

Assessing student-learning gains from video lessons in a flipped calculus course

Cassandra Williams
James Madison University

John (Zig) Siegfried
James Madison University

Abstract

The flipped classroom has garnered attention in post-secondary mathematics in the past few years, but much of the research on this model has been on student perceptions rather than its effect on the attainment of learning goals. Instead of comparing to a “traditional” model, in this study we investigated student-learning gains in two flipped sections of Calculus I. In this paper, we focus on the question of determining learning gains from delivering content via video outside of the classroom. In particular, we compare student-learning gains after watching more conceptual videos versus more procedural ones. We share qualitative and quantitative data gathered from surveys and quizzes, as well as results from in-class assessments. We conclude by sharing some implications for future research.

Keywords: Flipped Classroom, Video Lessons, Learner-Centered Teaching, Calculus

Background

Learner-centered or active classrooms are those which change the role of the instructor from “sage on the stage” to “guide on the side” and encourage students to construct their own meaning while engaging in authentic problem-solving. Recent research has consistently showed that active classrooms improve student learning in a variety of fields. For example, in 2014 the National Academy of Sciences published a meta-study of 225 studies on student performance and failure rates in undergraduate science, technology, engineering, and mathematics [STEM] classrooms employing active learning components. Their analysis suggests that students in traditional lecture classrooms are 1.5 times more likely to fail than students in classrooms including any type of active learning technique, and failure rates in lecture classrooms are 55% higher than in active classrooms (Freeman et al., 2014). In addition, a 2013 report from the President’s Council of Advisors on Science and Technology [PCAST] called for 1 million more college graduates in STEM over the next decade (PCAST STEM Undergraduate Working Group, 2013). Given the current retention rate in STEM majors during the first two years of college and the decreased failure rate in active learning classrooms reported in Freeman et al. (2014), a majority of this goal could be met by employing more active learning and less lecturing during classroom time.

The *flipped* (or *inverted*) classroom structure is one example of an active learning method that has become increasingly popular. This classroom structure takes on many forms, but the common trait is that most of the initial content delivery happens outside of the classroom while in-class time is spent solving problems, often in small groups, to assimilate the new knowledge and to deepen understanding. Some instructors deliver content through assigned readings from a text or other source, while others use videos curated from those available online or create their own videos. The core idea is to use classroom time for challenging problem-solving where students can draw support from their peers and instructor; this design more effectively uses the experience and knowledge of the instructor to guide students through the topic at hand.

Literature Review

Much of the initial literature on flipped classrooms only described the varying structures of such classrooms or the particular technologies employed by teachers using a flipped classroom. The controlled studies published on this classroom model have often focused on student perceptions of and attitudes towards the structure rather than its effect on the attainment of learning goals. For example, Foertsch, Moses, Strikwerda, and Litzkow (2002) described the use of a specific video streaming software in an engineering classroom, and reported student opinions of the videos and software, and Ford (2015) described the activity structure in a math content course for pre-service elementary teachers. Strayer (2007) gathered data on a traditional and flipped introductory statistics classroom to evaluate the learning environment of each structure, and found that students enjoyed the innovation and cooperation in the flipped class, but had a low “comfortability” with the learning activities in this environment. Roach (2014) found that 76% of students in an economics class believed that video lectures helped them learn, and the same percentage would take another class using the flipped format. Bishop and Verlager (2013) did a meta-analysis of the literature on flipped classrooms in all areas of STEM, as well as economics and sociology, and found that there were few studies examining student achievement and advocated for more controlled research.

