CrossMark

Investigating Strategies for Pre-Class Content Learning in a Flipped Classroom

Jamie L. Jensen[1] · Emily A. Holt[2] · Jacob B. Sowards[1] · T. Heath Ogden[3] · Richard E. West[4]

Abstract

In a flipped classroom model, learning of basic content is shifted before class while in-class time is used for concept application. Empirical and controlled research studies are lacking on the best strategies to provide the necessary pre-class content instruction. In this study, we tested three methods of pre-class content learning—interactive online tutorials, video lectures, and textbook-style readings—while holding the content and the in-class application activities constant. Identical introductory, non-majors biology classes were manipulated at both a public, open-enrollment institution and a private, highly selective institution. We found that video lectures offer a small advantage to overall student learning over interactive tutorials or textbook-style readings. Although our two populations differed in their ability to effectively learn from pre-class activities, through a student-centered flipped classroom approach, students at both institutions demonstrated equal learning gains by the final assessment. Potential reasons for some observed differences are suggested.

Keywords Flipped classroom · Undergraduate · Active learning · Content learning

Introduction

It is well known that active learning can reduce failure rates and increase student performance (Freeman et al. 2014), but

✉ Jamie L. Jensen
Jamie.Jensen@byu.edu

Emily A. Holt
Emily.Holt@unco.edu

Jacob B. Sowards
jbsowards15@gmail.com

T. Heath Ogden
Heath.Ogden@uvu.edu

Richard E. West
rickwest@byu.edu

[1] Department of Biology, Brigham Young University, 4102 LSB, Provo, UT 84602, USA

[2] School of Biological Sciences, University of Northern Colorado, Campus Box 92, Greeley, CO 80639, USA

[3] Department of Biology, Utah Valley University, SB 242M, Orem, UT 84058, USA

[4] Department of Instructional Psychology and Technology, Brigham Young University, 150-H MCKB, Provo, UT 84602, USA

many teachers struggle to find the time to incorporate active learning in the classroom. One potential strategy that has become popular is the "flipped classroom" model, where students learn basic content before class in the form of instructional videos, recorded lectures, readings, etc. (Hamdan et al. 2013). Then, instructors spend time in class applying the material through complex problem solving, deeper conceptual coverage, and peer interaction (Strayer 2012; Tucker 2012; Gajjar 2013; Sarawagi 2013). Instructors use this model to improve cognitive load management of their students (Abeysekera and Dawson 2015), encourage greater independence of their students, and free class time for active strategies (Seery 2015).

Research has shown a generally positive effect of flipped classrooms on student attitudes and performance compared to a passive lecture model (Bergstrom 2011; Strayer 2012; Tune et al. 2013; O'Flaherty and Philips 2015; Seery 2015). However, while there is much research on how to use the recaptured class time for more effective active (Roehl et al. 2013; Baepler et al. 2014; Hung 2015) and student-centered learning (Kim et al. 2014; McLaughlin et al. 2014), there is very little research on how best to provide the necessary content instruction at home. For flipped teaching to be successful, teachers must have confidence that their students learned the necessary information and skills for the active learning

activities upon arrival in the classroom. This lack of clarity about the optimal out-of-class content instruction may be one reason why teachers are hesitant to adopt flipped teaching.

In this paper, we report our findings about three different methods for providing at home content learning (interactive tutorials, video lectures, and textbook-style readings), in support of flipped learning experiences. We also contextualize our findings according to various learning theories to better understand the pros and cons of these various approaches.

Literature Review

To provide context for this study, we first review the research on flipped classrooms in general, and then focus on strategies for out-of-class content learning, including learning theories that can support these various strategies.

The Evidence on Flipped Classrooms

There has been much recent excitement over the concept of flipping classrooms to emphasize more mentored, active learning during class periods. Flipped classroom instruction is defined simply as an instructional strategy where students learn content prior to class, allowing them to come to class prepared and ready for mentored, active, and experiential learning experiences (Hamdan et al. 2013). Abeysekera and Dawson's (2015) "lowest common denominator" definition of a flipped classroom suggests that the strategy must include three key components: (1) information-transmission outside of class time, (2) class time dedicated to active, collaborative activities, and (3) student accountability for in-class activities through pre- or post-class activities. This first piece, information-transmission, is almost always in the form of a pre-recorded lecture or screencast (Pierce and Fox 2012; O'Flaherty and Philips 2015; Seery 2015; Zainuddin and Halili 2016). Other methods of pre-class content dissemination include readings, Blogs, Google Docs, Google Hangouts, and some interactive computer software programs such as MyITLab (Davies et al. 2013) or Integrated Learning Accelerator Modules (McLaughlin et al. 2014).

While research into flipped classroom instruction, as specifically defined, is recent, many positive effects have been found, particularly for science education (Fautch 2015; Tomory and Watson 2015) and students seem to like the approach overall (Phillips and Trainor 2014; Jeong and González-Gómez 2016), often preferring it over traditional learning (Gilboy et al. 2015). Most of this research has been in the K-12 realm (Leo and Puzio 2016; Olakanmi 2017), but other studies have been conducted in higher education. In one example, Long et al. (2017) conducted a qualitative case study analysis of instructor experiences using this method and found that instructors perceived that

their students enjoyed this approach to learning. Indeed, for some, the motivation to flip their classrooms was because of student dissatisfaction with traditional lecture approaches. The instructors also felt the method provided more active learning and better student support, but a major challenge was that not all students learned the content on their own to a sufficient degree before class in order to make the in-class activity effective.

Schwarzenberg et al. (2017) conducted a quantitative study on flipped classroom effects and found similar modest results. They found slightly higher achievement in the flipped classroom experience compared to conventional classes, but the design of the flipped experience was important. In particular, Schwarzenberg et al. (2017) noted that in-class activities should focus on active learning and that the level of pre-class preparation from students was important—echoing the concern raised by Long et al. (2017).

Adding further evidence, Davies et al. (2013) found that a flipped classroom was more effective in teaching concepts than a simulation-based approach or traditional classroom approach for introductory information systems students; Zainuddin and Halili (2016) found in their meta-synthesis of 20 articles that flipped classrooms overall brought positive impacts on learning and motivation. However, some scholars have found more mixed results. Harrington et al. (2015) found no significant differences in the learning outcomes for nursing students randomly assigned to either a flipped or traditional teaching style; Jensen et al. (2015a) found similar results for introductory biology students between a flipped and non-flipped classroom. A systematic review of 21 nursing studies Betihavas et al. (2016) similarly found overall themes of neutral or positive academic outcomes and mixed results for satisfaction. These mixed results can lead to conclusions that flipped classroom can be an effective strategy but that the real difference may be in how flipped classrooms are implemented.

