Learning and Instruction 45 (2016) 61–71

Contents lists available at ScienceDirect

Learning and Instruction

journal homepage: www.elsevier.com/locate/learninstruc

The effects of flipped instruction on out-of-class study time, exam performance, and student perceptions

Wenliang He [a,*], Amanda Holton [b], George Farkas [a], Mark Warschauer [a]

[a] School of Education, 3200 Education Bldg, University of California, Irvine, CA 92697, USA
[b] Department of Chemistry, 580C Rowland Hall, University of California, Irvine, CA 92697, USA

ARTICLE INFO

Article history:
Received 11 November 2015
Received in revised form
21 June 2016
Accepted 8 July 2016

Keywords:
Flipped instruction
Chemistry
Study time
Exam performance
Perception

ABSTRACT

This study assessed the impact of flipped instruction on students' out-of-class study time, exam performance, preference, motivation, and perceptions in two sections of a large undergraduate chemistry course. Flipped instruction caused a shift in student workload without appreciably changing the overall study time. The treatment impact on student performance gradually diminished over time, showing a small but statistically significant effect with the final exam. No marked interaction was identified, indicating that flipped instruction benefited students of diverse backgrounds uniformly. Students in the flipped section showed mixed feelings with about one fifth of them displaying polarized attitudes. Open-ended student survey responses revealed non-compliance with pre-class studying as a serious implementation issue: By slowing down the overall pace of the class, it negatively affected students with different study behaviors and characteristics in ways that partly explained the small, diminishing treatment effect and absence of marked interaction.

1. Introduction

In recent years, flipped instruction, also known as inverted instruction, has attracted growing attention from both teaching and research communities as a promising instructional technique. Before 2013, we have only identified a small handful of empirical studies in higher education addressing this topic (see a brief review below). The past two years have seen a surge in the number of both case studies and empirical research, as universities worldwide are adopting the flipped pedagogy. The excitement over flipped instruction stems from its capacity to combine purported benefits of online instruction with active learning techniques into a new pedagogy. The essence of the flipped pedagogy is to stage learning of new material before class to free up class time for more in-depth explanation, practice, and productive use of knowledge.

Researchers have not yet reached a consensus on a formal definition of flipped instruction. For early pioneers (Bergmann & Sams, 2008), the intent of flipped instruction was to eliminate lectures in class (i.e., students could watch video podcasts at home before class) so that material that had traditionally been assigned as homework could be completed in class with more student-centered and inquiry-based activities. This idea has led to the general conception of flipped instruction as "events that have traditionally taken place inside the classroom now take place outside the classroom and vice versa" (Lage, Platt, & Treglia, 2000). Bishop and Verleger (2013) suggested including only the studies with computer-based pre-class instruction. We believe qualifying instructional medium is unnecessary and unjustified. It had been shown in a quasi-experimental study with over 800 students that pre-class reading assignments supplemented with worksheets could be as effective as pre-class videos in increasing exam performance using the flipped pedagogy (Moravec, Williams, Aguilar-Roca, & O'Dowd, 2010).

Based on the discussion above, in this study, we define flipped instruction as having three attributes. Flipped classrooms should feature (a) mandatory pre-class learning of new material followed by (b) in-depth explanation, practice, and productive use of knowledge in class through active learning techniques, where (c) class attendance is mandatory. All three features are necessary. First, pre-class learning is an integral part of instruction. Long before flipped instruction was studied as a distinct pedagogy, instructors were known to assign textbook material for students to read before class. In traditional classrooms, however, pre-class learning was often not enforced and instructors would cover the

* Corresponding author.
E-mail address: wenlianh@uci.edu (W. He).

http://dx.doi.org/10.1016/j.learninstruc.2016.07.001
0959-4752/

pre-assigned material in class anyway. In a flipped classroom, pre-class instruction is designated for teaching (factual) knowledge that will not be repeated in class except for brief reviews. Secondly, productive use of knowledge should dominate class time in order to promote conceptual understanding. Traditional classrooms also adopt active learning techniques. Due to a packed schedule, however, active learning often accounts for a small proportion of the class time. In a flipped classroom, since lectures are, to a large extent (if not entirely) replaced by active learning activities, class time is largely reserved for productive use of knowledge. Finally, class attendance must be mandatory. In-class instruction is geared towards promoting conceptual understanding, which is a crucially important aspect of learning. Therefore, a "flipped" classroom that adopts an optional attendance policy is not genuinely flipped. It resembles an online class more than a flipped class, since the instruction is already offered online and hence a student can afford not to attend class. In other words, pre-class study and in-class activities should be integrated. They complement each other and are integral parts of learning as a whole.

