

# A Course-Embedded Comparison of Instructor-Generated Videos of Either an Instructor Alone or an Instructor and a Student

Katelyn M. Cooper,<sup>†</sup> Lu Ding,<sup>‡</sup> Michelle D. Stephens,<sup>†</sup> Michelene T. H. Chi,<sup>‡§</sup> and Sara E. Brownell<sup>†§\*</sup>

<sup>†</sup>Biology Education Research Lab, School of Life Sciences, and <sup>‡</sup>Learning and Cognition Lab, Mary Lou Fulton College of Education, Arizona State University, Tempe, AZ 85281

## ABSTRACT

Instructor-generated videos have become a popular way to engage students with material before a class, yet this is a relatively unexplored area of research. There is support for the use of videos in which instructors tutor students, but few studies have been conducted within the context of a classroom. In this study, conducted in a large-enrollment college physiology course, we used a randomized crossover design to compare the impact of two types of instructor-generated videos that students watched as part of their preclass assignments. We compared videos featuring only an instructor (instructor-only videos) with videos featuring an instructor tutoring a student (instructor-tutee videos). We analyzed student survey responses and weekly physiology quiz scores and found that students preferred, enjoyed, and valued the instructor-only videos significantly more than the instructor-tutee videos. In contrast to prior literature, students with a grade point average (GPA) below the median (3.49) performed significantly better on physiology quizzes after watching instructor-only videos compared with instructor-tutee videos. Students with a GPA at or above the median performed equivalently on physiology quizzes after watching instructor-only or instructor-tutee videos. We present this study as an example of bringing cognitive science studies into the context of a real physiology classroom.

## INTRODUCTION

Active learning has been shown to improve student learning and reduce failure rates compared with traditional lecture (Freeman *et al.*, 2014). As such, college biology classes are increasingly transitioning from traditional lecture to active learning, wherein students engage in constructing their own knowledge during class (American Association for the Advancement of Science, 2011). One way of creating active learning classrooms is to “flip” the class so that students are expected to learn content on their own before class and then apply the concepts during in-class active-learning activities in which they solve complex problems and work with peers (Brame, 2013; Jensen *et al.*, 2015). While some classrooms are completely flipped and all content attainment happens outside class, many classrooms have a partially flipped structure in which instructors require students to learn some content before class and then deliver additional content during class and integrate active-learning activities that allow students to practice applying content to new problems. While this is not the only way to make classrooms more student centered, several studies have shown that this flipped-classroom model can improve student learning (Freeman *et al.*, 2007) and decrease achievement gaps between students (Freeman *et al.*, 2011; Eddy and Hogan, 2014). However, there is a need for greater understanding of what particular aspects of flipped classrooms lead to student learning gains, and *why* flipped classrooms lead

Cynthia Brame, *Monitoring Editor*

Submitted January 2, 2018; Revised February 28, 2018; Accepted March 7, 2018

CBE Life Sci Educ June 1, 2018 17:ar31

DOI:10.1187/cbe.17-12-0288

\*Address correspondence to: Sara E. Brownell (sara.brownell@asu.edu).

<sup>§</sup>These senior authors contributed equally.

© 2018 K. M. Cooper *et al.* CBE—Life Sciences Education © 2018. The American Society for Cell Biology. This article is distributed by The American Society for Cell Biology under license from the author(s). It is available to the public under an Attribution–Noncommercial–Share Alike 3.0 Unported Creative Commons License (<http://creativecommons.org/licenses/by-nc-sa/3.0>).

“ASCB®” and “The American Society for Cell Biology®” are registered trademarks of The American Society for Cell Biology.

to increased student learning gains is a growing area of research (Gross *et al.*, 2015; Jensen *et al.*, 2015).

In partially flipped and fully flipped classrooms, students commonly learn content on their own before class by reading the textbook or articles, watching animations, or viewing recorded lectures. Few studies have compared how different methods of content delivery may affect students' attitudes and performance. However, in one study of students enrolled in an introductory college biology class, researchers used a two-by-two study design and compared the effects of video preclass assignments with textbook preclass assignments and the effect of in-class lecturing versus an active learning-style class with student problem solving. The researchers found that, for students in the active-learning condition, there was no difference in exam performance between students who were assigned to read the textbook before class and students who were assigned to watch a video (Stockwell *et al.*, 2015). However, students who were assigned to watch the videos were significantly more satisfied with their preclass assignment than students who were assigned to read the textbook. In another study, which solely explored student perceptions of instructor-generated videos that were used to introduce weekly topics to students in a graduate-level research methods course, students reported that the videos improved their understanding of the material and helped them to feel closer to the instructor (Rose, 2009). These studies and others have led to recommendations for using videos as an educational tool to help prepare students for class (Rackaway, 2012; Stockwell *et al.*, 2015). However, if an instructor is interested in using videos as part of a preclass assignment, how should these videos be designed and implemented?

In the literature, tutees learning from tutors is considered the gold standard, in that it exceeds all other forms of instruction in helping students achieve learning gains (Bloom, 1984; VanLehn, 2011; Wood and Tanner, 2012). Tutoring has been shown to have effect sizes ranging from 0.79 to 2.0 SD (Bloom, 1984; VanLehn, 2011). On the basis of this literature, Chi and colleagues (2008) have argued that capturing tutor-tutee dialogue in videos (hereafter referred to as "instructor-tutee videos") and re-using it for students to watch would result in greater learning for the student watching the videos (observing student) than only watching a video of an instructor alone (hereafter referred to as "instructor-only videos"). We present three distinct reasons for including a tutee in instructional videos to enhance observing student learning. 1) The tutee can provide a zone of proximal representation match for the observing students; that is, the tutee can help break down the instructor/tutor's cognitive processes during the video. Research shows that experts usually cannot accurately gauge novice learners' levels of understanding, which limits their ability to scaffold concepts to maximize student learning (Chi *et al.*, 2004), and previous studies have shown that students learn more when problems are attempted by novices instead of experts (Schunk *et al.*, 1987; Monaghan and Stenning, 1998). Thus, videos that include tutees may help translate difficult concepts in ways that are more accessible for the observing students than videos featuring only an expert instructor. 2) Observing students' reaction to cognitive conflict or tutees struggling with content may enhance student learning. Specifically, a tutee's expression of a miscon-