Indeed, in most of these studies, a key finding has been that the benefits depend on the ability of students to come to class well prepared in having learned the content and thus ready for the in-class active learning. The difficulty in preparing students to learn the content before class is amplified because some students may lack the self-regulatory skills to learn content effectively on their own. Sletten (2017) asked 76 students in a flipped introductory biology class about their self-regulated learning strategy use and found that students largely preferred the in-class active learning but did not think highly of the out-of-class video lectures. Sletten (2017) suggested that students may not be well prepared to learn independently via the out-of-class lecture videos. However, Porcaro et al. (2016) similarly used screencast-recorded video lectures as the pre-

class content learning in two iterations of a hematology course but found the vast majority (89 and 93%) of students in their study to be well prepared for class. This may indicate, of course, that different classes of students just respond differently to different video lectures, but it also suggests that more research is needed into how, when, and why students engage deeply with the out-of-class content learning critical to flipped learning success.

Variety of Strategies for Out-of-Class Content Learning

As mentioned above, many studies have noted that there is great variance in how well students learn content out of class, which is crucial to the overall success of the flipped classroom model. However, surprisingly few studies have discussed the various methods for teaching this out-of-class content. Some studies have shown benefits of video tutorials (e.g., He et al. 2012; Kay and Kletskin 2012), and indeed video tutorials or lectures seem to be the most common method. However, only one systematic, controlled comparison has been made between various pre-class content delivery methods to determine which is most effective for student learning (Moravec et al. 2010), and in this study, the instructors only implemented three flipped class sessions in an otherwise non-flipped semester to investigate performance differences between students taught outside of class via video lectures versus readings.

Another strategy for out-of-classroom instruction is to utilize interactive tutorials. This strategy is based in Constructivist theory (Dewey 1938; O'Donnell et al. 2006)—the idea that students are not empty buckets to be filled with knowledge but rather that students must construct knowledge for themselves (Piaget 1985; Lawson 2002). This construction of knowledge is a process of dynamic equilibration or interaction between an individual and their environment, where innate mental structures are reorganized as gaps and contradictions are recognized (O'Donnell et al. 2006).

Allowing students to work through the information in interactive tutorials, rather than telling them the information, can influence conceptual development by posing critical cognitive conflict that disturbs equilibrium and forces the individual to restructure their cognitive architecture (Pulaski 1980; Damon 1984; Doise and Mugny 1984; Kubli 1989; Lumpe 1995). However, students quite often do not recognize the gaps or contradictions in their knowledge, or they may recognize them but choose not to act upon them (O'Donnell et al. 2006). This recognition and call to action can occur by designing materials to promote student construction of knowledge, which forces students to search for equilibration and drive cognitive development (Pulaski 1980).

Alternatively, utilizing video lectures allows instructors to tap into two modes of information processing, both visual and auditory. Dual coding theory (Paivio 1990) provides a theoretical framework for how differing content (video or readings) is processed by the student. According to this theory, the more sensory pathways that a student can use to interact with the material, the more likely they are to remember the content (Clark and Paivio 1991). By laying down two memory traces to the information, verbal and image, the information is more accessible to the learner (see Thomas 2014 for a review). Yadav et al. (2011) suggested that video may be a more powerful medium for cognitive and affective processing compared to text reading alone, because auditory and visual information are redundant bisensory stimuli that collectively contribute to learning (Moreno and Mayer 2002).

In addition, motivational theory (Keller 1983) may also explain the success of videos. In Keller's theory, motivational factors include relevance, attention, confidence, and satisfaction; the latter three factors are especially applicable for video lecture success. To gain attention and satisfaction, a video lecture may include quick and concise attractive images or animation which can serve as entertainment (Keller 2009). Further, learning the material via online video lectures can boost a student's confidence by using a medium of delivery to which they may be accustomed in their personal lives—online videos. Using this model, students may be more motivated to learn if their education includes the attention-grabbing technology with which they are familiar in everyday life, potentially leading to greater learning gains.

Lastly, textbook-style readings are frequently used by instructors to offer a short, easy-to-skim, straightforward delivery of low-level content. Textbooks are the backbone of almost every college course, they are easier to read than primary literature, and they are perceived by both instructors and students as an integral part of the learning experience (see Besser et al. 1999). Not only are textbook readings easily searchable (as opposed to a video lecture), but because of their brevity, it may be possible that students can easily access them repeated times getting more exposure to the material. In fact, textbooks are often the source of the majority of student studying (Besser et al. 1999). Textbooks may in fact be the modality to which students have been most exposed, having been a fundamental aspect of education for centuries. However, this treatment, in the flipped format, differs from the traditional classroom where reading is assigned before class and student responsibility is not enforced because the instructor lectures on the same content included in the pre-class reading (He et al. 2016). In traditional non-flipped courses, students often do not read the assignment (e.g., Sikorski et al. 2002; Clump et al. 2004), especially weaker students (Phillips and Phillips 2007). In a flipped classroom, however, student accountability for completing the

reading before class is often built in through pre-class reading assignments.

All three of the above-described content learning methods, situated before active face-to-face class time, adhere to the definition of a "lowest common denominator" flipped class. In this study, we compare these three strategies for out-of-classroom content learning in a flipped classroom approach while keeping in-class application activities the same. Our research question is, *what is the differential effectiveness of these strategies*? Answering this question about pre-class content learning can address an important piece of the flipped model, allowing for practical recommendations for the best method for implementing a flipped classroom.

Methods

Subjects

To investigate the mechanisms underlying the effectiveness of content learning methods during the "at home" content attainment phase of a flipped classroom, we collected data from 657 undergraduate students enrolled in non-science majors general biology courses at two large institutions in the Western USA. One institution is a large private university (~ 30,000 students) with highly competitive admission criteria reflected by incoming freshman average high school grade point average (GPA) of 3.8 and an average American College Test (ACT) score of 28.3 in 2014. It will hereafter be referred to as the "private institution." The second institution is a large public university (~ 32,000 students) with open-enrollment and an incoming freshman average high school GPA of 3.27 and an average ACT score of 23 (collected from a random sample of self-reported data from 250 students in an introductory biology course). It will hereafter be referred to as the "public institution."

Three instructors participated in the data collection, one at the private institution and two at the public institution (Table 1). Due to unbalanced sample sizes (Table 1), redundant treatment sections within the same public institution, but taught by different instructors, were combined. The decision to pool these data was supported by a lack of difference between instructors at the public institution within the common treatment for each dependent variable (Explore Assessments,

$p = 0.34$; Apply Assessments, $p = 0.46$; Final Exam, $p = 0.44$). Instructor 2 conducted the Video Lecture treatment (Table 1); however, we opted not to include these data in our analysis. The decision to exclude this class was made prior to completion of the semester (i.e., before summarizing the data) after three independent instructors interacted with the class and determined that it was anomalous in its lack of engagement and motivation; thus, we anticipated that any notable difference may be attributed to this larger class personality rather than a treatment effect. Students enrolled in the course at both institutions are generally non-science majors and range from freshman to seniors.

Quasi-Experimental Design

We used three treatments or modalities of pre-class content teaching. Each treatment represented a flipped classroom but the method of learning of material before class (i.e., at home) varied. The treatments, replicated by each instructor (Table 1), are described below.