Over the years, researchers have proposed several benefits of flipped instruction. Since many flipped classrooms use online videos to stage instruction before class, students can learn the material anytime, anywhere, and at their own pace (McDonald & Smith, 2013), which is a desired quality that has been attributed to online instruction (Bourne, Harris, & Mayadas, 2005). In addition, some studies have shown that pre-class learning reduced perceived difficulty of a course, while raising perceived instructional clarity and the overall course value (Narloch, Garbin, & Turnage, 2006; Stelzer, Gladding, Mestre, & Brookes, 2009). It is conceivable that by offloading some learning material to pre-class, flipped instruction effectively prepares students to engage with more challenging content during class time. Moreover, case studies have shown that flipped instruction increased teacher-student interaction during class and that students were more actively involved in the learning process in flipped classes (Herreid & Schiller, 2013).

Despite the enthusiasm, many of the studies hitherto published on flipped instruction were conducted in a single classroom with no reference to a comparison group and the purported benefits of flipped instruction have thus largely come from cases studies, descriptive surveys, theorized conjectures, and anecdotal evidence. Empirical studies employing validated controls with reasonably large sample sizes that directly assess the effectiveness of flipped instruction are sorely needed. Our study aims to address this issue. To begin with, we will review relevant research assessing the treatment effect of flipped instruction on student exam performance. Based on the review, our research questions are proposed.

1.1. Prior research on treatment effect

Several studies have reported surprisingly large effect sizes of flipped instruction. Deslauriers, Schelew, and Wieman (2011) compared two sections of a first-year physics course taken by 538 undergraduate engineering students at University of British Columbia (UBC). Two experienced professors, one for each section, taught with traditional lectures in a similar manner. Flipped instruction was implemented only in the twelfth week of the semester, when a postdoctoral researcher with limited teaching experience replaced one professor. Students' pretest measures were practically identical. A twelve-item posttest was administered and results showed that the average scores were 41% ($SD = 13\%$, $N = 171$) in the control section and 74% ($SD = 13\%$, $N = 211$) in treatment, which gives a staggering effect size of about 2.5 standard deviations. Moreover, during the twelfth week, class attendance increased by 20% and in-class student engagement nearly doubled in the experimental section, whereas both measures remained unchanged in the control. These results are interesting in that they imply that even "novice" flipped instructors could outperform their more experienced traditional counterparts in raising student performance, class attendance and engagement.

Such encouraging results, however, should be viewed with caution. To start with, the authors themselves cautioned that the immediate end-of-treatment posttest primarily reflected the result of learning achieved from pre-class study and the class itself. Other studies with smaller effect sizes often measured student performance with end-of-term final exams that reflected all the learning done inside and outside of the class. Since learning is a multifaceted process, the impact of a specific learning channel is likely to be diluted when more opportunities are open for students to acquire knowledge. Another contributing factor could be a sense of novelty induced by the presence of a young and energetic instructor with a distinctly different teaching style. If the flipped pedagogy had been assessed in a more authentic setting (i.e., same instructor throughout, class flipped over an entire semester, and performance measured using end-of-term high-stakes final exams), it is unclear if the same positive results would have been achieved.

Partial evidence for the above conjecture was in fact provided by research from the same authors. In another study (Deslauriers & Wieman, 2011), flipped instruction was implemented for the entire duration of an eleven-week summer course with two cohorts of students at UBC. A superb, award-winning lecturer taught a traditional classroom in 2008. The following summer, an unspecified instructor taught the same course using the flipped pedagogy. In an ungraded mock-up exam given one week prior to the final, 62 treatment and 48 control students averaged 85% ($SD = 14\%$) and 67% ($SD = 18\%$) respectively, which corresponds to an effect size of 1.14 standard deviations, still quite large but less than half of the previous estimate.

Furthermore, two studies involving the same instructors teaching in authentic settings showed more moderate impact. Day and Foley (2006) assessed flipped instruction using two sections of the same course with 18 control and 28 treatment students. The reported statistics suggest an estimated effect size of 0.69 standard deviations with the final exam ($p = 0.055$). In another study, Mason, Shuman, and Cook (2013) assessed flipped instruction in a senior-level mechanical engineering course with 40 students in two consecutive cohorts (i.e., 20 students in each). Seventeen problem pairs from quizzes, midterms and finals were matched and used to assess performance. The overall effect size was 0.75 standard deviations. Although an effect size of 0.6–0.8 is still quite large (Cohen, 2013), the limited sample sizes of the two studies make their results somewhat less convincing.

In addition, existing studies have not adequately explored possible heterogeneity of treatment effect on student performance. Mason et al. (2013) looked into the differentiated impact of flipped instruction by question type and found that their flipped classrooms were about twice as effective in improving student performance on design-based problems compared to non-design-based ones. Not enough research has been conducted examining the possibly differentiated impact of flipped treatment by student demographics, such as gender, ethnicity, and prior grades.

Finally, most studies have not carefully equated the overall instruction time and/or measured the out-of-class study time. So far, only two studies (Day & Foley, 2006; Street, Gilliland, McNeil, & Royal, 2015) have explicitly controlled for instruction time by reducing the number of class meetings commensurate with the added hours of pre-class instruction. Although directly manipulating class time might encounter administrative difficulties, researchers should measure and compare out-of-class study effort