While lecturing has been a staple of academia for close to a millennium, the flipped classroom structure might be seen as a return to an even older system of teaching where classroom time was centered around academic debate and discussion rather than the transmission of information. Modern flipped classrooms are now returning to this classroom structure and also taking advantage of newer technologies like video and the Internet. This recent resurgence dates to at least the mid-1990s when Eric Mazur, a physics professor at Harvard, started using team learning and in-class activities as ways to stop lecturing (Mazur, 1996). Jonathan Bergmann and Aaron Sams (2012) started using video lectures in the mid-2000s and are often credited with pioneering the flipped classroom and its current popularity. Since then, many educators in a variety of fields and at a wide range of institutions have started using this structure. For example, Gaughn (2014) wrote about their experiences running a flipped history classroom, and Findlay-Thompson and Mombourquette (2014) published research from their flipped business classroom. Additionally, research has been done on flipped classrooms at levels ranging from high school (Johnson, 2013; Moore, Gillett, & Steele, 2014) to upper division medical courses (Sharma, Lau, Doherty, & Harbutt, 2015). Education-focused video repositories like Khan Academy are available on the web, and many have spoken about their experiences with various forms of the flipped classroom at local and national professional meetings (e.g., in 2014 the Joint Mathematics Meetings included a session titled *Flipping the Classroom* with 37 different talks).

As the flipped classroom has gained popularity among undergraduate STEM educators, more research studies are using classroom data to evaluate the success of flipped classrooms. Lape et al. (2014) and Mason, Shuman, and Cook (2013) compare grades on individual assessment questions in engineering between flipped and traditional sections of the same course and found few cases of statistically significantly higher scores in the flipped classroom, but no cases where students in a lecture section outperformed students in a flipped section. Similarly, Day and Foley (2006) compared grades on several course components in a senior level computer science elective and found that the flipped section earned higher average scores on every component of the grade, with statistically significant differences in the case of homework based on lectures/video lectures. Moravec et al. (2010) found statistically significant score increases

over previous years in matched exam questions related to topics delivered in an inverted fashion in a large introductory biology course.

In mathematics in particular, McGivney-Burelle and Xue (2013) flipped a unit in a Calculus II course and showed that student grades on exams and homework were higher for the flipped section than the traditional section. Wilson (2013) found that students in a flipped section of statistics outperformed their lecture counterparts on exams and the course post-test. Love, Hodge, Grandgenett, and Swift (2014) found that students in a flipped linear algebra course had greater improvement in exam scores than those in a traditional section, and had higher averages on the final exam. Additionally, Problems, Resources, and Issues in Mathematics Undergraduate Studies (PRIMUS) has a forthcoming special issue on research in flipped classrooms that will increase the literature within mathematics education.

Some researchers have also considered the format, use, and effectiveness of video lectures, both in flipped classrooms and in general. For example Zappe, Leicht, Messner, Litzinger, and Lee (2009) investigated how students used online lecture videos to learn in an undergraduate engineering course, including the percentage of videos watched, students reviewing unclear segments, and time spent per video. Mayer and colleagues have published a number of papers considering specific attributes of videos, like the use of graphics and animations, or the style and tone of the voice in the video, and how they help or hinder student learning (e.g., Mayer, Hegarty, Mayer, & Campbell, 2005; Mayer, Sobko, & Mautone, 2003).

Research Question

Since students in the flipped classroom model do introductory learning of topics outside of the classroom, it is prudent to investigate the effectiveness of the content delivery method. The classroom in our study most often introduced new content outside of class through the use of the instructor's own video-recorded lessons. In this study we investigate the effectiveness of these videos on the learning gains made by students enrolled in two sections of a standard first semester undergraduate calculus course. In particular, we explore student-learning gains from watching videos outside the classroom to determine students' development of conceptual understanding and procedural skills in calculus.

Methods

Participants

The participants were undergraduate students in a first semester calculus course at a large comprehensive public university in the Mid-Atlantic United States. Of the 59 students in the study, 51 (86%) were freshmen, 5 (8%) were sophomores, 2 were juniors, and 1 was a senior. The majority of the students were male (64% male, 36% female). Four students withdrew from the course before the end of the semester. More than 80% of the students had previously had a course in calculus, generally in high school. The majority of the students were majoring in STEM fields. The students were divided into two sections (34 students in one section, 25 in the other) and generally covered the same material on the same days.

Classroom

The data was collected during the instructor's third semester running a flipped Calculus I classroom. Before each class, students had a pre-class assignment, such as watching a video or completing a reading. Nearly all class sessions started off with a short open-note quiz related to their pre-class assignment. The majority of class time was spent on group-work activities.