While the strategy for teaching pre-class material differed depending on the treatment, all content and in-class activities were identical among all treatments. Class time was dedicated to concept application activities in peer groups to apply the material explored in the pre-class activities. Pre-class material was not re-presented during class to encourage students to rely upon their online assignment for course preparation.

Curriculum

To assess which method for learning out-of-class content was most effective in our study classrooms, we designed three contrasting methods for the "at home" portion of instruction: interactive tutorials, video lectures, and non-interactive, textbook-style readings (Fig. 1). Identical curricular materials were developed collectively by the three instructors and used at both institutions. In all curricular materials, we used a learning cycle (Bybee 1993), constructivist approach utilizing the 5 Es: (1) Engage and (2) Explore, where students are initially given the opportunity to wrestle with the concepts and build their own understanding; (3) Explain, during which students are introduced to the terms for which they have built meaning; (4) Elaborate, where students apply the content to new

Table 1 Experimental setup

Instructor	Institution	Interactive tutorials	Video lectures	Textbook readings
1	Private	$N = 200$	$N = 113$	$N = 206$
2	Public	$N = 36$	–	$N = 29$
3	Public	$N = 23$	$N = 22$	–

The experiment was run with three instructors at two institutions (private and public) using three different methods of content delivery. Dashes indicate that the treatment was not run by that particular instructor or excluded for reasons explained in the text

Fig. 1 Three treatment conditions on sex-linked inheritance: Interactive Tutorials posed questions and solicited feedback from students; Video Lectures consisted of the same material but was presented by one of the instructors in a video format; and Textbook-style Readings consisted of the same material but was written in the form of a textbook passage; the latter two treatments required no interaction by students

situations; and (5) Evaluate, during which students participate in both formative and summative assessments. In our design, Engage, Explore, and Explain occurred during pre-class activities at home, and we collectively referred to as "Explore" activities. Additional Explanation and Elaboration activities took place in class and are collectively referred to as "Apply" activities. Evaluation occurred both directly following every pre-class activity at home (41 Explore Assessments) and at the end of each unit, roughly every 2 weeks (eight Apply Assessments).

Content quality and quantity of pre-class activities was controlled by using essentially identical scripts and examples to focus on the method of learning as well as the constructivist active nature among the three treatments. Certainly, the interactive tutorials were more active than the other two treatments; however, the constructivist nature (i.e., exposing students to a puzzling phenomenon and reasoning through it before introducing terminology and explanations) was conserved in all three treatments. Students were assigned to treatments based on their section enrollment utilizing a quasi-experimental design.

Students completed all pre-class assignments online, in one of the three modalities, utilizing content designed with online survey software (Qualtrics© 2015, Provo, UT) linked through their learning management system (LMS). Completion of

these assignments was incentivized with course credit. Following each pre-class assignment, all students took an identical Explore Assessment that covered material introduced in the pre-class activities, irrespective of treatment. These were short, online quizzes written at mostly low levels of Bloom's Taxonomy of Learning (87% Remember and Understand). We asked students to complete a first attempt of these Explore Assessments without the aid of any notes or information, telling them that they should use it as a way of testing their own knowledge. Following the first attempt, students were given unlimited open-note attempts to achieve 100%, if they so desired. The number of attempts was recorded for analysis.

Items on these assessments were assigned to levels of Bloom's Taxonomy (Bloom 1984) by two independent raters trained in Bloom's and familiar with the curriculum. "Remember" and "Understand" items were categorized as "Low-Level" questions; "Apply" items and above were categorized as "High-Level" items. Raters independently rated the items, discussed differences in ratings, and came to agreement such that inter-rater reliability was 96.4%.

Roughly every 2 weeks, Apply Assessments were administered online through the LMS. These eight Apply Assessments included questions at a variety of Bloom's levels with 65.8% being Apply or above. Apply Assessments were

open-note, but students only had a single attempt and a limit of 60 min to complete it. Apply Assessments were not comprehensive but focused on the material presented since the previous Apply Assessment. A single, comprehensive final assessment was administered either in class or in a proctored testing center with a time limit of 2 h and was closed-note. However, students were allowed one page of notes (8.5 in. × 11 in. sheet of paper) that they personally constructed and upon which they were allowed to put any information they deemed necessary for the exam. The final assessment contained 63% high-level questions and 37% low-level questions.

While the assessments, in-class activities, and pre-class content were identical in all classes, the teaching method of pre-class content differed by treatment. The three modalities are described below. Excerpts from each treatment for one representative topic are shown in Fig. 1.

Interactive Tutorials

These online activities were designed to be active and constructivist in nature (Piaget 1985). Frequent, embedded questions required students to make hypotheses, analyze data, draw conclusions, and make connections. The online system prevented students from advancing forward in the assignment until they answered each question. Assignments were graded for completion and general accuracy, to ensure that students were actively and meaningfully participating in the activity. Periodic short video clips or readings were included that were identical to the other two treatments, yet students in this treatment often needed to make predictions, collect data, and draw conclusions on their own.

Video Lectures

As is more traditional for the flipped classroom (e.g., Pierce and Fox 2012; O'Flaherty and Philips 2015; Seery 2015; Zainuddin and Halili 2016), this condition consisted of video lectures recorded of two of the participating instructors, following an essentially identical script as the interactive tutorials and textbook-style readings. In this case, however, the students watched a video of the instructor talking over a slideshow and visuals depicting the same information as the other two treatments. Total video run-time ranged from 15 to 45 min, and the longer videos were divided into multiple shorter videos. Students were awarded credit for merely watching the video (as measured by viewing the page with the weblink to the video file) and were not required to interact with any of the material.

Textbook-Style Readings

This condition was designed to introduce the same material but in a much less interactive fashion. Patterned to resemble a

textbook passage, students read the material without interacting with it (i.e., no input or answers were requested from students during the assignment). Students were awarded credit for viewing all pages of the reading assignment. The material was presented in the same order and the same questions were posed as in the other treatments. However, unlike the interactive tutorials, answers to all questions were immediately provided as part of the assignment, rather than requiring students to think and answer the questions on their own. While this treatment did include some images or short videos, visuals were considerably fewer in this method than the other two.

Measures

Covariate

To test for group equivalence and control for potential differences that may exist due to our quasi-experimental design, Lawson's Classroom Test of Scientific Reasoning (LCTSR; Lawson 1978, ver. 2000) was used as a covariate in all treatments. The LCTSR is a content-independent assessment of scientific reasoning ability that has been positively related to performance in biology courses (Johnson and Lawson 1998; Lawson et al. 2000b, 2007). The LCTSR was graded on a 24-point scale. Scoring procedures, validity, and reliability of the test are discussed in Lawson et al. (2000a).

Dependent Measures

To determine the effect of learning pre-class content using each method, we analyzed student performance in three ways. First, we used the average first attempt score of 41 pre-class Explore Assessments for each student. Second, we averaged eight Apply Assessment scores for each student. Finally, we used the comprehensive final exam score. All assessments (41 Explore Assessments, 8 Apply Assessments, and the final) were identical for all treatments across both institutions, based on a common set of learning outcomes.