ception followed by a tutor's correction of the misconception may enhance observing students' learning because previous studies have shown that, when observing students experience tutors and tutees expressing different ideas, they tend to engage in knowledge-building when trying to think through why the tutee has a different understanding than the tutor (Chan *et al.*, 1997; Chi *et al.*, 2017). 3) The tutee provides a model for learning, which can influence observing student learning. The observing students may identify with the tutee in the video and consequently mimic the tutee's learning skills. Previous research has shown that observing students can learn skills and behaviors to improve learning, such as asking questions, from tutees in the video (Craig *et al.*, 2000; Ryokai *et al.*, 2003). Thus, if students employ the learning skills they learn from the tutees, they may learn more after watching a video of a tutor and tutee compared with a video featuring only a tutor.

Cognitive science studies have supported the hypothesis that students learn from watching other students being tutored. For example, in two separate studies, observing students who were watching a video to learn about a particular topic learned as well as the individual tutees in the videos; this result was demonstrated in two different science, technology, engineering, and mathematics (STEM) domains: learning to solve physics problems and learning to explain diffusion in chemistry (Chi *et al.*, 2008; Muldner *et al.*, 2014). Additionally, several other laboratory studies have found that, for observing students, videos in which a student is being tutored can be more beneficial for learning than videos featuring only an instructor (Driscoll *et al.*, 2003; Fox Tree, 1999). While these findings suggest that instructors should create instructor-tutee videos to enhance student learning, these prior studies documenting that students learn more when watching instructor-tutee videos compared with instructor-only videos were situated in a controlled laboratory setting that was removed from the context of a real class. To our knowledge, no studies have been conducted over consecutive weeks in a college classroom.

However, there have been recent calls for more collaborations between cognitive scientists and discipline-based education researchers to see whether lab-based cognitive science findings can be replicated in a formal class environment and to see what additional information we can glean from applying these theories to real classrooms (McDaniel *et al.*, 2017; Mestre *et al.*, 2018). College classrooms have additional variables that are not accounted for in controlled laboratory experiments, including the multiple demands on students from other courses and outside-school commitments that affect the time a student can dedicate to a task, as well as student motivation to complete a task. Thus, we need to conduct both lab-based experiments and classroom-based experiments to fully understand a phenomenon and make recommendations to instructors.

The primary goal of this study was to build on previous laboratory findings and compare the impact of instructor-tutee videos and instructor-only videos on student performance in the context of a large-enrollment, upper-division, active-learning, college physiology course. Further, to our knowledge, no studies have explored student affect toward instructor-tutee and instructor-only videos, so we also probed what types of videos students prefer and why. Specifically, our research questions were:

- To what extent do students appreciate instructor-generated videos and why?
- What do students identify as advantages and disadvantages of instructor-only videos and instructor-tutee videos?
- Which type of video—instructor-only videos or instructor-tutee videos—do students prefer and why?
- To what extent do students perform differently on weekly physiology quizzes when watching instructor-only videos compared with instructor-tutee videos?

## METHODS

This study was conducted in the context of a large-enrollment, upper-division physiology course with 280 students at a large southwestern research university taught in Fall 2017. The study was piloted in the same physiology course in Fall 2016, and all instruments, including quizzes and surveys, were revised before being deployed again in Fall 2017. The overall structure and content of the Fall 2016 and Fall 2017 courses were identical. All data reported in this article are from the Fall 2017 study.

### Course Description

The physiology course met in person 3 days per week: Tuesday, Thursday, and Friday. On Tuesday and Thursday, all students attended the same 50-minute class held in a large lecture-style auditorium. On Friday, students attended one of three recitation sections, which were 75 minutes long with ~90 students each. All classes on Tuesday, Thursday, and Friday were taught in an active-learning way. During every class, students frequently engaged in group work while answering clicker questions or completing an in-class worksheet.

Students were required to complete preclass assignments before each class. The goal of the preclass assignments was primarily to provide students with a foundation of material so that the instructor could build upon this information during class, and students could use the information while problem solving in the classroom (O'Flaherty and Phillips, 2015). Before the Tuesday and Thursday sessions, students were asked to read sections of the textbook or popular science articles or to watch short videos that were not developed specifically for the course (e.g., animations about physiological processes, videos of patients describing their symptoms). Outside class, through the Mastering Physiology portal (Pearson, New York), students answered online instructor-generated questions or questions from a question bank about what they had read or watched. Before each Friday recitation, students were required to watch an instructor-generated video (see *Video Intervention and Assessment Development* section for a more specific description of the videos). While watching each of these instructor-generated videos, students were asked to complete a worksheet that aligned with the problems being discussed in the video. At the beginning of the Friday recitation, students turned in their completed video worksheets and took a quiz focusing on the content covered in the video. Together, the video, video worksheet, and quiz were considered students' preclass assignment for Friday.

