahl experiment presented in a partial flip format, while the other section used a traditional lecture. We compared responses on formative assessments with clicker questions during class (diagnostic, low-stakes assessments with no grade or low number of points [Suskie 2010]) and summative after-class homework assignment questions (graded, higher stakes with greater number of points [Suskie 2010]) related to the experiment. Additionally, we evaluated performance on summative examination questions (exam 3 and exam 4) related to the experiment from the two sections. Overall, we showed that the partial flip students performed better on formative in-class questions and some summative exam multiple-choice and essay questions. These data support better immediate and long-term retention of Meselson and Stahl content with a partial flip intervention. We found a gender-related effect where the partial flip classroom appeared to be more effective for immediate performance in females but effective for later exam performance in males. Overall, these findings offer insight into benefits that may be derived from incorporating flipped or partial flipped classrooms into large introductory biology courses.

Methods

Participants and consent

This study was evaluated and approved by the Duquesne University (Pittsburgh, PA) Institutional Review Board (protocol #2015-08-1). Participants were undergraduate students at Duquesne University enrolled in the fall 2015 introductory General Biology 111/111L course. This course consisted of two separate sections, A and B, henceforth referred to as 'Control' and 'Partial-flip', respectively. The majority of students were first-year students (88%). These students' course of study included forensics science and law, physical therapy, occupational therapy, pre-pharmacy, biochemistry, chemistry, physician assistants, athletic training, health management systems, chemistry, biomedical engineering, biology and psychology (see supplemental Table 1 for details).

Participants were enrolled in the study at the beginning of the term by an experimenter not involved in teaching or grading the course. Students read and signed consent forms. Section A—221 total students; 171 consent; 33 excluded from study because no demographic data; 8 excluded from study because they did not attend class on experimental day. In this section, a majority of the students included in the study (~75%) had pharmacy as their course of study. Section B—239 total students; 201 consent; 32 excluded from study because no demographic data; 11 excluded from study because they did not attend class on experimental day. In this section, a majority of the students (~60%) were from three health sciences: physical therapy, physician assistant and occupational therapy. Demographic data collected from consenting participants included gender (listed as 'sex' on survey form), age, year at institution and major (course of study).

Study design

All students in section A were assigned to the 'lecture-only' control group. All students in section B were assigned to the 'partial-flip' experimental group. All students in each section received the same instruction and assessments even if they did not provide informed consent. The manipulation implemented is considered standard teaching practice and the lessons were specifically designed so that all students were exposed to the exact same material and provided similar opportunities to learn the material. The main differences between the groups were the timing of the 'lecture' and worksheet components of the lesson. The experimental manipulation centred on teaching Meselson and Stahl's 'Most Beautiful Experiment in Biology' (Meselson and Stahl 1958). This area was identified in previous years as a challenging lesson for students (data not shown). Two days prior to the experiment, students in both sections were asked to read the relevant section in the course textbook (Morris et al. 2013). Students in the control group then answered two multiple-choice questions about the material with unlimited attempts to get the correct answer prior to class. Students in the partial flip group watched a

16-min lecture that was recorded by the instructor demonstrating the Meselson and Stahl experiment. Following the recording, students answered the same two multiple-choice questions as the control group along with a 'password' question, the answer to which was embedded in the recorded lecture. The password question was used to ensure that students actually watched the recorded lecture prior to class.

During the class session, the control group received the same content from the recorded lecture in a live lecture format from the course instructor (lecture time 20 min). At the end of the lecture, the students were given two questions, answered using personal response system (PRS) remotes (Turning Point Technologies). Students without remotes were asked to provide answers on 3×5 cards. Students received participation credit for answering the questions. For the partial flip group, students worked on a group worksheet for 20 min. This worksheet challenged students to recreate the experiment and determine conclusions based on hypothetical situations and results. During the worksheet component of the lesson, four individuals (two course instructors plus two graduate-level teaching assistants) walked around the room to aid students. To facilitate discussion between staff and students, four rows in the auditorium-style classroom were left empty to act as aisles for the staff. At the end of the worksheet, the students were given three PRS questions. Students without remotes were asked to provide answers on 3×5 cards.

Following class, for both groups, students completed an online assignment that covered the Meselson and Stahl experiment. The control group completed identical questions to the partial flip group's in-class worksheet plus an additional two 'novel' questions. The partial flip group completed the two 'novel' questions only. On two subsequent exams (exams 3 and 4 of the term), all students received identical multiple-choice questions and an essay question about the Meselson and Stahl experiment. Cognitive level of questions ranged from knowledge-based to application-type questions.

Recording and Classroom Observation Protocol for Undergraduate STEM (COPUS) scoring

Both sections of the course were recorded during the relevant period of the manipulation for post hoc COPUS analysis (Smith et al. 2013). Recording was announced prior to class and camera was directed away from the first four rows in the classroom. Students who did not wish to be recorded were asked to sit in these first four rows.

Videos were analysed by three individuals following a standardisation and training session. COPUS scoring (Smith et al. 2013) was performed using the online GORP software (Tools for Evidence-Based Action, University of California, Davis). The following student actions were scored: listening (Student-L), waiting (Student-W), individual thinking/problem-solving (Student-Ind), student answering a question posed by instructor with rest of class listening (Student-AnQ), student asks question (Student-SQ), student making prediction about a demo or experiment (Student-Prd), group activity (Student-OG), students answering clicker question (Student-CG), students working in groups on worksheet (Students-WG) and other student actions (Student-Other). The following instructor actions were scored: lecturing (Instructor-Lec), waiting (Instructor-W), real-time writing on board or projector (Instructor-RtW), administrative activity (Instructor-Adm), demonstration or experiment (Instructor-DV), listening to and answering question, while entire class listens (Instructor-AnQ), providing follow-up feedback on clicker question (Instructor-FUp), moving through class guiding ongoing active learning task (Instructor-MG), answering 1 on 1 questions from students during active learning task (Instructor-1o1), posing a clicker question (Instructor-CG), posing a non-clicker question (Instructor-PQ). Actions were scored in two-min bins across the entire video; actions were then averaged between the three individual scorers.