In an attempt to measure how much effort students dedicated to learning the material on their own prior to class, and as a way to detect any deficiencies in learning between the methods (interactive tutorials, video lectures, or textbook-style readings), we analyzed the number of attempts that students took on the Explore Assessments. With unlimited attempts and the only feedback being their score (i.e., they were not shown which items they missed nor which were correct), many students retook these Explore Assessments multiple times to earn a score as close to 100% before the deadline. Students' average number of attempts on the 41 Explore Assessments was our final response variable to measure effort.

Method of Analysis

A 2 × 3 analysis of covariance (ANCOVA) was conducted on all data combined with LCTSR as the covariate, using Institution (Public or Private) and Treatment condition (Interactive Tutorials, Video Lectures, or Textbook-style Readings) as between-subjects variables, to determine if the method by which content was learned affected performance on each of the outcome variables: Explore Assessment first attempt scores, Explore Assessment number of attempts, Apply Assessment scores, and Final Exam scores. Post hoc analyses were done using a Bonferroni correction to account for alpha inflation.

Results

Scientific Reasoning Ability

LCTSR score distributions varied between institutions (Fig. 2), with the private institution having higher scores overall than the public institution ($M_{Private} = 18.3$, $M_{Public} = 13.6$, $F(1, 655) = 128.64$, $p < 0.001$, $\eta_p^2 = 0.170$).

Explore Assessments

The first attempt on pre-class Explore Assessments, an indicator of content learned from the pre-class materials, differed by institution and treatment. A 2 × 3 ANCOVA showed a significant interaction between the treatment and the institution ($F(2, 622) = 6.87$, $p = 0.001$, $\eta_p^2 = 0.022$) (Fig. 3a). Multiple comparisons within the main treatment effect demonstrated significant differences at the $p < 0.017$ level, our Bonferroni adjustment. At the private institution, the Interactive Tutorial treatment [$M = 69.9$, 95% CI (68.2, 71.6)] outperformed the Video Lecture treatment [$M = 67.6$, 95% CI (65.6, 69.6), $p = 0.016$], while the Textbook-style Readings [$M = 68.9$, 95% CI (67.4, 70.4)] treatment was not significantly different than either Interactive Tutorials or Video Lectures ($p = 1.0$, $p = 0.058$), respectively. Although this difference may be statistically significant, it may not be practically significant as the Interactive Tutorial improvement was roughly 0.14 points on an average seven-item Explore Assessment. At the public institution, none of the differences reached significance. The Video Lecture treatment [$M = 63.3$, 95% CI (58.2, 68.4)] was not significantly different than the Interactive Tutorial treatment [$M = 55.7$, 95% CI (51.7, 59.7), $p = 0.027$], and the Textbook-style Readings [$M = 59.4$, 95% CI (53.9, 64.9)] treatment was not significantly different than either Interactive Tutorials or Video Lectures ($p = 0.124$, $p = 1.0$), respectively.

Explore Assessment Attempts

The number of attempts students took on the Explore Assessments, at both institutions, did not significantly differ by treatment ($F(2, 622) = 0.06$, $p = 0.944$). However, there was a significant institution effect with students at the public institution taking significantly more attempts [$M = 8.5$, 95% CI (7.8, 9.2)], than students at the private institution [$M = 3.7$, 95% CI (3.6, 3.8), $F(1, 622) = 285.46$, $p < 0.001$, $\eta_p^2 = 0.315$] (Fig. 3b). There was no treatment by institution interaction [$F(2, 622) = 0.36$, $p = 0.696$].

Fig. 2 Boxplots of LCTSR scores by institution
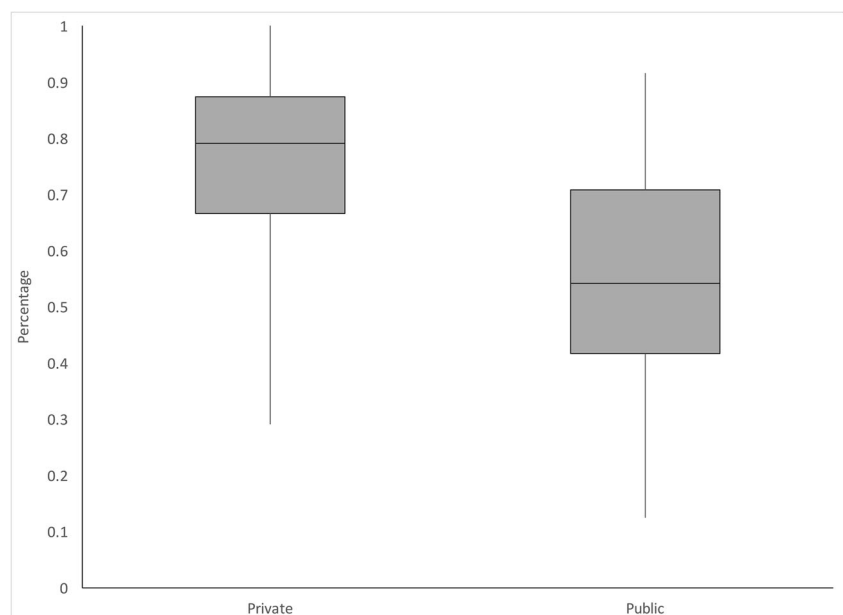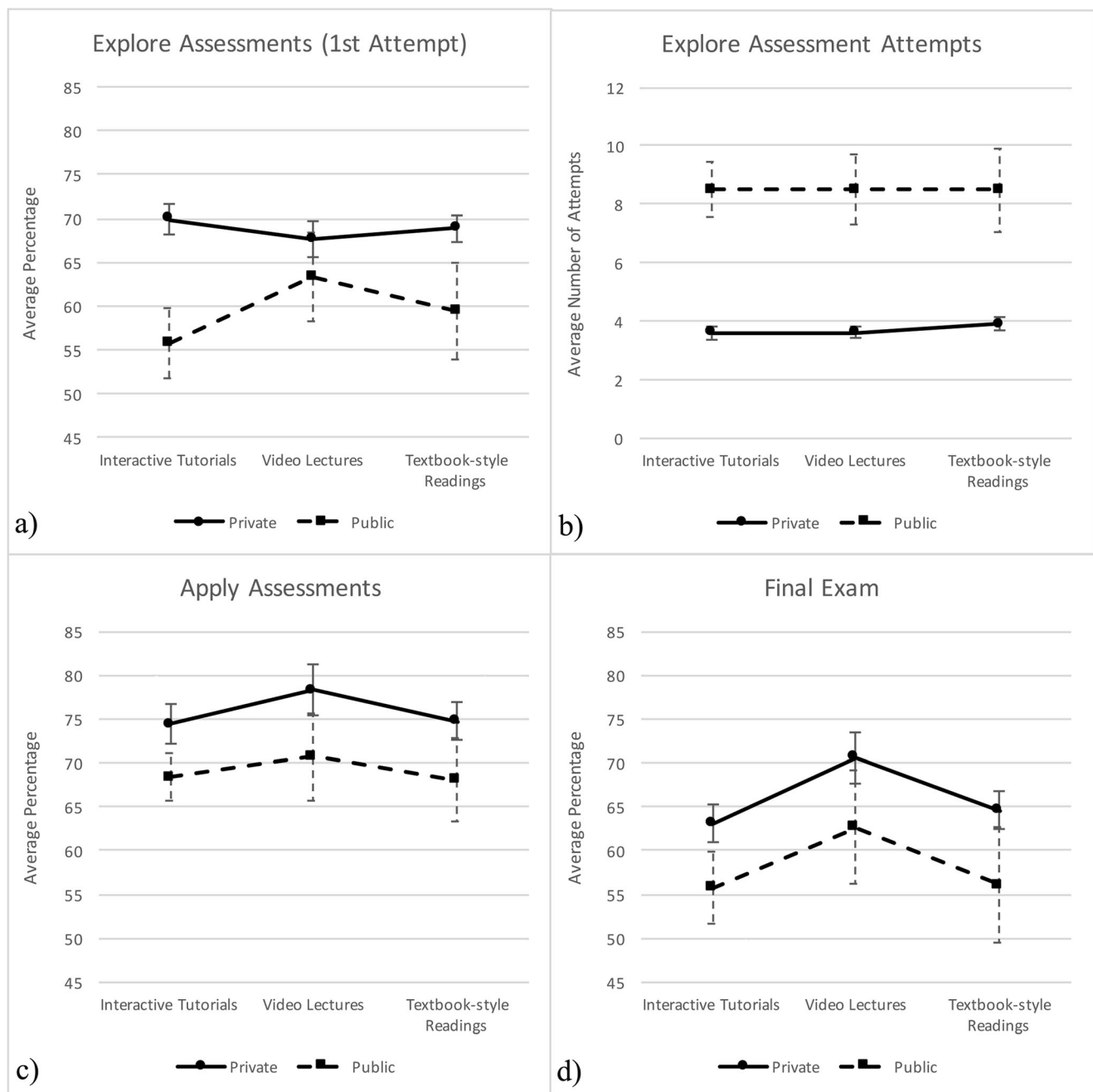
Fig. 3 Raw scores of four dependent variables are shown by treatment condition at both a private (solid lines) and public (dashed lines) institution. Error bars represent 95% confidence intervals. Panel **a** shows average scores on the first attempt of pre-class Explore Assessments. Significant findings include a significant interaction ($p = 0.001$) wherein only Interactive Tutorials are different between institutions ($p < 0.001$). Panel **b** shows the average number of attempts students made on the pre-class Explore Assessments; the public institution students made significantly more attempts than the private institution students on all treatments ($p < 0.001$). Panel **c** shows the average scores on Apply Assessments; no differences were found to be significant. Panel **d** shows the average scores on the Final Exam; the Video Lecture Treatment was significantly greater than the Interactive Tutorials treatment ($p = 0.007$)