### Instructor-Generated Videos: Instructor-Only Videos and Instructor-Tutee Videos

The instructor of the course created two different sets of videos for each week to prepare students for Friday's recitation: instructor-only videos, videos that teach physiology with only

the instructor of the course present (Figure 1, A and B), and instructor-tutee videos, videos that teach physiology with the instructor of the course tutoring a former physiology student (Figure 1, C and D). The videos were not lectures but were recorded problem-solving exercises in which the instructor or the student being tutored by the instructor worked through five to seven physiology problems. Both sets of videos used the same physiology problems, and the videos were paired with the same worksheet that students were required to complete while watching the videos (see the Supplemental Material for an example worksheet). As a result, both sets of videos covered the same physiology content. All videos were professionally made in collaboration with instructional designers using a green screen or green Chromakey. PowerPoint slides with the physiology content and physiology problems were displayed in the background of the video, and the instructor or student could write or draw on the slides using a Wacom Tablet. The annotations and slides were recorded with Camtasia, and the files were put into Adobe Premier, combined with the raw studio video, and edited for audiovisual effects, so the images of the instructor or instructor and student were projected onto the PowerPoint slides (Figure 1).

### Experimental Design

To determine whether students learned more after watching the instructor-only videos or the instructor-tutee videos, we used a randomized crossover design. Students in the physiology course were randomized into group A or group B upon enrolling in the course. Group A watched instructor-only videos during weeks 1–4 of the semester and then watched instructor-tutee videos during weeks 5–8. Conversely, group B watched instructor-tutee videos during weeks 1–4 of the semester and watched the instructor-only videos during weeks 5–8. Only the specific video that a student was assigned to watch was automatically uploaded to the student's personal Blackboard account, so the student watched the correct type of video each week (see *Instructor-Tutee Videos* and *Instructor-Only Videos* subsections below for more specific information about the videos).

### Procedures

Every week, students in groups A and B were required to fill out the same video worksheet while they watched either an instructor-only video or an instructor-tutee video as part of their preclass assignment. Students received half of their Friday preclass assignment points for turning in the completed video worksheet at the beginning of the Friday recitation; video worksheets were graded based on completion, not accuracy. At the beginning of every Friday recitation, students in both groups completed the same video quiz after turning in their video worksheet. The physiology quizzes, which were graded on accuracy, made up the other half of students' Friday preclass assignment points.

To capture student opinions about the videos, we gave students a survey at the end of week 4, after they had watched one type of video, and another survey at the end of week 8, after they had watched both types of videos (see *Survey of Student Demographics and Opinions about the Videos* subsection for more specific information). Students were awarded a small number of course points for completing the surveys. See Figure 2 for a depiction of the experimental design over the 8 weeks.

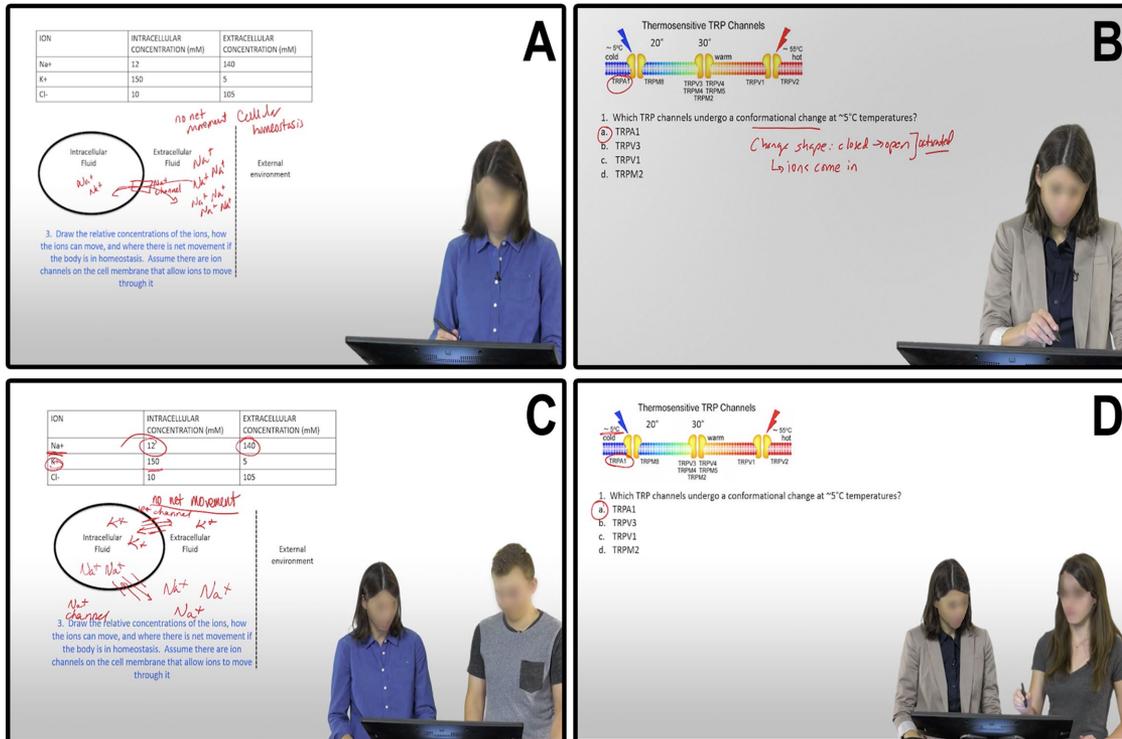


FIGURE 1. Screen capture of instructor-only videos (A, B), in which the instructor alone was present, and instructor-tutee videos (C, D), in which the instructor appeared with a student. In both sets of videos, slides with content information and physiology problems appeared in the background. The Wacom Tablet allowed for the instructor’s annotations to appear on the slide so that observing students who were watching the video could see the instructor’s writing. In instructor–tutee videos, both the instructor and the student annotated the slides on the screen.