Data analysis and statistics

Statistical analysis was performed on Prism (GraphPad version 6). All data analysis was performed at the end of the term to avoid any biases in grading. Data were analysed by unpaired *t*-test. *P* values ≤ 0.05 were considered statistically significant. All data are presented as means \pm SEM.

Results

Qualitative analysis of instructor and student activity

Using the COPUS analysis (Smith et al. 2013), we quantified how time was spent by the students and instructor in the control section compared to the partial flip section. Here, both the students and instructor were recorded during the presentation of the Meselson and Stahl material and their actions were scored. Using COPUS, we found that both the instructor and students in the control

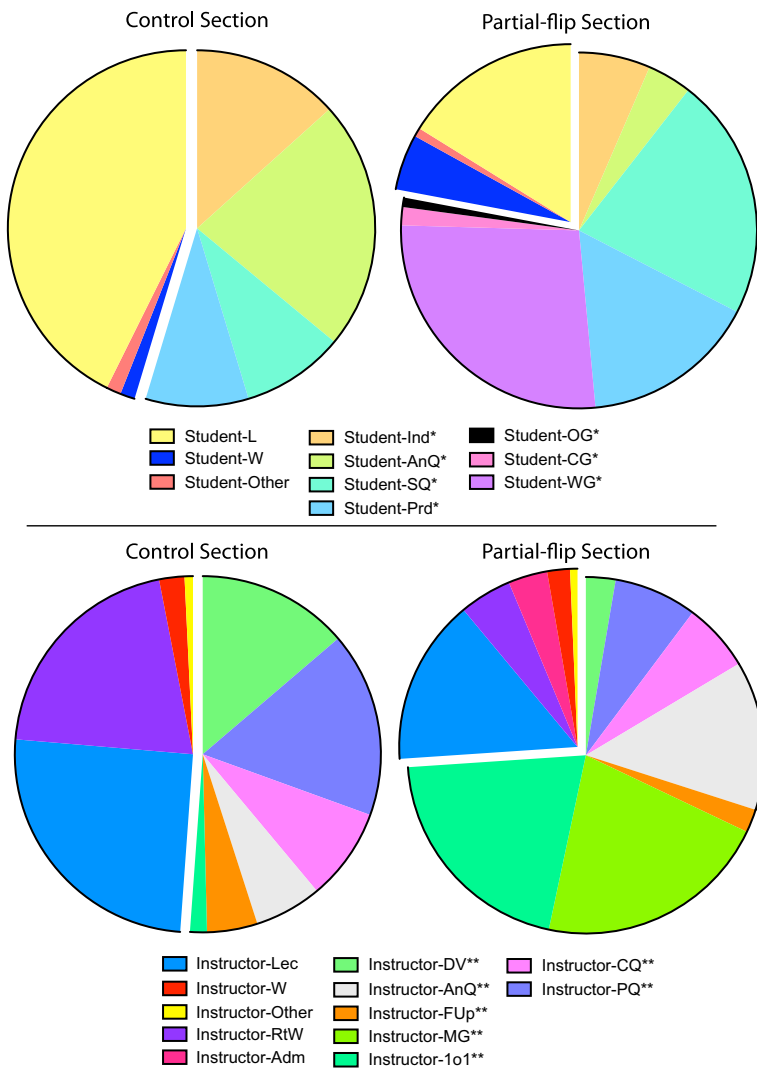


Figure 1. Quantification of student and instructor actions.

Notes: (A) Student and (B) instructor actions were analysed using the COPUS scoring system (Smith et al. 2013) for the control section and the partial flip section. Actions were only scored during the in-class period that covered the Meselson and Stahl experiment. *criteria associated with active learning; **criteria associated with active teaching. Student-L: listening; Student-W: waiting; Student-Ind: individual thinking/problem-solving; Student-AnQ: student answering a question posed by instructor with rest of class listening; Student-SQ: student asks question; Student-Prd: student making prediction about a demo or experiment; Student-OG: group activity; Student-CG: students answering clicker question; Students-WG: students working in groups on worksheet; Student-Other: other student action; Instructor-Lec: lecturing; Instructor-W: waiting; Instructor-RtW: real-time writing on board or projector; Instructor-Adm: administrative activity; Instructor-DV: demonstration or experiment; Instructor-AnQ: listening to and answering question while entire class listens; Instructor-FUp: providing follow-up feedback on clicker question; Instructor-MG: moving through class guiding ongoing active learning task; Instructor-1o1: answering 1 on 1 questions from students during active learning task; Instructor-CG: posing a clicker question; Instructor-PQ: posing a non-clicker question.

section spent only about 50% of class time involved in active learning/teaching activities (Student: Ind, AnQ, SQ, Prd, OG, CG, WG; Instructor: DV, AnQ, FUp, MG, 1o1, CQ, PQ). Not surprisingly, in the partial flip section, both the students and professor spent approximately 75% of class time in active learning/teaching (Figure 1).

Overall effect of the partial flip on learning

After showing that the students and instructor in the partial flip section did spend more time involved in active learning activities, we measured learning outcomes in both groups. When looking at in-class formative assignments, students in the partial flip section performed better on an in-class clicker question that required them to predict an experimental outcome (Figure 2(A), unpaired t -test, $***p < 0.0001$). These same students also had higher confidence in the material (Figure 2(B), unpaired t -test, $***p < 0.0001$) and they viewed the partial flip as an effective way to learn the material (1–5 Likert-type scale response to ‘Partial-flip was effective’ 1 = Strongly disagree, 5 = Strongly agree; mean = 4.1 ± 0.06). The partial flip group also showed a trend in performing better on a post-class assignment (see supplemental Figure 1(A)). When looking at performance on the exams after the flip (exams 3 and 4), there was no overall difference in performance on multiple-choice questions (see supplemental Table 2 for exact multiple choice questions with associated Bloom’s taxonomy level for both exams 3 and 4) related to the Meselson and Stahl experiment between the sections (see supplemental Figure 1(B) and (C)). Analysis of individual questions, however, did point to a difference between the two groups. In question 1 of exam 3, the partial flip section scored significantly higher compared to the control section (Figure 2(C), unpaired t -test, $*p < 0.05$). Additionally, on the essay question included in exam 4, students in the partial flip section scored significantly higher than the students in the control group (Figure 2(D), unpaired t -test, $*p = 0.04$). These results show that there were