These activities were often sets of questions designed to reinforce, clarify, deepen, and extend the content in the pre-class assignment, address misconceptions, and provide practice. At times throughout the semester, activity were discovery or guided-inquiry, meant to allow students to develop a key concept or idea on their own through the use of a carefully chosen set of leading questions and problems. The students worked in groups of two to four students and the instructor would interact with the groups one-on-one. Students were also given homework and practice problems to be completed outside of class.

Data Sources

Over the course of the semester, we gathered qualitative data from the students, including student feedback about specific video lectures (for example, questions like “What did you find confusing?” or “What helped clear up confusion?”), student answers to post-video or post-activity questions or problems (calculus content questions to evaluate learning gains), and student surveys about their perceptions of the class structure and their learning gains. Aggregate quantitative data, such as assessment scores and exam grades, were also recorded. We used video recording on certain class days to help the instructor objectively evaluate and improve student-teacher interactions in the classroom. Collected data was used to make changes to course structure and activities in order to increase potential learning gains.

Analysis

We created rubrics to analyze the students’ responses to assessment questions. For example, the rubric shown in Table 1 was used to analyze responses to a conceptual question asking students to describe L’Hôpital’s Rule. We then used two-tailed pairwise comparisons ($\alpha = 0.05$) to compare groups of students (e.g., students who had previously viewed a more conceptual video about the mathematical content versus students who had viewed a more procedural video) or to compare pre- and post-assessment results. Written responses were also categorized so that we could view trends in the data.

Table 1.

Rubric Used for Scoring Responses to Conceptual L’Hôpital’s Rule Question

Score	Explanation
0	Answer was blank or made no mention of tangent lines.
1	Answers either lack "functions act like their tangent lines", or say something about tangent lines but neither "slope" nor "compare".
2	Answer states that functions act like their tangent lines near a point, and that one can find limits of $f(x)/g(x)$ (or compare $f(x)$ and $g(x)$) which have indeterminate forms by comparing the slopes of their tangent lines.

Students were also given in-class surveys consisting of Likert-scale and multiple-choice questions. The surveys generally asked students about their perceptions of the class structure and their learning gains. Their answers were categorized to look for trends in the responses. Aggregate quantitative data, such as scores on specific questions from class assessments, were also used as a measure of student learning gains.

Results

In this section we share a subset of results from our larger study. In particular, we share students’ overall opinions about video use and data on content learning for three individual

topics, including one class period specifically designed to help us see differences in the ways students learn conceptual and procedural content via video.

First, we share data on the students' beliefs about video usage. Several times throughout the semester, students were given surveys where they could voice their opinions about the structure of the class. When asked to compare learning a new topic outside of class via reading assignment versus watching a video, students overwhelmingly preferred videos (86%). However, when asked what part of their class structure had the greatest positive impact on their learning, 56% of students said the pre-class videos and readings, whereas 46% said the in-class activities and interactions.¹ We also asked the students to state their beliefs on how the videos increased both their conceptual understanding and computational skills in the class (see Table 2). For both questions, the majority of the class believed the videos greatly or significantly helped their mathematical understanding and skills, although more of the students found video helpful for their conceptual understanding than their computational skills.

Table 2.
Students' Beliefs About Video Usage

	Greatly	Significantly	Moderately	Slightly
Conceptual understanding	38%	38%	24%	0%
Procedural skills	20%	40%	30%	10%

The data suggests that the students believed the videos contributed to their content learning, but what objective evidence for learning gains can be seen in the students' work in the classroom? Prior to an in-class activity about L'Hôpital's Rule, we had the students watch an introductory video about the topic. However, we split the classes into two groups: one group watched a more conceptual video, and the other watched a more procedural video ($n = 23$ for each group). At the beginning of class, the students were given a content-driven assessment about L'Hôpital's Rule, with one question asking for a more conceptual explanation ("Describe how L'Hôpital's Rule works geometrically.") and the other asking for a more procedural explanation ("How does one calculate a limit using L'Hôpital's Rule?"). We then assigned the students to groups of two to three so that each group contained at least one student who had watched each video. We videotaped the class session to capture the students' interactions with and explanations to each other. At the end of class, students were given the same assessment as before to measure what changes in their understanding occurred due to their group discussions.