Apply Assessments

The method of content learning had no effect on average Apply Assessment scores, with a non-significant treatment effect [$F(2, 622) = 1.28$, $p = 0.278$] and institution effect [$F(1, 622) = 0.32$, $p = 0.571$]. We also found the treatment ×

institution interaction was not significant [$F(2, 622) = 0.06$, $p = 0.945$]. At the private institution, students scored an average of 74.4% [95% CI (72.8, 76.0)] in the Interactive Tutorials condition, 78.3% [95% CI (76.3, 80.3)] in the Video Lectures, and 74.7% [95% CI (73.1, 76.3)] in the Textbook-style readings condition. At the public institution, students scored an

average of 68.4% [95% CI (65.8, 71.1)] in the Interactive Tutorials condition, 70.7% [95% CI (65.8, 75.6)] in the Video Lectures, and 68.0% [95% CI (63.3, 72.7)] in the Textbook-style readings condition (see Fig. 3c).

Final Assessment

The Final Assessment was comprehensive and meant to serve as a summative assessment of students' overall understanding at the conclusion of the course. A 2×3 ANCOVA revealed a significant treatment effect ($F(2, 622) = 4.69$, $p = 0.009$, $\eta_p^2 = 0.015$). Pairwise comparisons were performed with a Bonferroni correction such that significance was accepted at $p < 0.017$. The Video Lecture treatment [$M = 69.3$, 95% CI (66.6, 72.0)] significantly outperformed the Interactive Tutorial treatment [$M = 61.4$, 95% CI (59.4, 63.4), $p = 0.007$], and the Textbook-style Readings [$M = 63.5$, 95% CI (61.4, 65.6)] treatment was not significantly different than either Interactive Tutorials or Video Lectures ($p = 0.647$, $p = 0.244$), respectively. There was no significant effect of institution [$F(1, 622) = 0.67$, $p = 0.415$], nor was the treatment $\times$ institution interaction significant [$F(2, 622) = 0.01$, $p = 0.994$] (Fig. 3d).

Discussion

In this study, we sought to explore differences in several of the most popular ways of flipping a classroom. Specifically, we tested three strategies for facilitating pre-class content learning: encouraging student interaction with online material at home, having students watch videos of lecture, or assigning textbook-style readings. Three main findings surfaced from our work: first, video lectures appear to offer a small advantage to overall student learning; second, populations at different institutions differ in their preparedness to effectively learn from pre-class activities; and third, despite this inequality in preparedness, both populations demonstrate equivalent learning gains after experiencing a student-centered, flipped classroom curriculum.

High-Stakes Outcomes

Although the method for content learning did not appear to differentially affect student performance on our higher-stakes unit exams (i.e., Apply Assessments), video lectures appeared to offer a clear advantage on the final summative assessment of the course, raising their grade by an average of nearly eight percentage points. This would be the difference between a B+ and an A, a practically significant difference. Both high-stakes assessments required students to both remember content (low-level items) but also apply the concepts in a higher-order fashion in over half the questions (high-level items). Appealing to

a dual coding theoretical rationale (Paivio 1990), it is possible that, in students who are self-motivated or academically prepared enough to gain information in the independent fashion required in a flipped classroom, receiving information through both a visual and auditory route may have an additive rather than competitive effect in terms of information gain (Yadav et al. 2011). Likewise, the combined audio and visual information may have contributed extra signaling to highlight important information to the benefit of Video Lecture students above the other treatments (Mautone and Mayer 2001).