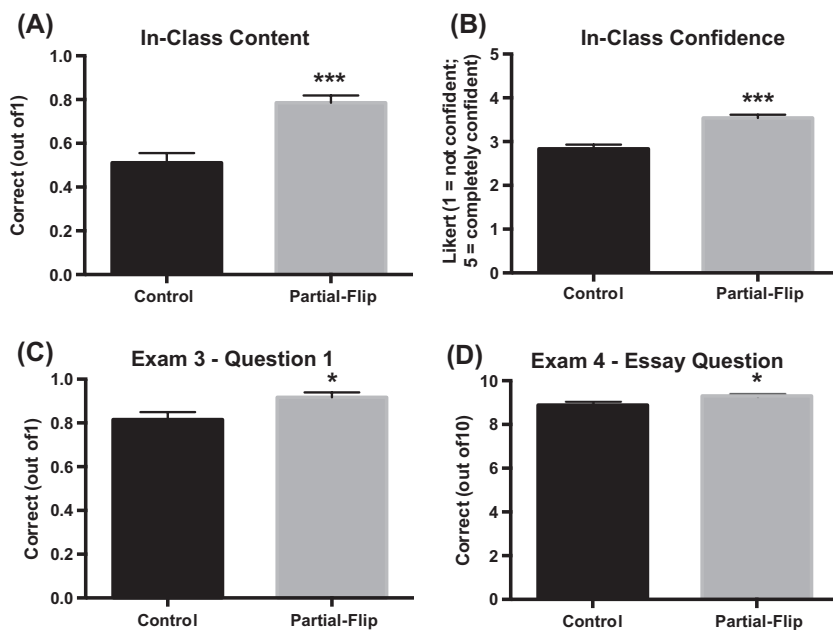


Figure 2. Partial flip exercise increases learning outcomes.

Notes: (A) Students in the partial flip section performed better as a class on an in-class clicker question involving prediction of an experimental outcome (t -test $***p < 0.0001$). (B) Students in the partial flip section reported have higher confidence in the Meselson and Stahl experiment compared to the control students (1–5 Likert-type scale 1 = Completely not confident, 5 = Completely confident; t -test $***p < 0.0001$). (C) Students in the partial flip course scored significantly higher on a single multiple-choice question (question 1) on exam 3 (t -test $*p = 0.01$). (D) Students in the partial flip course scored significantly higher on an essay question about the Meselson and Stahl experiment on exam 4 (t -test $*p = 0.04$).

performance differences with in-class response to questions, post-class assignment results and exam question responses of the partial flip and control group, with the partial flip group performing better.

Gender-specific effects of the partial flip on activities and assignments

Next, we looked at effects of the partial flip environment on male versus female students. In analysing in-class and out-of-class activities/assignments, we saw that both males and females who were in the partial flip group performed significantly better on in-class content questions related to the Meselson and Stahl experiment (Figure 3(A) and (B), unpaired *t*-tests, $*p < 0.05$ male, $***p < 0.0001$ female). Females in the partial flip group reported more confidence in the material compared to female students in the control section (Figure 3(C), unpaired *t*-test, $***p < 0.0001$), while males showed no difference in confidence between the sections (Figure 3(D), unpaired *t*-test, n.s.). Moreover, females in the partial flip section showed a trend for scoring higher on a post-class assignment compared to the control section (Figure 3(E), unpaired *t*-test, $+p = 0.07$), while males from both sections scored equally well on the assignment (Figure 3(F), unpaired *t*-test, n.s.).

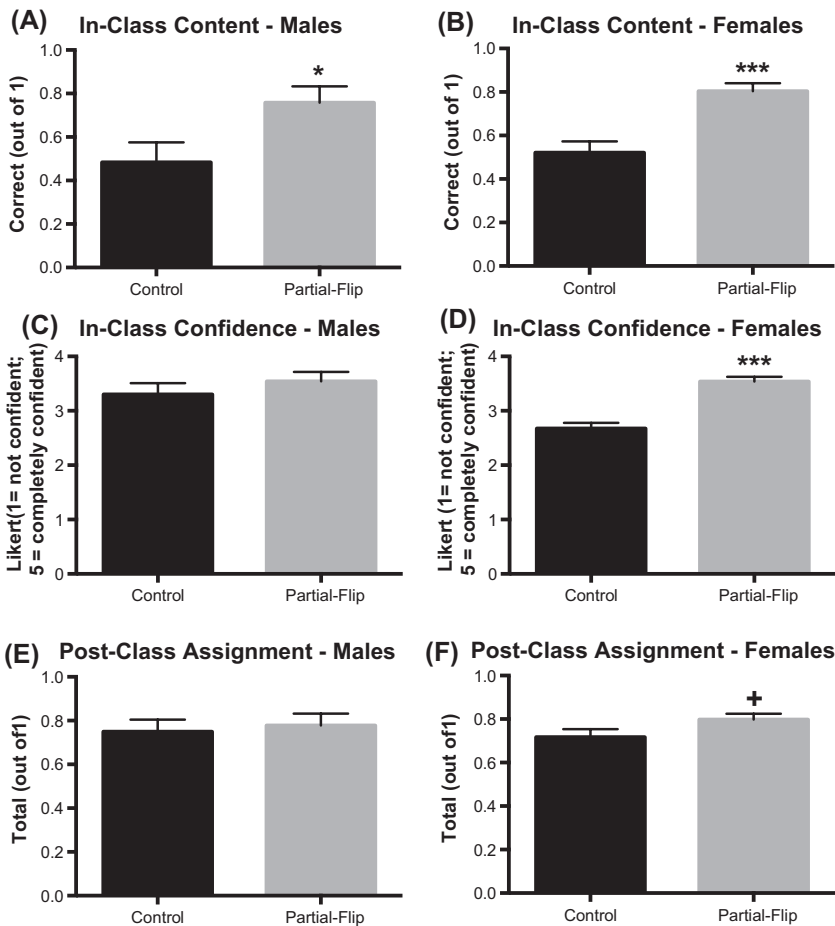


Figure 3. Partial Flip manipulation had different effects in male vs. female students for in-class and post-class evaluations.