### Instructor–Tutee Videos

In the instructor–tutee videos, the instructor tutored a student during the video. All four students who appeared in the videos (referred to as “tutees” from here forward) had completed the same physiology course in Fall 2015 and thus were familiar with the content of the course. However, the tutees were not primed about what specific content they would be asked about in the video so that they would not study the content beforehand and would more authentically struggle with the problems. Of the tutees, two were male and two were female. Two students identified as white, one student identified as Black, and one student identified as white and Asian. The instructor of the course considered all of the tutees to be above-average students; three of the students had received an “A” in the physiology course, and one student had received a “B.” Only one tutee interacted with the instructor in each video. There were eight instructor–tutee videos, one for each week of the experiment. Each of the four tutees appeared in two of the eight videos, so that the observing students in both group A and group B watched a video with each of the four tutees.

In the instructor–tutee videos, the instructor would first introduce the tutee, then the instructor would pose a physiology question and allow a few minutes for the tutee and the physiology students watching the videos (observing students) to think about how to answer the question. The tutee would then attempt to solve the problem, and the instructor would ask guiding questions so that the student fully elaborated on his or

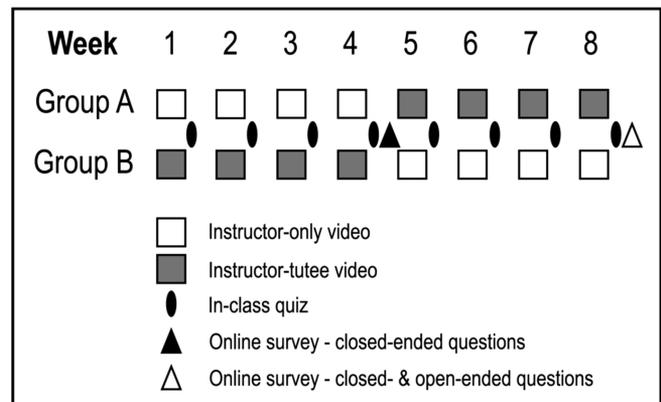


FIGURE 2. Depiction of experimental design. During weeks 1–4, students in group A watched instructor-only videos and students in group B watched instructor–tutee videos. During weeks 5–8, students in group A watched instructor–tutee videos and students in group B watched instructor-only videos. At the beginning of every Friday recitation, both groups of students completed the same in-class quiz about the content presented in the respective Friday videos. At the end of week 4, after watching one type of video, and again at the end of week 8, after watching the other type of video, students completed an online survey about their opinions of the type of videos that they had recently watched, either instructor-only videos or instructor–tutee videos. The survey at the end of week 8 also asked students open-ended and closed-ended questions about the videos.

her thoughts. Simultaneously, the tutee would write out his or her thoughts, and this writing was displayed on the PowerPoint slide. The tutees attempted five to seven physiology problems in each video. After the tutee attempted a problem, the instructor corrected any misconceptions brought up by the tutee and articulated or elaborated on the correct solution to the physiology problem. The instructor–tutee videos ranged from 16 to 27 minutes and averaged 21 minutes.

### Instructor-Only Videos

Instructor-only videos featured only the instructor of the course. Similar to the instructor–tutee videos, the instructor would guide the observing students through the same five to seven problems that were presented in that week’s instructor–tutee videos and give the observing students a few minutes of think-time to consider how to answer a problem before working through the problem by writing and talking out the answer. The instructor-only videos ranged from 12 to 21 minutes and averaged 17 minutes. An independent *t* test revealed that the instructor-only videos were significantly shorter than the instructor–tutee videos ( $t = 2.3393, p = 0.0347$ ).

### Video Worksheets and Quizzes

In both the instructor-only videos and the instructor–tutee videos, the instructor posed a set of physiology problems and observing students simultaneously completed a worksheet. The worksheet had a series of question numbers on it; none of the problems was printed on the worksheet, but they were instead presented in the videos to encourage students to engage with the videos and not complete the worksheet independent of the videos. Regardless of whether an observing student was watching an instructor-only video or an instructor–tutee video during a particular week, all students completed the same worksheet, because the physiology questions in each type of video were the same. Worksheets were distributed to students at the end of class on Thursdays, and a copy of the worksheet was posted on Blackboard in case the student did not attend Thursday class or lost the worksheet. Students were encouraged to watch the videos during the time between Thursday class and Friday class, but videos were posted on Tuesdays, because some students needed a longer time frame to watch the videos. Once a video was posted, students had access to that video until the end of the semester.

Students turned in the video worksheets at the beginning of the Friday recitations. All worksheets were graded for completion. At the beginning of each Friday recitation, immediately after turning in the video worksheets, all students were asked to individually complete a 10- to 12-item quiz. The quizzes were the same for all students (groups A and B), because the content in both sets of videos was the same. Quizzes were created by the course instructor based on the content that was presented in the corresponding videos. Some quiz questions were taken from existing concept inventories, but most of the questions were written by the instructor so that the questions would be specific to the course material presented in the video. The quiz questions were piloted with students in Fall 2016, and we observed a ceiling effect in student scores on the quizzes. We revised the quiz questions to make them more difficult and added questions for Fall 2017 that were intended to be more challenging. On the basis of student feedback, we also removed

a few quiz questions that students interpreted differently than we intended. Final versions of quiz questions for Fall 2017 were reviewed by a member of the research team (M.D.S.) to check for clarity before deployment. All quizzes were graded for accuracy, and students were not allowed to use any resources during the quiz.

The quizzes were used to measure student learning; the crossover design allowed us to compare all students’ scores on quizzes that they completed after watching instructor-only videos with their scores on quizzes that they completed after watching instructor–tutee videos (see the *Analyses* subsection for additional information). Thus, it was important to ensure that the quizzes in weeks 1–4 were not more or less cognitively challenging than the quizzes in weeks 5–8. Using the Blooming Biology Tool (Crowe *et al.*, 2008), two researchers calculated the cognitive difficulty based on Bloom’s taxonomy (Bloom *et al.*, 1956). Questions testing knowledge or comprehension were categorized as low level; questions testing application, analysis, or evaluation were categorized as high level. The researchers had an interrater reliability of 85.5% before discussing and coming to consensus on all questions. There were no significant differences between the number of low-level Bloom’s questions or high-level Bloom’s questions on the quizzes given to students during weeks 1–4 or 5–8 (weeks 1–4: 47.6% of questions were low level, and 52.4% were high level; weeks 5–8: 45.7% of questions were low level, and 54.3% of questions were high level;  $p = 0.1847, Z = 0.85716$ ).