We scored their responses to the pre/post assessment using rubrics similar to the one described above (0–2 scale). The students' average results can be found in Table 3. The results indicate that students who watched the more conceptual video were able to answer the more conceptual question on the pre-class assessment, but were not able to answer the more procedural question. The opposite was true for the students who had watched the procedural video.

Table 3.
Average Scores on L'Hôpital's Rule Assessment

Group	Conceptual Question		Procedural Question	
	Pre	Post	Pre	Post

¹ Percentages add up to more than 100% because students could choose more than one answer.

Watched conceptual video	1.39	1.48	0.09	1.26
Watched procedural video	0.04	1.35	1.74	1.57
Significantly different? (<i>p</i> -value)	Yes $p < 0.001$	No $p = 0.210$	Yes $p < 0.001$	No (barely) $p = 0.057$
Effect Size (<i>r</i>)	0.839	0.119	0.909	0.235

After working with their peers, both groups of students were generally able to answer the conceptual and procedural questions. No statistically significant differences were found in the two groups' post-class assessment average scores. However, the difference in their post-assessment scores for the procedural question was just barely insignificant. (We will discuss this finding further in the next section.) These are preliminary analyses, but they seem to indicate that students gained mathematical knowledge from watching the videos and were able to share that knowledge with other students.

We also found that the learning gains from the L'Hôpital's Rule videos were similar to the gains from other videos in the class. For example, two of the videos the students watched covered the formal definition of the limit and the intermediate value theorem. After each, the students had an in-class activity to explore the topic in more depth. This was similar to how the students were introduced to L'Hôpital's Rule

After watching each of these two videos, students took post-video surveys. For the definition of the limit, they were asked to write the definition in their own words. For the intermediate value theorem, students were asked to explain the importance of assuming continuity in the statement of the theorem. We created rubrics (0–2 scale) and scored their responses on these items (see Table 4). The average score on the limit definition survey was 1.07, and the average score on the intermediate value theorem survey was 1.26. However, these scores may hide the range of solutions given by the students. For example, on the limit definition survey more than 75% of the students had at least some understanding of the limit definition. In both cases, students exhibited at least moderate content learning gains after only watching the videos.

Table 4.
Score Distribution for Two Surveys

Score	0	1	2
Limit Definition survey	18%	57%	25%
IVT survey	10%	43%	36%

For the limit definition and for L'Hôpital's rule, we gave the students an additional assessment question after they had watched the video and discussed the topic in class. (For the limit definition, this was an exam question, and for L'Hôpital's rule this was the post-activity assessment.) We found the students had similar results on these assessments, with an average score of 71% on the L'Hôpital's Rule question and 75% on the definition of limit question.

Last, we compared the students' solutions to specific questions on their final exam (see Table 5). We looked at the results from the final exam about three different topics from the course: two of which were introduced via video and one of which was introduced via a guided-inquiry activity, the second most common form of content delivery in this classroom. Topics in

this table are listed in the order they were covered in the course. The students' final exam scores on these three topics were virtually identical, which seems to indicate learning gains for these two methods of content delivery are nearly equivalent for these students.

Table 5.
Final Exam Results

Topic	Delivery Method	Percent
Definition of Limits	Video	76%
Definition of Derivatives	Guided-Inquiry Activity	78%
L'Hôpital's Rule	Video	75%

In summary, the results from our data imply that students are gaining at least some conceptual and procedural understanding of the mathematical content from video lessons. In the next section we discuss some possible implications of these results, along with areas of future research.

Discussion

Our study contains many different types of research data that may at first seem disconnected. This was our first attempt to quantify student learning gains in a flipped calculus classroom, specifically via video lectures. We collected data from a wide range of sources to help narrow down the research questions we wanted to explore in more depth in later studies. As such, this data may give us an initial overall picture of student learning gains through the semester but it does not allow us to go into depth for any one topic. However, we will use this data to help us design future studies to investigate the role that video lectures play in student learning in a flipped classroom.

In reviewing the results of the pre-activity assessment for the videos on L'Hôpital's Rule, we were not surprised by how well the students did on the question that related to the video they had viewed. However, more than 80% of the students in the class had taken at least one calculus class before, so we predicted that some students would initially be able to answer both questions successfully, which was not the case. Also, we were surprised by the students' improvement in both conceptual and procedural understanding after working in groups. Our results seem to indicate that students learned conceptual and procedural content from the videos and were able to share that knowledge effectively with their peers.