Notes: (A) Both male and (B) female students in the partial flip section performed better on an in-class clicker question involving prediction of an experimental outcome compared to control section students (*t*-test $***p < 0.0001$, $*p = 0.02$). (D) Females but not (C) males in the partial flip section reported higher confidence for the Meselson and Stahl experiments compared to the control section (*t*-test $***p < 0.0001$). (F) Females but not (E) males in the partial flip section scored a trend for a higher score in the post-class assignment (*t*-test $+p = 0.07$).

Gender-specific effects of the partial flip on examination performance

Looking at examination performance between males and females from both sections, we found that males from the partial flip group showed a difference in multiple-choice performance with males in the partial flip group scoring higher on these questions compared to the control group (Figure 4(A), unpaired t -test, $*p < 0.05$). Females from both the control and partial flip groups, however, scored

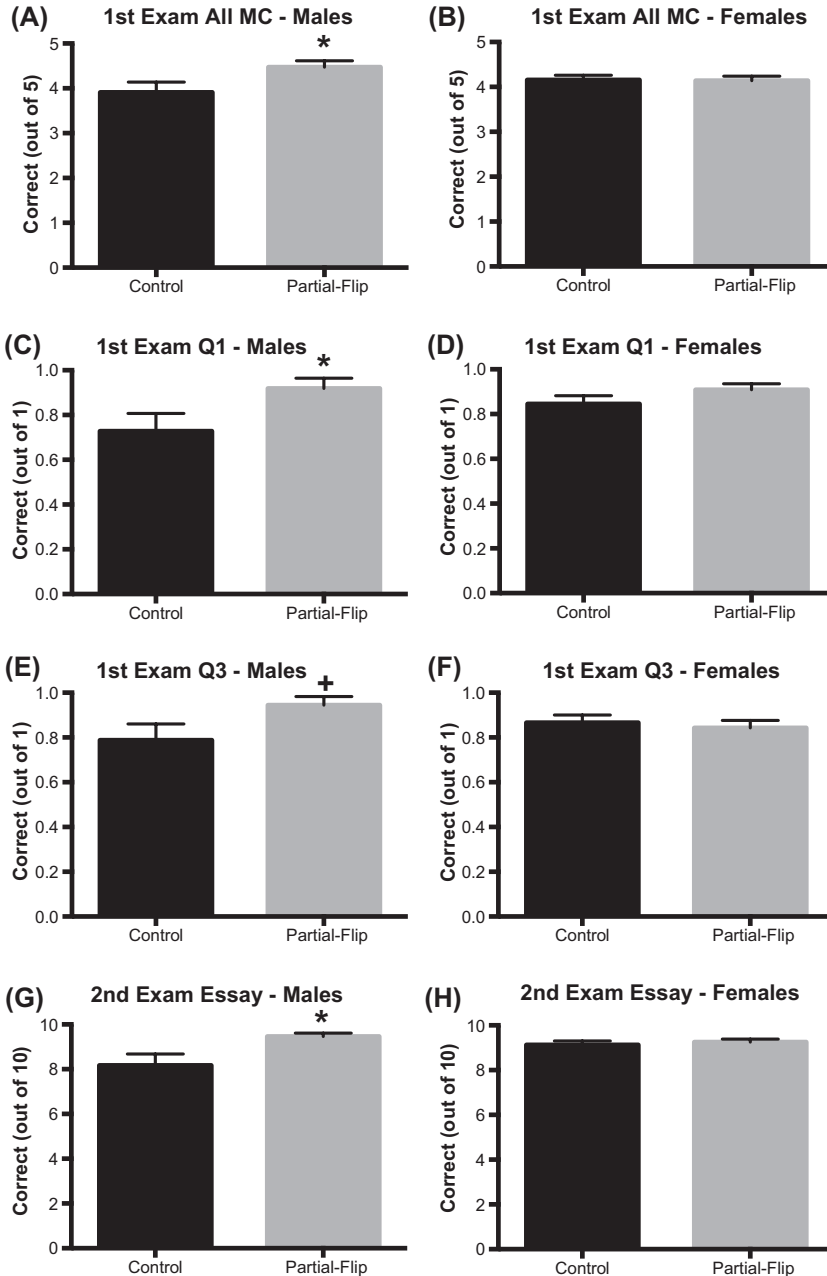


Figure 4. Partial Flip manipulation had different effects in male vs. female students on examinations.

Notes: (A) Male but not (B) female students in the partial flip section performed better on exam 3 considering all multiple-choice (MC) questions compared to control groups (t -test $*p = 0.04$). (C, E) Males but not (D, F) females in the partial flip section scored higher on questions 1 (Q1) and 3 (Q3) compared to control groups (t -test C: $*p = 0.03$, E: $+p = 0.05$). (G) Male but not (H) female students in the partial flip section performed better on the exam 4 essay question compared to control groups (t -test $*p = 0.01$).

equally well on all five multiple-choice exam 3 questions related to Meselson and Stahl (Figure 4(B), unpaired t -test, n.s.). Of these five multiple-choice questions, two in particular showed significant differences (Q1 and Q3). While, females from the control section and partial flip section showed no differences in performance on these two questions (Figure 4(D) and (F), unpaired t -tests, n.s.), for both of these questions, males in the partial flip group scored significantly higher than the control section (Figure 4(C) and (E), unpaired t -test, $P \leq 0.05$). Interestingly, on the second exam after the partial flip (exam 4), neither females nor males showed any difference with performance on multiple-choice questions related to the Meselson and Stahl experiment (see supplemental Figures 2(A) and (B)). However, males from the partial flip group scored higher on an essay question compared to males from the control section (Figure 4(G), unpaired t -test, $*p < 0.05$). This difference was not seen in female students (Figure 4(H), unpaired t -test, n.s.).