### Survey of Student Demographics and Opinions about the Videos

A survey consisting of Likert-scale questions was administered to students after week 4 and a survey consisting of Likert-scale and open-ended questions was administered to students after week 8 of the course (see the Supplemental Material for copies of the surveys). Both surveys asked students about their perceived usefulness of the videos and the extent to which they enjoyed watching the videos. Five Likert-scale items measuring student-perceived usefulness and five Likert-scale items measuring student-perceived enjoyment of watching the videos were adapted from the Intrinsic Motivation Inventory (IMI; Ryan, 1982); each of them was rated from 1 (not at all true) to 7 (very true). IMI has been widely used in social studies, and its validity has been well established (McAuley *et al.*, 1989). Items were slightly reworded to reflect the video context of this study. Table 1 presents examples of the original IMI items and the reworded items used in the current study for the two constructs. Reliabilities (Cronbach’s  $\alpha$ ) of the two constructs in the current study were at an acceptable level. Cronbach’s  $\alpha$  values are shown in Table 1.

On the survey given to students in week 8, after they had watched both types of videos, students answered a Likert-scale question about the extent to which they appreciated the instructor-generated videos and a follow-up open-ended question about why they did or did not appreciate the videos. Additionally, students answered open-ended questions about the advantages and disadvantages of *both* instructor-only and instructor–tutee videos and the type of video they preferred and why. To understand how the instructor-generated videos compared with other ways of introducing content to students before class, we also asked students to indicate how valuable (1 = lowest value,

**TABLE 1. Sample items and reliability of survey used to measure student-perceived usefulness of the videos and the extent to which students enjoyed watching the videos**

| Construct  | Sample item   | Reliability     |                  |
|------------|---|-----------------|------------------|
|            |   | Instructor-only | Instructor-tutee |
| Usefulness | <i>Original item</i><br>I think this is an important activity.                        | 0.879           | 0.869            |
|            | <i>Reworded item</i><br>I think watching the Friday videos was an important activity. |                 |                  |
| Enjoyment  | <i>Original item</i><br>I would describe this activity as very interesting.           | 0.899           | 0.902            |
|            | <i>Reworded item</i><br>I would describe watching the Friday videos as interesting.   |                 |                  |

5 = highest value) they perceived the different types of preclass assignments were for their learning. Additionally, student demographic information was collected, including gender, race/ethnicity, and college grade point average (GPA). The surveys were initially piloted to students in the physiology class during the Fall 2016 semester, and questions were revised for Fall 2017.

### Analyses

*Student Appreciation of the Videos, Advantages/Disadvantages of Video Types, and Student Video Preference.* We intended to conduct paired-sample *t* tests to compare the extent to which students valued watching the instructor-generated videos outside class compared with reading the textbook or completing other types of reading such as reading popular news articles. However, the assumption of normality was violated for paired *t* tests; therefore, three nonparametric Wilcoxon signed-rank tests were carried out instead.

Three authors (K.M.C., M.D.S., and S.E.B.) reviewed all student responses to open-ended survey questions about why students appreciate instructor-generated videos and the advantages and disadvantages of instructor-only and instructor-tutee videos. Using open-coding methods, the authors identified common themes in student responses and created a rubric to code each question (Strauss and Corbin, 1990). The authors used constant-comparison methods (Glesne and Peshkin, 1992) to develop the rubric; quotes were assigned to a category and were continuously compared to ensure that the description of the category was inclusive of all quotes and that student quotes were not different enough from one another to warrant a different category. See the Supplemental Material for a coding rubric for each question. After all rubrics were established, two authors (M.D.S. and K.M.C.) coded 25% of student responses for each question, and the interrater reliability for each question ranged from 92 to 98%. One author (M.D.S.) coded the remaining responses for each question.

We used content analysis to code students' reasoning for why they preferred a specific type of video—either instructor-only or instructor-tutee (Bogdan and Biklen, 1998). To code students' reasoning for why they preferred instructor-only videos, we used the categories that emerged from analyzing student-reported advantages of instructor-only videos and student-reported disadvantages of instructor-tutee videos. Similarly, to code students' reasoning why they preferred instructor-tutee videos, we used the categories that emerged from analyzing student-reported advantages of instructor-tutee vid-

eos and student-reported disadvantages of instructor-only videos. For each preference question, two authors (K.M.C. and M.D.S.) reviewed a subset of student responses together using the respective previously established rubrics for student-reported advantages and disadvantages. Then, the authors independently reviewed 25% of student responses. Their interrater reliability was 94% for the question about why students preferred instructor-only videos and 98% for why students preferred instructor-tutee videos. One author (M.D.S.) coded the remaining responses for each question. We report out themes that were identified by at least 10% of students for each question unless otherwise noted.

*Student-Reported Usefulness and Enjoyment of the Videos and Performance on Video Quizzes.* To compare the effect of instructor-only videos and instructor-tutee videos on the student-reported usefulness and enjoyment and student performance, the data from student surveys about perceived usefulness and enjoyment of each type of video and student performance on all eight video quizzes were reorganized into two data sets: instructor-only and instructor-tutee. That is, group A survey data from weeks 1–4 and group B survey data from weeks 5–8 were combined and renamed as instructor-only video survey data. Group A survey data from weeks 5–8 and group B survey data from weeks 1–4 were combined and renamed as instructor-tutee video survey data. The same reorganization was applied to the students' performance data on the physiology quizzes.