Beyond signals of the text within the video (e.g., arrows, zooming into particular areas of a slide, or highlighting text), the instructor shown in the video often gesticulated, which can provide additional cognitive aids (Singer and Goldin-Meadow 2005). In addition, inflection patterns that we commonly include in our speech patterns may have offered students additional unintended cues to important information or places to pause that encouraged thought, where these cues are largely absent from written passages which predominated both the Interactive Tutorials and Textbook-style Reading conditions. Psychological research suggests that prosody in speech can convey the importance of phrases to students outside the actual lexical channel (i.e., the meaning of the actual words being spoken; Johar 2016). In other words, unconscious cues found both in the tone, pitch, and tempo of our speech as well as in hand gestures or facial expressions, all visible throughout in the Video Lectures, often convey meaning to students in the course of normal dialogue. Video Lectures were the only method where human speech was present for pre-class delivery of content. Social Agency Theory posits that a human voice integrating into learning can increase student motivation (Mayer et al. 2003) and attentiveness (McLaren et al. 2011). Additionally, the presence of a human face has been shown to positively impact student learning through increased comprehension via the opportunity to lip read and apprehend positive social cues (Kizilcec et al. 2015). These benefits likely outweigh demonstrated drawbacks, distractions, or the presence of human agency during a lecture video (Schroeder and Traxler 2017). The combination of factors associated with a human pedagogical agent could overtly influence the effectiveness of video lectures versus silent written text or short videos set to music and lacking instructors' facial expressions and voice, regardless of the considerable overlap in content.

In addition, due to the nature of the method, information may be more easily re-accessed from textbook-style readings and even the interactive tutorials as we originally hypothesized, but this access may be done in a superficial way. These were both available as easily skimmable pdf documents via their LMS after the Explore Assessment deadline passed. However, to re-visit information for studying from a video lecture, students would necessarily have to re-watch the videos leading to unintended deeper processing. While students had control over pacing and replaying of the video, in

this "learner-attenuated system-paced" instruction model (Schroeder and Traxler 2017), students had to make decisions on where to return in the video and undoubtedly watched more than anticipated. We suggest that through these more complete repeated visits to the material in preparation for the final exam, students in the Video Lecture condition better solidified their understanding over students in the other modalities.

Interestingly, among the high-stakes outcomes, the advantage of the video treatment was only observed in the Final Assessment and not the Apply Assessments. Both assessments were written with roughly the same proportion of high and low Bloom level items, and both were timed. The key differences, however, were that the Final Assessment was closed-note and comprehensive. We suspect that students likely revisited the videos, with more unintended exposure, in preparation for the final, because they could not rely on notes and it may have been months since they originally learned the material. In contrast, students may minimally or fail entirely to review pre-class content in preparation for Apply Assessments, assuming that they could access it during the assessment, if needed, or that the material was learned recently enough (within 2 weeks) to remember. Regardless of whether the video benefits lie in dual coding, extra cuing, or the presence of a human pedagogical agent, certainly allowing for greater exposure to more content through re-watching video content provides a long-term benefit to students in our sample.

Low-Stakes Outcomes

Interestingly, when comparing initial attainment of materials from pre-class activities (using Explore Assessment scores, a largely low-level assessment format) from the three methods, a significant interaction emerged between treatment and institution. At the private institution, although the Video Lecture treatment showed a slight disadvantage to an interactive tutorial, and this trend has precedence in the literature (Abeysekera and Dawson 2015), the differences were minimal (0.14 points on a 7-point assessment, a difference that would not likely change their grade) bringing into question whether it was practically significant. At the public institution, no method offered an advantage or disadvantage for initial attainment of information. These Explore Assessments reflect the amount of learning (albeit, rather low-level) students achieved independently, which may reflect many things outside of content learning. Specifically, student performance may simply mirror students' prior knowledge or their study strategies, which then reveals a student's preparedness for college-level instruction or online learning. Research shows that many students are less than ready to engage in a flipped classroom upon entering college (Hao 2016). In fact, Yilmaz (2017) showed that a student's e-learning readiness was a direct predictor of their motivation in a flipped classroom.

Perhaps at a highly selective institution, students' overall readiness to engage independently in learning activities prior to instructor-scaffolded class activities makes the constructivist, interactive tutorial treatment more successful in initial attainment. Whereas, at an open-enrollment institution, the pre-class, content learning method seemed to have little effect on their success with scores being consistently lower on initial attainment than the private institution, regardless of the mode of delivery. Dramatic differences in the number of attempts on Explore Assessments (i.e., significantly more in the public institution than in the private institution) provide further evidence of differences in preparedness between institutions.

Population Differences

As was seen by pre-LCTSR scores, these institutions differed dramatically in students' incoming scientific reasoning ability. Scientific reasoning skills are tightly correlated with achievement in college biology (e.g., Johnson and Lawson 1998; Lawson et al. 2007) and differ significantly between those who pursue STEM degrees and those who do not (Jensen et al. 2015b). Despite these differences in initial reasoning ability between our two institutions, no institutional differences were observed in final assessment performance. In other words, following completion of a student-centered, flipped biology course, students at the public institution reached indistinguishable levels of understanding on the summative final assessment as students from the selective private institution. While students at both institutions benefited from the student-centered, flipped classroom approach (Strayer 2012; Tucker 2012; Gajjar 2013; Sarawagi 2013), it seems that lower performing, underprepared students may benefit even more and can achieve equal learning on the final course assessment as higher-performing students.

Limitations and Future Directions

Results from this research are intriguing, but they also create more questions to be answered. One of the limitations of this study is that it was conducted on only two student populations. However, one represented a more selective, high-achieving group, while the other was more representative of an open-enrollment, public institution, allowing two ends of a continuum to be tested. Both student populations lacked sufficient representation of underrepresented minorities leaving it open-ended as to whether social group may be an interacting factor in the success of each model. It is certainly worth further study. Another potential limitation is that 70% of the video lectures were prepared by one of the instructors of the courses with the other 30% being prepared by the second instructor and 0% being prepared by the third instructor. This means that in some of the courses, the students watched video lectures by a professor other than their own. Whether this played a role in their engagement with the video is unknown.

Several causal mechanisms have been suggested for the advantage of video lectures. These causal mechanisms represent future avenues for valuable research. For example, we suggest that extra cuing, through prosody of speech, may give students an advantage in cuing them to important information. Trials in which prosody is altered may shed further light on this mechanism. Additionally, the presence of a pedagogical agent in the videos may have played a role in the videos' success. Removing the pedagogical agent on screen and, instead, only including audio source material may shed light on this causal mechanism. Lastly, one of the major causal mechanisms suggested is the different depth to which information may have been re-accessed in preparation for the final exam. Monitoring re-access more closely may allow researchers to tease out this causal mechanism in more detail.

Conclusions

The overall effectiveness of a flipped classroom, in general, may be highly dependent upon the student population, which can vary in academic preparation, scientific reasoning ability, or self-directed learning skills. However, we found that the method for pre-class content learning may play a differential role in effectiveness, with the Video Lecture strategy being superior to interactive or textbook-style readings on final assessment performance. Our findings are in contrast to Moravec and others' (2010) work that detected no performance differences between video lectures and readings. Yet our study used flipped strategies throughout the entire semester, rather than three isolated days, which may better represent what happens in a fully flipped classroom. Currently, our data cannot suggest a causal mechanism for this benefit and more research of the benefits of video lecturing, compared to other flipped methods of "at home" content learning, is warranted.