Partial flip effects on above and below average students

Finally, we evaluated whether the partial flip pedagogy would differently help students performing below average or above average in the class. Below or above average groups were determined by dividing each section based on the overall course average score on exam 2 (the exam prior to the experimental manipulation) and statistically splitting the class into two data-sets: below average and above average. For below average students, we found that students in the partial flip group performed better on an in-class question related to the Meselson and Stahl experiment (Figure 5(A); $p < 0.01$) and rated their confidence in the material as higher (Figure 5(B); $p < 0.05$) compared to the below average students in the control section. Similarly, in the above average analysis, we also found that students in the partial

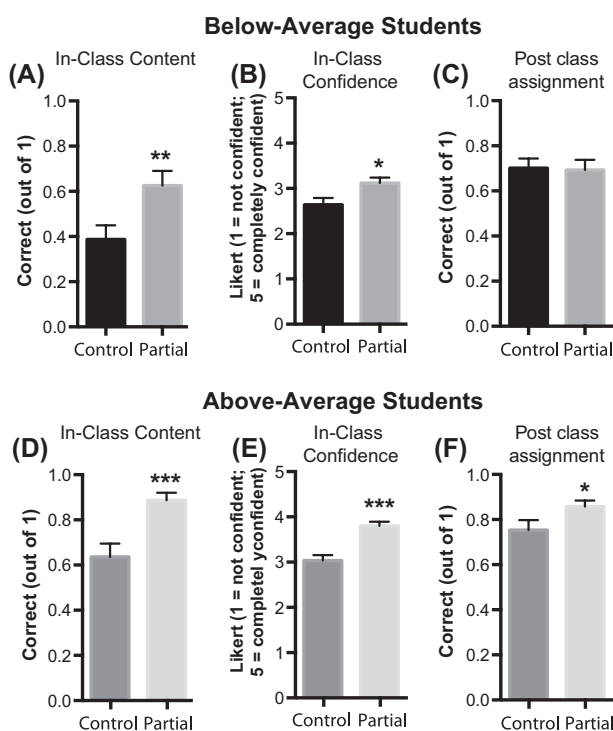


Figure 5. Partial flip pedagogy aids both below and above average students primarily in formative assessments.

Notes: Below average students in the partial flip section ($n = 56$) scored higher than below average students in the control section ($n = 62$) on formative in-class (A) content questions and (B) confidence questions but not on a summative (C) post-class assignment. Above average students in the partial flip section ($n = 98$) scored higher than above average students in the control section ($n = 67$) on an in-class (D) content question, (E) an in-class confidence question, and a summative (F) post-class assignment. (unpaired t tests compared to control group $*p < 0.05$, $**p < 0.01$, $***p < 0.001$).

flip group performed better on an in-class questions related to the Meselson and Stahl experiment (Figure 5(D); $p < 0.001$) and rated their confidence in the material as higher (Figure 5(E); $p < 0.001$) compared to the above average students in the control section. Thus, for our formative assessments, the partial flip helped both below and above average students. This is in contrast to data from summative assessments of this material. In the post-class summative assignment, below average students in the partial flip group performed similar to the below average students in the control section (Figure 5(C); $p > 0.05$). In contrast, above average students in the partial flip group performed better on the post-class assignment compared to the above average students in the control group (Figure 5(F); $p < 0.05$). Above and below average students scored equally well on exam 3 multiple choice related to the Meselson and Stahl experiment (see supplemental Figure 3(A) and (B)).

Differences between partial flip and control students

We analysed if the control and partial flip groups were different from each other in ways other than performance on Meselson and Stahl material. Since the same instructional technique was used for the rest of class material in both of the sections, we expected there to be no differences between them. Indeed, this was the case when looking at the overall performance on exam 3 (exam immediately following experimental manipulation): both the control and partial flip groups scored equally well on the whole exam (control: $83.61\% \pm 0.967$, $N = 130$, partial flip: $85.30\% \pm 1.022$, $N = 158$, unpaired t -test, n.s. $p = 0.2352$). However, we did see differences in overall exam 4 performance. In this parameter, the partial flip group scored significantly higher than the control group (control: $76.16\% \pm 1.030$, $N = 130$, partial flip: $80.45\% \pm 0.898$, $N = 158$, unpaired t -test, $**p = 0.0018$). While, we were able to show significant differences between outcomes relating to the Meselson and Stahl material, these data from exam 4 suggest that there may have been other differences between the two sections in terms of overall performance.

Discussion

The flipped classroom increases active learning

Flipped classrooms have recently emerged as a method to increase student engagement with class material (Hamdan et al. 2013). To our knowledge, no studies to date have investigated how this methodology could be incorporated into already existing large introductory biology classes in a 'partial-flip' model to supplement existing lecture material on complex topics. Here, we show that presenting the Meselson and Stahl experiment in a flipped classroom format helps increase performance on both formative and summative assessments. Furthermore, we demonstrate that the partial flip classroom model may have different effects on male versus female students depending on the specific assessment. Finally, we show that the partial flip pedagogy helps students from both ends of the grade spectrum.

Active learning strategies, such as the flipped method, that focus on shifting the acquisition of fact-based information outside of the classroom are greatly beneficial because they allow the class to spend more time engaging with the material (Michael 2006). This allows students to interact with the professor and each other by asking questions and working on problems rather than just listening to a presentation during class (Huba and Freed 2000). As was shown in the COPUS analysis, compared to a traditional lecture on the same material, using the partial flip, students and the instructor spent about 25% more time in active learning actions during that class period. This shows how easily a different instructional method can be used to present the same material and increase the amount of active learning. These 'hybrid-lecture' classes are gaining popularity because active learning has been shown to increase exam grades, attendance rates and student perception of the class (McFarlin 2008; Riffell and Merrill 2005; Riffell and Sibley 2004).