Before formal analyses were conducted, data sets were screened and modified for missing values. Within the instructor-only video group, 27.2% of students' performance data and 14.3% of survey data were missing; within the instructor-tutee video group, 29.5% of students' performance data and 15.7% of survey data were missing. Although this was not surprising, given the multiweek design of our study with multiple data points and a missing rate of 15–20% being common in educational studies (Enders, 2003), failing to treat missing values properly may result in biased, unreliable results and insufficient power of the study (Rubin, 1987). Little's MCAR test (1988) was performed, and the results indicated that the missing values in the data sets were missing completely at random. Multiple imputation was used to impute missing values, because it provides robust estimates when the data are missing completely at random (Schafer and Graham, 2002). For each data set, five imputations were generated, and the results were pooled from the new imputed data sets. That is, each missing value was

replaced by a list of five estimated values. Each of the five data sets was then analyzed by the same analysis method, and the pooled results were generated by the combined results from all the data sets.

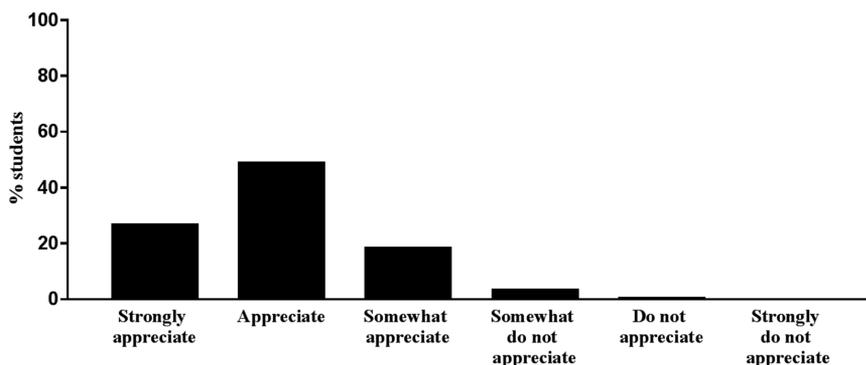
The assumption of normality was violated for paired-sample *t* tests, and thus the alternative nonparametric test, the Wilcoxon signed-rank test, was conducted to compare student-perceived usefulness and enjoyment after they watched the instructor-only videos and the instructor-tutée videos and to compare student performance on the quizzes. Because research has shown that, compared with high-performing students, low-performing students tend to benefit more from engaged-learning activities (Carini *et al.*, 2006; Freeman *et al.*, 2011), we divided students into two groups—lower-achieving students with a prior GPA below the median course prior GPA ( $GPA < 3.49$ ) and higher-achieving students with a GPA at or above the median ( $GPA \geq 3.49$ ). A significant interaction effect was observed between student achievement and performance after watching the two types of videos ( $p = 0.014$ ). Therefore, we investigated whether there was any difference in student performance within the lower- and higher-performing groups after they watched the instructor-tutée and instructor-only videos, respectively. Two Wilcoxon signed-rank tests were carried out to examine student learning gains. All analyses were performed using the IBM Statistical Package for the Social Sciences (SPSS) 24.0.

This study was done in accordance with an approved IRB.

## RESULTS

### Student Demographics

Out of the 280 students enrolled in the physiology course, 217 students consented to participate in the study (77.5%). Of the students who consented to participate in the study, 37.4% were male and 62.6% were female. Students identified as either American Indian (1.5%), Asian (19.5%), African American (4.1%), Latino/a (16.4%), Pacific Islander (2.1%), or white (51.8%), and 4.6% of students identified with a race or ethnicity other than those listed. Students' prior college GPAs ranged from 1.69 to 4.0, with a mean of 3.39 and a median of 3.49. There were 53 students in group A with a GPA lower than 3.49, and 61 students with a GPA greater than or equal to 3.49.



**FIGURE 3.** Extent to which students appreciate instructor-generated videos. Twenty-seven percent of students strongly appreciated the videos, 49.3% of students appreciated the videos, and 18.8% of students somewhat appreciated the videos. Nearly 4% of students somewhat did not appreciate the videos, and 1.0% of students did not appreciate the videos. No student reported that he or she strongly did not appreciate the videos.

**TABLE 2.** Average student value for different types of preclass assignments<sup>a</sup>

| Preclass assignment                            | Mean and SD         |
|--|---------------------|
| Instructor-generated videos                    | M = 3.90, SD = 0.99 |
| Other types of reading (such as news articles) | M = 2.73, SD = 1.28 |
| Textbook reading                               | M = 2.26, SD = 1.32 |

<sup>a</sup>Students rated the extent to which they valued different types of preclass assignments from 1 (lowest value) to 5 (highest value). Students rated instructor-generated videos ( $M = 3.90$ ,  $SD = 0.99$ ) significantly higher than other reading ( $M = 2.73$ ,  $SD = 1.28$ ,  $Z = -8.73$ ,  $p < 0.001$ , Cohen's  $d = 0.76$ ) or textbook reading ( $M = 2.26$ ,  $SD = 1.32$ ,  $Z = -9.84$ ,  $p < 0.001$ , Cohen's  $d = 1.00$ ). Students also significantly favored other reading over textbook reading ( $Z = -3.88$ ,  $p < 0.001$ , Cohen's  $d = 0.30$ ).

Within group B, 54 students had a GPA lower than 3.49 and 49 students had a GPA greater than or equal to 3.49. There was no significant difference in demographic composition (gender, race/ethnicity, or college GPA) of the consenting students in the randomized groups A and B (see the Supplemental Material for a breakdown of student demographics in each group).