Compliance with Ethical Standards

Conflict of Interest The authors declare that they have no conflict of interest.

Ethical Approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent Informed consent was obtained from all individual participants included in the study.

References

Abeysekera, L., & Dawson, P. (2015). Motivation and cognitive load in the flipped classroom: definition, rationale and a call for research. *Higher Education Research and Development, 34*(1), 1–14.

Baepler, P., Walker, J. D., & Driessen, M. (2014). It's not about seat time: blending, flipping, and efficiency in active learning classrooms. *Computers in Education, 78*, 227–236.

Bergstrom, G. (2011). Content vs. learning: an old dichotomy in science courses. *Journal of Asynchronous Learning Networks, 15*, 33–44.

Besser, D., Stone, G., & Nan, L. (1999). Textbooks and teaching: a lesson from students. *Journalism & Mass Communication Educator, 53*(4).

Betihavas, V., Bridgman, H., Kornhaber, R., & Cross, M. (2016). The evidence for 'flipping out': a systematic review of the flipped classroom in nursing education. *Nurse Education Today, 38*, 15–21.

Bloom, B. S. (1984). *Taxonomy of educational objectives*. Boston: Allyn and Bacon.

Bybee, R. (1993). *An instructional model for science education developing biological literacy*. Colorado Springs, Colorado: Biological Sciences Curriculum Studies.

Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review, 3*(3), 149–210. https://doi.org/10.1007/BF01320076.

Clump, M. A., Bauer, H., & Bradley, C. (2004). The extent to which psychology students read textbooks: a multiple class analysis of reading across the psychology curriculum. *Journal of Instructional Psychology, 31*(3), 227–232.

Damon, W. (1984). Peer education: the untapped potential. *Journal of Applied Developmental Psychology, 5*(4), 331–343.

Davies, R. S., Dean, D. L., & Ball, N. (2013). Flipping the classroom and instructional technology integration in a college-level information systems spreadsheet course. *Educational Technology Research and Development, 61*(4), 563–580.

Dewey, J. (1938). *Experience and education*. New York: Collier Books.

Doise, W., & Mugny, G. (1984). *The social development of the intellect*. Oxford: Pergamon Press.

Fautch, J. M. (2015). The flipped classroom for teaching organic chemistry in small classes: is it effective? *Chemistry Education Research and Practice, 16*(1), 179–186.

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *PNAS, 111*(23), 8410–8415.

Gajjar, N. B. (2013). The role of technology in 21st century education. *International Journal for Research in Education, 2*, 23–25.

Gilboy, M. B., Heinerichs, S., & Pazzaglia, G. (2015). Enhancing student engagement using the flipped classroom. *Journal of Nutrition Education and Behavior, 47*(1), 109–114.

Hamdan, N., McKnight, P., McKnight, K., Arfstrom, K. M. (2013). A review of flipped learning. Flipped Learning Network, Pearson Education, and George Mason University. http://www.flippedlearning.org/research. Accessed 15 July 2013.

Harrington, S. A., Bosch, M. V., Schoofs, N., Beel-Bates, C., & Anderson, K. (2015). Quantitative outcomes for nursing students in a flipped classroom. *Nursing Education Perspectives, 36*(3), 179–181.

Hao, Y. (2016). Exploring undergraduates' perspectives and flipped learning readiness in their flipped classrooms. *Computers in Human Behavior, 59*, 82–92.

He, W., Holton, A., Farkas, G., & Warschauer, M. (2016). The effects of flipped instruction on out-of-class study time, exam performance, and student perceptions. *Learning and Instruction, 45*, 61–71.

He, Y., Swenson, S., & Lents, N. (2012). Online video tutorials increase learning of difficult concepts in an undergraduate analytical chemistry course. *Journal of Chemical Education, 89*(9), 1128–1132.

Hung, H. T. (2015). Flipping the classroom for English language learners to foster active learning. *Computer Assisted Language Learning, 28*(1), 81–96.

Jensen, J. L., Kummer, T. A., & Godoy, P. D. d. M. (2015a). Improvements from a flipped classroom may simply be the fruits of active learning. *CBE Life Sciences Education, 14*, 1–12.

Jensen, J. L., Neeley, S., Hatch, J. B., & Piorczynski, T. (2015b). Learning scientific reasoning skills may be key to retention in science, technology, engineering, and mathematics. *Journal of College Student Retention Research, Theory, & Practice, 0*, 19–11. https://doi.org/10.1177/1521025115611616.

Jeong, J. S., & González-Gómez, D. (2016). Students' perceptions and emotions toward learning in a flipped general science classroom. *Journal of Science Education and Technology, 25*(5), 747–758.

Johar, S. (2016). Emotion, affect and personality in speech: The Bias of language and paralanguage. New York: Springer International Publishing AG.

Johnson, M. A., & Lawson, A. E. (1998). What are the relative effects of reasoning ability and prior knowledge on biology achievement in expository and inquiry classes? *Journal of Research in Science Teaching, 35*(1), 89–103.

Kay, R., & Kletskin, I. (2012). Evaluating the use of problem-based video podcasts to teach mathematics in higher education. *Computers in Education, 59*(2), 619–627.

Keller, J. M. (1983). Motivational design of instruction. In C. M. Reigeluth (Ed.), *Instructional-design theories and models: an overview of their current status*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Keller, J. M. (2009). Motivational design for learning and performance: the ARCS model approach. Springer Science & Business Media.

Kim, M. K., Kim, S. M., Khera, O., & Getman, J. (2014). The experience of three flipped classrooms in an urban university: an exploration of design principles. *The Internet and Higher Education, 22*, 37–50.

Kizilcec, R. F., Bailenson, J. N., & Gomez, C. J. (2015). Motivation as a lens to understand online learners: toward data-driven design with the OLEI scale. *ACM Transactions on Computer-Human Interactions, 22*(2) No. 6.

Kubli, F. (1989). Piaget and interest in science subjects. In P. Adey, J. Bliss, & M. Shayer (Eds.), *Adolescent development and school science*. New York: Falmer Press.

Lawson, A. E. (1978). The development and validation of a classroom test of formal reasoning. *Journal of Research in Science Teaching, 15*(1), 11–24.

Lawson, A. E. (2002). *Science teaching and development of thinking*. Belmont, CA: Wadsworth/Thompson Learning.

Lawson, A. E., Alkhoury, S., Benford, R., Clark, B. R., & Falconer, K. A. (2000a). What kinds of scientific concepts exist? Concept construction and intellectual development in college biology. *Journal of Research in Science Teaching, 37*(9), 996–1018.

Lawson, A. E., Clark, B., Cramer-Meldrum, E., Falconer, K. A., Sequist, J. M., & Kwon, Y.-J. (2000b). Development of scientific reasoning in college biology: do two levels of general hypothesis-testing skills exist? *Journal of Research in Science Teaching, 37*(1), 81–101.