The partial flipped classroom helps improve student learning outcomes

Flipped classrooms have the advantage of increased interaction with the professor and other students, leading to a better understanding of the material (Warter-Perez and Dong 2012). While the empirical data on flipped classrooms is limited, flipped classes have been shown to increase standardised test performance in mathematics (Fulton 2012) and decrease failure rates in reading, mathematics, science, social studies and writing (Green 2012). However, some studies have shown that there are no differences in students' knowledge in a flipped classroom (Frederickson, Reed, and Clifford 2005) and that the flipped model may not be best used in introductory classes where students do not have a deep interest in the subject (Strayer 2012). Looking at our results and the overall effects of the partial flip compared to traditional lecture, we were able to show some benefits of this model. Beyond in-class content questions, we found statistically significant differences class-wide on performance on only one multiple-choice question on exam 3 (the first exam following the manipulation) but did find enhanced performance on an essay question on exam 4 (given on the second exam after the manipulation). These data show that the effects of the partial flip may be stronger in higher level thinking situations—the partial flip improved scores on an essay question where students had to explain their reasoning. In contrast, on lower level multiple-choice style questions, students from both sections scored equally well on all but one question from both exams. These results agree with data that show active learning strategies increase deep learning but have minimal effects on surface learning (Dangel and Wang 2008; Dolmans et al. 2015; Spencer and Jordan 1999). Our results may also suggest that presentation of the Meselson and Stahl experiment using the flip classroom helps with long-term retention of the material; however, additional studies are needed to determine if the partial flip is better for higher order thinking, long-term understanding or both. Moreover, our results are based on a small number of questions pertaining to the Meselson and Stahl experiment covered in the partial flip. While the addition of more questions relating to the Meselson and Stahl experiment would have been ideal, only a limited number of questions related to that material were included, to avoid excluding material from other lectures. Despite the small number of questions however, our results show that the partial flip did not deleteriously affect learning outcomes, and that the partial flip may work as well as, if not better than, traditional lecture.

A consideration of any new pedagogy is the impact on faculty. In this context, the development and implementation of the partial flip experience was not overly burdensome on the course instructor. The total amount of extra time dedicated to the development of this technique was 3 h including time to develop the worksheet (~2 h) and record, edit and upload the pre-lecture video. Fortunately, after an instructor develops a partial flip worksheet and video, those materials can be used each year with minimal additional effort.

Males and females respond differently to partial flip model

Interesting effects of the partial flip model were also seen when analysing male students compared to females. Meta-analyses have shown that gender-related differences in cognitive performance in science and mathematics are small and declining (Linn and Hyde 1989). While these differences in cognitive ability are almost non-existent, hands-on activities have been shown to be more valuable to females in science and mathematics education and lead to better performance (Burkam, Lee, and Smerdon 1997; Peterson and Fennema 1985). Our data show that the hands-on nature of the partial flip differentially affected male and female students. Both genders in the partial flip section scored higher on the in-class content questions compared to the control section, but the partial flip females reported significantly more confidence in the material and also performed better on post-class assignments. Males did not show differences in confidence, suggesting that the partial flip helps female students be more comfortable with the material. Subsequently, the partial flip seems to have helped male students more than female students on exams. This effect was driven by two specific questions (Q1 and Q3), questions that were on different levels of Bloom's taxonomy (Q1—identify, Q3—apply; supplemental

Table 2). Females showed no differences in any of these exam questions. This means that the overall class effects on exam 3 multiple-choice questions were likely driven by the increased performance in males. This same difference held true for the essay on exam 4—males from the partial flip scored significantly higher than control male students on the essay, while female students showed no such difference between groups.

These data highlight some potentially interesting phenomena related to the partial flip model. Our results suggest that the partial flip gave female students more reported confidence in the material while not actually generating any differences in summative assessments. Conversely, male students did not report confidence with the material following the partial flip, but actually performed better on summative assessments. In other words, while not realising it, males may have gained a benefit from the partial flip. Females, on the other hand, only had a perception of a greater understanding. Reasons behind this phenomenon are unknown. Other studies have found similar results, with females showing trends for lower confidence in STEM fields (Strenta et al. 1994). Females tend to be more critical of their abilities, while males tend to have unrealistic estimates of their knowledge (Pomerantz, Altermatt, and Saxon 2002). In our study, female students with lower expectations of themselves may have rated their confidence as higher after going through the partial flip session. On the other hand, males who are already over confident do not gain any additional confidence from the partial flip because they are already at a high level. Additional studies are required to tease apart the mechanisms behind the effects we found.

The partial flip classroom effects students on varying ends of the grade spectrum

Our results also indicate some differences of the partial flip methodology on above versus below average students. Previously, it has been shown that hybrid classrooms lead to relatively equal benefits for students of different aptitudes and abilities, and may help students on the lower end of the grade spectrum more significantly (Baepler, Walker, and Driessen 2014). From our analysis, the partial flip helped both ends of the spectrum on formative assessments. This was in contrast to some of our results from the summative assessments. While below average students from the partial flip and control sections performed equally well on the post-class assignment, above average students from the partial flip group scored higher than above average students in the control section. However, this may reflect a recency effect for the partial flip manipulation that existed for the above average students only. There were no differences in either the above average or below average students between the partial flip and control section on exam 3 multiple-choice questions. These results were expected because more responsibility falls onto students to learn the basics outside of class in the partial flip model. Students who were above average and already earning higher grades may be more motivated to do this than below average students. The partial flip class would be more rewarding to a student who did the assigned work outside of class. Being able to take away more from the in-class experience, they would have been able to do better on a post-class assignment. Future studies could examine the impact when students have gained experience in using the partial flip model several times throughout the course to see how this affects learning outcomes.

Considerations on the classes used for this study

Finally, we analysed both the control and partial flip sections to determine if the groups were different in other dimensions. One challenge in science education research is the many uncontrolled variables that exist when introducing different pedagogical strategies across groups. In the present study, we utilised the fact that the general biology course included two sections taught by the same instructor. This allowed us to have an experimental and control section of the course (chosen randomly). However, this experimental design carries the *a priori* assumption that the two groups of students were of similar academic aptitude. We found evidence of similarities supporting this assumption as well as differences between the groups. Both the control and partial flip sections scored equally well on exam 3 overall,

allowing the differences in specific questions related to Meselson and Stahl to be interpreted as a result of the partial flip. However, we did find some evidence of differences between the groups that should be taken into account when interpreting these data. The areas of course of study differed between the groups. The control section included ~75% pharmacy students, while the partial flip included ~60% health sciences. While both of these programmes have enhanced admissions criteria compared to general admission at the university, there may be more inherent differences between students going into pharmacy practice versus health sciences with regards to understanding Meselson and Stahl.