### Finding 1: Students Appreciate Instructor-Generated Videos

In response to a closed-ended, Likert-scale survey question, students rated the extent to which they perceived different types of preclass assignments were valuable to their learning. Students rated three types of assignments from the physiology course—textbook reading, other types of reading (such as popular news articles), and watching instructor-generated videos—on a Likert scale from 1 (lowest value) to 5 (highest value). Students rated instructor-generated videos ( $M = 3.90$ ,  $SD = 0.99$ ) significantly higher than either other types of reading, including news articles ( $M = 2.73$ ,  $SD = 1.28$ ,  $Z = -8.73$ ,  $p < 0.001$ , Cohen's  $d = 0.76$ ), or textbook reading ( $M = 2.26$ ,  $SD = 1.32$ ,  $Z = -9.84$ ,  $p < 0.001$ , Cohen's  $d = 1.00$ ). Students also significantly favored other types of reading compared with textbook reading ( $Z = -3.88$ ,  $p < 0.001$ , Cohen's  $d = 0.30$ ; Table 2).

We were also interested in exploring to what extent students appreciated the instructor-generated videos broadly. In response to a Likert-scale question on the survey, the majority of students (95.2%) indicated that they appreciated the instructor-generated videos (Figure 3). Students were asked to explain in an open-ended survey question why they either appreciated or did not appreciate the instructor-generated videos. There were too few responses ( $n = 10$ ) for why students did not appreciate the videos to code into specific categories. The primary reason that students appreciated the instructor-generated videos was that students perceived that the videos improved their understanding of physiology (37.6% of responses). Additionally, 28.4% of students provided broadly positive comments about the videos, such as the videos were helpful and informative or the videos were a valuable teaching tool.

## Finding 2. Advantages and Disadvantages of Instructor-Only and Instructor–Tutee Videos

Students were asked about the advantages and disadvantages of instructor-only and instructor–tutee videos in four open-ended survey questions: 1) What are the advantages of instructor-only videos? 2) What are the disadvantages of instructor-only videos? 3) What are the advantages of instructor–tutee videos? 4) What are the disadvantages of instructor–tutee videos? Coding of student responses to each of these questions identified four advantages and one disadvantage of the instructor-only videos (Table 3) and four advantages and four disadvantages of the instructor–tutee videos (Table 4).

Interestingly, the student-identified advantages of the instructor-only videos were opposites of the four student-identified disadvantages of the instructor–tutee videos. 1) Nearly 20% of students highlighted that an advantage of the instructor-only videos was that they were short or went by quickly, and 32.4% of students reported that the instructor–tutee videos felt as though they went by slowly or took longer to watch. However, this is not particularly surprising, given that, on average, the instructor–tutee videos were 4 minutes and 31 seconds longer than the instructor-only videos. 2) Students highlighted that the information that was presented in the videos could either be an advantage or a disadvantage depending on whether the information was correct. We found that 12.6% of students

reported that an advantage of the instructor-only videos was that the instructor presented only accurate information about physiology. However, in the instructor–tutee videos, the tutees did not always have a completely accurate understanding of the physiology content and would sometimes answer the physiology questions incorrectly, which 26.1% of students highlighted as a disadvantage. 3) Nearly half of the students (47.8%) emphasized the straightforward or direct presentation of information as an advantage of the instructor-only videos, while 19.3% of students described the instructor–tutee videos as presenting information in an indirect or disorganized way, which they viewed as a disadvantage. Specifically, students noted that, in instructor–tutee videos, a tutee would sometimes incorrectly describe his or her answer to a physiology problem, but the observing student did not immediately recognize the response as incorrect. It was only when the instructor corrected the tutee’s response that the observing student learned that the tutee’s answer was inaccurate. Some students recounted that they felt misled or later had trouble remembering which of the two answers was correct. 4) Students also stated that having the information presented in an indirect way or watching a student provide incorrect information in the instructor–tutee videos led them to feel confused. In fact, 27.5% of students noted that the instructor–tutee videos were confusing or confused their understanding of physiology content. However, this

TABLE 3. Student-perceived advantages and disadvantages of instructor-only videos<sup>a</sup>

| Code  | All students<br>( <i>n</i> = 207) | Description  | Example quote  | Example quote   |
|---|-----------------------------------|--|--|---|
| <b>Advantages of instructor-only videos</b>             |                                   |  |  |   |
| Presents information in a straightforward or direct way | 47.8%                             | Student indicates that presentation of information is straightforward, direct, or organized.   | “The videos were straightforward and to the point.”  | “Direct and clear presentation of processes and concepts”   |
| The video is short                                      | 19.8%                             | Student indicates that the instructor-only video is shorter or goes by faster than the instructor–tutee videos.  | “[The instructor-only video] was shorter which is nice.”   | “The [instructor-only] videos are shorter.”   |
| Improves observing student understanding of content     | 14.0%                             | Student indicates that, as a result of watching the instructor-only video, his or her learning or understanding of the concept improved.   | “[The instructor-only video] helped me thoroughly understand the concepts presented in class.”   | “I felt like I was better able to retain the information [after watching the instructor-only videos].”  |
| The information presented is correct                    | 12.6%                             | Student indicates that only correct information is provided in the videos.   | “There is very little room for miscommunication and inaccurate information.”   | “[The instructor-only videos] rarely had wrong information to throw you off about what the right answer is.”  |
| <b>Disadvantages of instructor-only videos</b>          |                                   |  |  |   |
| Is missing a student’s perspective                      | 18.8%                             | Student indicates that a student’s perspective is missing from the videos. Thus, observing students’ questions are not answered or misconceptions, mistakes, or incorrect thought processes are not corrected. | “[The instructor] doesn’t work through the problems commonly had by students, only how a seasoned biologist sees the problems (which we are not).” | “[The instructor-only videos] are without student perspective, so there is a lot of assumption of knowledge that might occur since [the instructor] is talking on her own.” |

<sup>a</sup>Only categories that were reported out by at least 10.0% of students are reported in this table. Students could write multiple reasons in response to the open-ended questions, and thus the percentages do not add to 100%.