Lawson, A. E., Banks, D. L., & Logvin, M. (2007). Self-efficacy, reasoning ability, and achievement in college biology. *Journal of Research in Science Teaching, 44*(5), 706–724.

Leo, J., & Puzio, K. (2016). Flipped instruction in a high school science classroom. *Journal of Science Education and Technology, 25*(5), 775–781.

Long, T., Cummins, J., & Waugh, M. (2017). Use of the flipped classroom instructional model in higher education: instructors' perspectives. *Journal of Computing in Higher Education, 29*(2), 179–200.

Lumpe, A. T. (1995). Peer interaction in science concept development and problem solving. *School Science and Mathematics, 95*(6), 302–309.

Mautone, P. D., & Mayer, R. E. (2001). Signaling as a cognitive guide in multimedia learning. *Journal of Education & Psychology, 93*(2), 377–389.

Mayer, R. E., Sobko, K., & Mautone, P. D. (2003). Social cues in multimedia learning: role of speaker's voice. *Journal of Education & Psychology, 95*(2), 419–425.

McLaren, B., DeLeeuw, K. E., & Mayer, R. E. (2011). A politeness effect in learning with web-based intelligent tutors. *International Journal of Human-Computer Studies, 69*(1–2), 70–79.

McLaughlin, J. E., Roth, M., Glatt, D. M., Gharkholonarehe, N., Davidson, C. A., Griffin, L. M., Esserman, D. A., & Mumper, R. J. (2014). The flipped classroom: a course redesign to foster learning and engagement in a health professions school. *Academic Medicine, 89*, 1–8.

Moravec, M., Williams, A., Aguilar-Roca, N., & O'Dowd, D. K. (2010). Learn before lecture: a strategy that improves learning outcomes in a large introductory biology class. *CBE Life Sciences Education, 9*(4), 473–481.

Moreno, R., & Mayer, R. E. (2002). Verbal redundancy in multimedia learning: when reading helps listening. *Journal of Education & Psychology, 94*(1), 156–163.

O'Donnell, A. M., Hmelo-Silver, C. E., & Erkens, G. (2006). *Collaborative learning, reasoning, and technology*. Mahwah, New Jersey: Lawrence Erlbaum Associates, Publishers.

O'Flaherty, J., & Philips, C. (2015). The use of flipped classrooms in higher education: a scoping review. *The Internet and Higher Education, 25*, 85–95.

Olakanmi, E. E. (2017). The effects of a flipped classroom model of instruction on students' performance and attitudes towards chemistry. *Journal of Science Education and Technology, 26*(1), 127–137.

Paivio, A. (1990). *Mental representations: a dual coding approach*. New York: Oxford University Press.

Phillips, B. J., & Phillips, F. (2007). Sink or skim: textbook reading behaviors of introductory accounting students. *Issues in Accounting Education, 22*(1), 21–44.

Phillips, C. R., & Trainor, J. E. (2014). Millennial students and the flipped classroom. *Journal of Business and Educational Leadership, 5*(1), 102.

Piaget, J. (1985). *The equilibration of cognitive structures: the central problem of intellectual development*. Chicago: University of Chicago Press.

Pierce, R., & Fox, J. (2012). Vodcasts and active-learning exercises in a "flipped classroom" model of a renal pharmacotherapy module. *American Journal of Pharmaceutical Education, 76*(10), 196.

Porcaro, P. A., Jackson, D. E., McLaughlin, P. M., & O'Malley, C. J. (2016). Curriculum design of a flipped classroom to enhance haematology learning. *Journal of Science Education and Technology, 25*(3), 345–357.

Pulaski, M. A. S. (1980). *Understanding Piaget: an introduction to children's cognitive development*. New York: Harper & Row, Publishers.

Qualtrics software. (2015). Qualtrics, Provo, Utah, USA. http://www.qualtrics.com.

Roehl, A., Reddy, S. L., & Shannon, G. J. (2013). The flipped classroom: an opportunity to engage millennial students through active learning. *Journal of Family and Consumer Sciences, 105*(2), 44–49.

Sarawagi, N. (2013). Flipping an introductory programming course—yes you can! Tutorial Presentation, CCSC: Northeastern Conference. http://delivery.acm.org/. Accessed 15 July 2013.

Schroeder, N. L., & Traxler, A. L. (2017). Humanizing instructional videos in physics: when less is more. *Journal of Science Education and Technology, 26*(3), 269–278.

Schwarzenberg, P., Navon, J., Nussbaum, M., Pérez-Sanagustín, M., & Caballero, D. (2017). Learning experience assessment of flipped courses. *Journal of Computing in Higher Education*, 1–22.

Seery, M. (2015). Flipped learning in higher education chemistry: emerging trends and potential directions. *Chemical Education Research and Practice, 16*(4), 758–768.

Sikorski, J. F., Rich, K., Saville, B. K., Buskist, W., Drogan, O., & Davis, S. F. (2002). Student use of introductory texts: comparative survey findings from two universities. *Teaching of Psychology, 29*(4), 312–313.

Singer, M. A., & Goldin-Meadow, S. (2005). Children learn when their teacher's gestures and speech differ. *Psychological Science, 16*(2), 85–89.

Sletten, S. R. (2017). Investigating flipped learning: student self-regulated learning, perceptions, and achievement in an introductory biology course. *Journal of Science Education and Technology, 26*(3), 347–358.

Strayer, J. F. (2012). How learning in an inverted classroom influences cooperation, innovation and task orientation. *Learning Environments Research, 15*(2), 171–193.

Thomas, J. T. (2014). Dual coding and common coding theories of memory. Stanford Encyclopedia of Philosophy. https://plato.stanford.edu/entries/mental-imagery/theories-memory.html. Accessed 8 June 2017.

Tomory, A., & Watson, S. L. (2015). Flipped classrooms for advanced science courses. *Journal of Science Education and Technology, 24*(6), 875–887.

Tucker, B. (2012). The flipped classroom. Education Next: 12. http://educationnext.org/the-flipped-classroom/. Accessed 15 July 2013.

Tune, J. D., Sturek, M., & Basile, D. P. (2013). Flipped classroom model improves graduate student performance in cardiovascular, respiratory, and renal physiology. *Advances in Physiology Education, 37*(4), 316–320.

Yadav, A., Phillips, M. M., Lundeberg, M. A., Koehler, M. J., Hilden, K., & Dirkin, K. H. (2011). If a picture is worth a thousand words is video worth a million? Differences in affective and cognitive processing of video and text cases. *Journal of Computing in Higher Education, 23*(1), 15–37.

Yilmaz, R. (2017). Exploring the role of e-learning readiness on student satisfaction and motivation in flipped classroom. *Computers in Human Behavior, 70*, 251–260.

Zainuddin, Z., & Halili, S. H. (2016). Flipped classroom research and trends from different fields of study. *International Review of Research in Open and Distributed Learning, 17*(3).