The other prominent difference between the sections was their grade on exam 4, where the partial flip group performed better than the control group overall. Since the partial flip experiment was a small component of the overall course, it is likely that other factors contributed to this difference. These factors may include differing interest or ability in the subject between pharmacy and health science students, time of day effects (control section began at 8AM, partial flip section began at 9:30AM) or varying background knowledge. Notably, while pharmacy and health science admissions have nearly identical SAT/GPA requirements, health science students are required to have four years of mathematics/science courses in high school prior to entering the programme, while this requirement does not exist for the pharmacy programme. Despite these caveats, we feel that a case/control comparison provides for more meaningful interpretation of a pedagogical intervention. Many studies would be improved using two sections comprised of similar students, but this is often a challenge. One way to address this in future studies is to collect other academic data on students prior to the start of a trial (eg SAT scores, high school GPAs) to ensure that aptitude is grossly equivalent across sections.

Educational implications

Overall, this study has two important take-home messages. First, while the one-time partial flip does not have a large effect, it does not hurt student performance in any noticeable manner, may help boost self-reported confidence in the material in females and helps to boost the performance of male students in the class in formative assessments. Second, the biggest effects of the partial flip are on immediate formative class assessments compared to longer term summative assessments.

The partial flip technique can easily be implemented since it does not require the complete redesign of a course, can be used even in a large auditorium-style course, does not harm students' performance, and can lead to small benefits to both above and below average students. This implies that the addition of any flipped classroom component to a curriculum is better than none and that the partial flip model could be used as a way to increase the amount of active learning in a classroom. An entire overhaul of a traditional lecture class into a fully flipped class is not necessary to yield some of the benefits. Overall, these results show that the partial flip pedagogy fits in well as a complementary technique to use in a hybrid lecture-style classroom as a transition to using more active learning methods.

Acknowledgements

We acknowledge funding for this study from the American Society for Cell Biology Mentoring in Active Learning and Teaching (MALT) programme (BJK fellow; JM mentor). We also acknowledge Dr Michael Burns and Amy Jarnot (College of Biological Sciences, University of Minnesota) for live training of COPUS scoring, Dr Dan L. Perlman (Department of Biology, Brandeis University) for study design assistance, Dr Laurel Willingham-McLain (Center for Teaching Excellence, Duquesne University) for study design assistance and COPUS scoring assistance and Dr Ian Quitadamo (Central Washington University) for classification of multiple-choice questions based on Bloom's Cognitive levels. We also thank our students for their participation in the study and the scientific process.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the American Society for Cell Biology Mentoring in Active Learning and Teaching (MALT).

References

- Baepler, Paul, J. D. Walker, and Michelle Driessen. 2014. "It's Not about Seat Time: Blending, Flipping, and Efficiency in Active Learning Classrooms." *Computers & Education* 78: 227–236.
- Bergmann, Jonathan, and Aaron Sams. 2012. *Flip Your Classroom: Reach Every Student in Every Class Every Day*. 1st ed. Arlington, VA: International Society for Technology in Education.
- Berrett, Dan. 2012. "How 'Flipping' the Classroom Can Improve the Traditional Lecture." *The Chronicle of Higher Education* 12: 1–14.
- Bevan, S. J., C. W. Chan, and J. A. Tanner. 2014. "Diverse Assessment and Active Student Engagement Sustain Deep Learning: A Comparative Study of Outcomes in Two Parallel Introductory Biochemistry Courses." *Biochemistry and Molecular Biology Education* 42 (6): 474–479. doi:10.1002/bmb.20824.
- Bloom B. S., M. D. Engelhart, E. J. Furst, W. H. Hill, and D. R. Krathwohl. 1956. *Taxonomy of Educational Objectives, Handbook 1: Cognitive Domain*. Edited by B. S. Bloom. New York: David McKay Company.
- Burkam, David T., Valerie E. Lee, and Becky A. Smerdon. 1997. "Gender and Science Learning Early in High School: Subject Matter and Laboratory Experiences." *American Educational Research Journal* 34 (2): 297–331.
- Dangel, Harry L, and Wang Charles Xiaoxue. 2008. "Student Response Systems in Higher Education: Moving beyond Linear Teaching and Surface Learning." *Journal of Educational Technology Development and Exchange* 1 (1): 93–104.
- Dolmans, D. H., S. M. Loyens, H. Marq, and D. Gijbels. 2015. "Deep and Surface Learning in Problem-based Learning: A Review of the Literature." *Advances in Health Sciences Education: Theory and Practice* Nov 13: 1–26. doi:10.1007/s10459-015-9645-6.
- Frederickson, Norah, Phil Reed, and Viv Clifford. 2005. "Evaluating Web-supported Learning versus Lecture-based Teaching: Quantitative and Qualitative Perspectives." *Higher Education* 50 (4): 645–664.
- Freeman, S., S. L. Eddy, M. McDonough, M. K. Smith, N. Okoroafor, H. Jordt, and M. P. Wenderoth. 2014. "Active Learning Increases Student Performance in Science, Engineering, and Mathematics." *Proceedings of the National Academy of Sciences* 111 (23): 8410–8415. doi:10.1073/pnas.1319030111.
- Fulton, Kathleen. 2012. "The Flipped Classroom: Transforming Education at Byron High School: A Minnesota High School with Severe Budget Constraints Enlisted YouTube in Its Successful Effort to Boost Math Competency Scores." *THE Journal (Technological Horizons in Education)* 39 (3): 18.
- Green, G. 2012. "The Flipped Classroom and School Approach: Clintondale High School." Annual Building Learning Communities Education Conference, Boston, MA. Available from <http://2012.blcconference.com/documents/flipped-classroom-school-approach.pdf>.
- Hamdan, N., P. McKnight, K. McKnight, and K. M. Arfstrom. 2013. *A Review of Flipped Learning*. Flipped Learning Network. Pearson Education, and George Mason University.
- Handelsman, Jo, Diane Ebert-May, Robert Beichner, and Peter Bruns. 2004. "Scientific Teaching." *Science* 304 (5670): 521–522.
- Holmes, Frederic Lawrence. 2001. *Meselson, Stahl, and the Replication of DNA: A History of "The Most Beautiful Experiment in Biology"*. London: Yale University Press.
- Huang, A. H., and R. G. Carroll. 1997. "Incorporating Active Learning into a Traditional Curriculum." *Am J Physiol* 273 (6 Pt 3): S14–S23.
- Huba, Mary E., and Jann E. Freed. 2000. "Learner Centered Assessment on College Campuses: Shifting the Focus from Teaching to Learning." *Community College Journal of Research and Practice* 24 (9): 759–766.
- Johnson, Larry, Samantha Adams Becker, Victoria Estrada, and Alex Freeman. 2015. *NMC Horizon Report: 2015 Higher Education Edition*. Austin, TX: The New Media Consortium.
- Klegeris, A., M. Bahniwal, and H. Hurren. 2013. "Improvement in Generic Problem-solving Abilities of Students by Use of Tutor-less Problem-based Learning in a Large Classroom Setting." *CBE Life Sciences Education* 12 (1): 73–79. doi:10.1187/cbe.12-06-0081.
- Linn, Marcia C., and Janet S. Hyde. 1989. "Gender, Mathematics, and Science." *Educational Researcher* 18 (8): 17–27.
- Margulieux, Lauren E., Keith R. Bujak, W. Michael McCracken, and D. M. Majerich. 2014. "Hybrid, Blended, Flipped, and Inverted: Defining Terms in a Two Dimensional Taxonomy." Paper accepted to the 12th Annual Hawaii International Conference on Education, Honolulu, January.
- McFarlin, Brian K. 2008. "Hybrid Lecture-online Format Increases Student Grades in an Undergraduate Exercise Physiology Course at a Large Urban University." *Advances in Physiology Education* 32 (1): 86–91.
- McLean, S., S. M. Attardi, L. Faden, and M. Goldszmidt. 2016. "Flipped Classrooms and Student Learning: Not Just Surface Gains." *Advances in Physiology Education* 40 (1): 47–55. doi:10.1152/advan.00098.2015.
- Meselson, M., and F. W. Stahl. 1958. "The Replication of DNA in *Escherichia coli*." *Proceedings of the National Academy of Sciences* 44 (7): 671–682.