There are still some open questions from the data. The L'Hôpital's Rule post-activity assessment scores between the two groups of students on the procedural question were just barely insignificantly different and students felt the videos helped them more with conceptual knowledge than with learning procedures. This could mean we need to take into consideration what content educators deliver via video. However, because of the small number of students in this study, more research needs to be done to determine if there is a statistically significant difference in learning gains from more procedural videos than more conceptual ones.

One result that stands out is the low average on the post-video assessment on the limit definition. While in most other cases, the average scores were typically 65–75% on post-video assessments, in this case the average was just over 50%. As this topic is one of the most difficult and most conceptual of the entire course regardless of delivery method, this may not be as surprising as it initially seemed. It is also significant to note that this topic appeared early in the

course, and it is possible that students had not yet developed productive ways of interacting with the video.

There are two important caveats to our findings. First, we had no control group against which to compare our results from the flipped course. We can anecdotally compare to our prior experiences teaching Calculus I at other institutions and with our colleagues' experiences in other sections of this course, but the goal of this study was not to make comparisons. Instead, we wanted to investigate learning gains in the flipped classroom, although we do advocate for research comparing learning gains from different pedagogical techniques. Second, while we attempted to specifically investigate learning gains from videos watched outside the classroom, we must remember the videos were not used in isolation. In some of our data, such as exam data, the effect of the videos on the students' understanding is difficult to separate from the effects of the other learning activities that happen afterwards (e.g., in-class discussion and activities, homework, office hours, studying for exams). As we continue our research, we hope to be able to isolate the learning gains from videos and also investigate how the structure of the video-watching experience affects student learning, which we will discuss in more detail below.

Last, teachers thinking about using videos in their classes should know that students will at least get a basic understanding from videos, whether the videos are more conceptual or procedural. Moreover, some may be disappointed that our results indicate that introducing material via video does not necessarily improve learning gains as compared to learning via other methods. However, our data does seem to indicate that student content learning gains from video are at least equivalent to those from the other content delivery methods used in this course.

Implications for Future Research

While the results of this preliminary study seem to indicate that students can and do learn mathematical content from video lessons, our data have also opened other lines of future inquiry. For example, one might investigate what balance of conceptual and procedural videos should be used to have the greatest impact on student-learning gains, or the effect of video-recorded lessons on specific student demographics. Another possible avenue of investigation would be to determine the effect of this classroom structure on student communities of learning. Anecdotal evidence suggests that students may form cohorts within a flipped classroom that persist in future courses.

It is important to note that flipped classrooms do not consist solely of video lessons. Moving primary instruction out of the classroom creates time in class for students to clarify and reinforce content through discussion with peers and to actively participate in authentic problem-solving. This inversion allows instructors to be present while students engage with deep mathematical ideas, which is a more effective use of their knowledge and instructional abilities.

In future research we plan to focus specifically on the ways in which students interact with videos. We want to determine if they are actively engaging with the video lessons or passively listening as though they are in a lecture. We are also curious about their video-watching habits and what they do when they are confused during a video. One of our primary goals is to find ways to structure the video-watching experience to improve student learning, including helping them build mathematical integrity (knowing what you do and do not know about mathematics). If students can accurately assess what they do and do not understand from the videos *before* each class session, the instructor can more productively run the class session. The results of this research could provide instructors with ways to make videos more effective and help students interact with videos more productively.

References

- Bergmann, J. & Sams, A. (2012). *Flip your classroom: Reach every student in every classroom every day*. Eugene, OR: International Society for Technology in Education.
- Bishop, J. L. & Verleger, M. A. (2013). The flipped classroom: A survey of the research. In *American Society for Engineering Education National Conference Proceedings*, Atlanta, GA. Retrieved from <https://www.asee.org/public/conferences/20/papers/6219/view>
- Day, J. A. & Foley, J. D. (2006). Evaluating a web lecture intervention in a human-computer interaction course. *IEEE Transactions on Education*, 49(4), 420–431.
- Findlay-Thompson, S. & Mombourquette, P. (2014). Evaluation of a flipped classroom in an undergraduate business course. *Business Education & Accreditation*, 6(1), 63–71.
- Foertsch, J., Moses, G., Strikwerda, J., & Litzkow, M. (2002). Reversing the lecture/homework paradigm using eTEACH[®] web-based streaming video software. *Journal of Engineering Education*, 91(3), 267–274.
- Ford, P. (2015). Flipping a math content course for pre-service elementary school teachers. *PRIMUS: Problems, Resources, and Issues in Mathematics Undergraduate Studies*, 25(4), 369–380.
- 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. *Proceedings of the National Academy of Sciences*, 111(23), 8410–8415.
- Gaughan, J. E. (2014). The flipped classroom in world history. *The History Teacher*, 47(2), 221–244.
- Johnson, G. B. (2013). *Student perceptions of the flipped classroom* (Unpublished master's thesis). University of British Columbia, Okanagan, BC.
- Lape, N. K., Levy, R., Yong, D. H., Haushalter, K. A., Eddy, R., & Hankel, N. (2014). Probing the inverted classroom: A controlled study of teaching and learning outcomes in undergraduate engineering and mathematics. In *American Society for Engineering Education National Conference Proceedings*, Indianapolis, IN. Retrieved from <https://www.asee.org/public/conferences/32/papers/9475/view>
- Love, B., Hodge, A., Grandgenett, N., & Swift, A. W. (2014). Student learning and perceptions in a flipped linear algebra course. *International Journal of Mathematical Education in Science and Technology*, 45(3), 317–324.

- Mason, G., Shuman, T. R., & Cook, K. E. (2013). Inverting (flipping) classrooms – Advantages and challenges. In *American Society for Engineering Education National Conference Proceedings*, Atlanta, GA. Retrieved from <https://www.asee.org/public/conferences/20/papers/7171/view>
- Mayer, R. E., Hegarty, M., Mayer, S., & Campbell, J. (2005). When Static Media Promote Active Learning: Annotated Illustrations Versus Narrated Animations in Multimedia Instruction. *Journal Of Experimental Psychology: Applied*, 11(4).
- Mayer, R. E., Sobko, K., & Mautone, P. D. (2003). Social Cues in Multimedia Learning: Role of Speaker's Voice. *Journal Of Educational Psychology*, 95(2).
- Mazur, E. (1996). Are science lectures a relic of the past? *Physics World*, 9, 13–14.
- McGivney-Burrelle, J. & Xue, F. (2013). Flipping calculus. *PRIMUS: Problems, Resources, and Issues in Mathematics Undergraduate Studies*, 23(5), 477–486.
- Moore, A. J., Gillett, M. R., & Steele, M. D. (2014). Fostering student engagement with the flip. *The Mathematics Teacher*, 107(6), 420–425.
- 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.
- PCAST STEM Undergraduate Working Group. (2013). *Engage to Excel: Producing One Million Additional College Graduates with Degrees in Science, Technology, Engineering, and Mathematics*, eds Gates, S. J. Jr., Handelsman, J., Lepage, G. P., & Mirkin, C. (Office of the President, Washington).
- Roach, T. (2014). Student perceptions toward flipped learning: New methods to increase interaction and active learning in economics. *International Review of Economics Education*, 17, 74–84.
- Sharma, N., Lau, C. S., Doherty, I., & Harbutt, D. (2015). How we flipped the medical classroom, *Medical Teacher*, 37(4), 327–330.
- Strayer, J. F. (2007). *The effects of the classroom flip on the learning environment: A comparison of learning activity in a traditional classroom and a flip classroom that used an intelligent tutoring system* (Unpublished doctoral dissertation). Ohio State University, Columbus, OH.
- Wilson, S. G. (2013). The flipped class: A method to address the challenges of an undergraduate statistics course. *Teaching of Psychology*, 40(3), 193–199.
- Zappe, S. E., Leicht, R. M., Messner, J. I., Litzinger, T., & Lee, H. W. (2009). “Flipping” the classroom to explore active learning in a large undergraduate course. In *American Society for Engineering Education National Conference Proceedings*, Austin, TX.