- Michael, Joel. 2006. "Where's the Evidence That Active Learning Works?" *Advances in Physiology Education* 30 (4): 159–167.
- Morris, James, Daniel L. Hartl, Andrew H. Knoll, and Robert Lue. 2013. *Biology : How Life Works*. New York: W.H. Freeman and Company, a Macmillan Higher Education company.
- Mortensen, C. J., and A. M. Nicholson. 2015. "The Flipped Classroom Stimulates Greater Learning and is a Modern 21st Century Approach to Teaching Today's Undergraduates." *Journal of Animal Science* 93 (7): 3722–3731. doi:10.2527/jas.2015-9087.
- O'Flaherty, Jacqueline, and Craig Phillips. 2015. "The Use of Flipped Classrooms in Higher Education: A Scoping Review." *The Internet and Higher Education* 25: 85–95.
- Pelaez, N. J. 2002. "Problem-based Writing With Peer Review Improves Academic Performance in Physiology." *Advances in Physiology Education* 26 (3): 174–184.
- Peterson, Penelope L., and Elizabeth Fennema. 1985. "Effective Teaching, Student Engagement in Classroom Activities, and Sex-related Differences in Learning Mathematics." *American Educational Research Journal* 22 (3): 309–335.
- Plendl, J., M. Bahramsoltani, O. Gemeinhardt, H. Hunigen, S. Kassmeyer, and P. Janczyk. 2009. "Active Participation instead of Passive Behaviour Opens up New Vistas in Education of Veterinary Anatomy and Histology." *Anatomia, Histologia, Embryologia* 38 (5): 355–360. doi:10.1111/j.1439-0264.2009.00954.x.
- Pomerantz, Eva M., Ellen Rydell Altermatt, and Jill L. Saxon. 2002. "Making the Grade but Feeling Distressed: Gender Differences in Academic Performance and Internal Distress." *Journal of Educational Psychology* 94 (2): 396.
- Prober, Charles G., and Salman Khan. 2013. "Medical Education Reimagined: A Call to Action." *Academic Medicine* 88 (10): 1407–1410.
- Riffell, Samuel, and John Merrill. 2005. "Do Hybrid Lecture Formats Influence Laboratory Performance in Large, Pre-professional Biology Courses?" *Journal of Natural Resources and Life Sciences Education* 34: 96.
- Riffell, Samuel K., and Duncan F. Sibley. 2004. "Can Hybrid Course Formats Increase Attendance in Undergraduate Environmental Science Courses?" *Journal of Natural Resources and Life Sciences Education* 33: 16.
- Smith, M. K., F. H. Jones, S. L. Gilbert, and C. E. Wieman. 2013. "The Classroom Observation Protocol for Undergraduate STEM (COPUS): A New Instrument to Characterize University STEM Classroom Practices." *CBE Life Sciences Education* 12 (4): 618–627. doi:10.1187/cbe.13-08-0154.
- Spencer, John A., and Reg K. Jordan. 1999. "Learner Centred Approaches in Medical Education." *BMJ* 318 (7193): 1280.
- Strayer, Jeremy F. 2012. "How Learning in an Inverted Classroom Influences Cooperation, Innovation and Task Orientation." *Learning Environments Research* 15 (2): 171–193.
- Strenta, A. Christopher, Rogers Elliott, Russell Adair, Michael Matier, and Jannah Scott. 1994. "Choosing and Leaving Science in Highly Selective Institutions." *Research in Higher Education* 35 (5): 513–547.
- Suskie, Linda. 2010. *Assessing Student Learning: A Common Sense Guide*. Oxford: Wiley.
- van Vliet, E. A., J. C. Winnips, and N. Brouwer. 2015. "Flipped-class Pedagogy Enhances Student Metacognition and Collaborative-learning Strategies in Higher Education but Effect Does Not Persist." *CBE-Life Sciences Education* 14 (3): ar26.
- Warter-Perez, Nancy, and Jianyu Dong. 2012. "Flipping the Classroom: How to Embed Inquiry and Design Projects into a Digital Engineering Lecture." Proceedings of the 2012 ASEE PSW Section Conference.