TABLE 4. Student-perceived advantages and disadvantages of instructor–tutee videos<sup>a</sup>

| Code  | All students<br>(n = 207) | Description  | Example quote   | Example quote  |
|---|---------------------------|--|---|--|
| <b>Advantages of instructor–tutee videos</b>            |                           |  |   |  |
| Student perspective                                     | 30.0%                     | Student indicates that the video includes a student’s perspective or the perspective from someone who is at a student’s level.   | “You had the opportunity to see how another student thinks about the material.”   | “It shows how a student would think of a solution to a question versus how a knowledgeable professor [would].”   |
| Tutee is relatable                                      | 22.2%                     | Student indicates that the tutee had a similar thought process, question, or mistake, or that the student could relate to the tutee.                                       | “Watching the [tutee] think through the problem similarly to how I would was helpful. It also helped me not feel alone with the things I was struggling with because I could see other people also struggle with it.” | “The videos with a [tutee] are very helpful because the [tutee] often will work through the problem in a similar way as I would.”  |
| Instructor tutors tutee                                 | 16.9%                     | Student indicates that the instructor tutors the tutee by correcting misconceptions or mistakes or by providing the scaffolding necessary to arrive at the correct answer. | “It was really interesting to see that [the instructor] guided and helped the [tutee] arrive to the answer. That was really helpful.”   | “At times the [tutee] would answer the question incorrectly, and [the instructor] would work with them through the problem, which was extremely helpful, especially with concepts that I was not fully understanding.” |
| Multiple perspectives                                   | 11.6%                     | Student indicates videos had multiple or different perspectives, but does not specifically mention a student’s perspective.  | “These videos showed two different forms of thinking.”  | “An advantage of the Friday videos with [the instructor] and a [tutee] is it provides multiple different thinking perspectives or processes.”  |
| <b>Disadvantages of instructor–tutee videos</b>         |                           |  |   |  |
| The video is slow or long                               | 32.4%                     | Student indicates the video is slow or long.   | “A disadvantage of the [instructor–tutee videos] is that it takes a longer time to get through the video.”  | “[The instructor–tutee videos] were longer and sometimes seemed to drag on.”   |
| Student is confused                                     | 27.5%                     | Student indicates that, as a result of watching the video, he or she did not have a good understanding of the content or felt confused.                                    | “I was often unable to distinguish the correct information from the misconception after reflecting on the information from the video at a later time.”  | “[The instructor–tutee videos] are sometimes confusing when presenting the material.”  |
| The tutee provides incorrect information                | 26.1%                     | Student indicates that the tutee provides incorrect or inaccurate information.   | “It doesn’t help me when [the tutee] is wrong.”   | “Some of the ideas the [tutees] threw out were just blatantly wrong.”  |
| Presents information in an indirect or disorganized way | 19.3%                     | Student indicates that the delivery of information is disorganized or that information is presented in a roundabout, misleading, or indirect way.                          | “The video spent a long time going down a wrong train of thought and then going back to correct it.”  | “The information was presented in a disorganized manner.”  |

<sup>a</sup>Only categories that were reported out by at least 10.0% of students were reported in this table. Students could write multiple reasons in response to the open-ended questions, and thus the percentages do not add to 100%.

contrasted with the instructor-only videos; 14% of students highlighted that the instructor-only videos improved their understanding of physiology content.

The one disadvantage of the instructor-only videos was the reverse of one of the advantages of the instructor–tutee videos: 1) Students acknowledged the presence of a student’s perspective (30.0%) as an advantage of the instructor–tutee videos, and 18.8% of students reported that the absence of a student’s perspective was a disadvantage of the instructor-only videos. Students who specifically recognized the presence of a student’s perspective as an advantage often contrasted watching how someone with expertise in physiology would solve a problem and the watching a student, with less expertise, think through a physiology problem, with the latter being beneficial. 2) An additional 11.6% of students highlighted that an advantage of the instructor–tutee videos was simply the presence of multiple perspectives. 3) Twenty-two percent of students said that an advantage of the instructor–tutee videos was that they were able to relate to the student in the videos. Specifically, some students explained that seeing a similar student solve physiology problems helped them realize that other students also struggle with physiology. 4) Finally, nearly 17% of students highlighted that watching the instructor tutor the student was an advantage of the instructor–tutee videos. Specifically, students reported that they benefited from watching the instructor guide the student to the correct answer by asking questions to focus the student’s thinking.

### Finding 3: The Majority of Students Prefer Instructor-Only Videos

The extent to which a student values or simply enjoys an active-learning task has been shown to be important when considering the extent to which a student will engage in the active-learning activity (Cooper *et al.*, 2017). Therefore, we were interested in which type of video students preferred. In response to a Likert-scale survey question, we found that the majority of students (59.9%) preferred the instructor-only videos, while 20.3% of students preferred the instructor–tutee videos, and 19.8% of students reported that they did not prefer one type of video over the other. These frequencies were significantly different ( $\chi^2(2, 207) = 65.77, p < 0.001$ ; Figure 4).

Students were asked to explain why they preferred a particular type of video in an open-ended question on the survey. We were interested in whether students focused on the advantages of the video they preferred or the disadvantages of the video they did not prefer when explaining why they preferred one type of video over the other. We used the themes that emerged from the disadvantages and advantages questions to code the preference question using content analysis. We reported out any set of opposite categories (e.g., instructor–tutee videos were too long, instructor-only videos were short) in which one of the two categories was mentioned by at least 10% of students. We also reported out any stand-alone category that was mentioned by at least 10% of students. See Table 5 for reasons why students prefer instructor-only videos and Table 6 for reasons why students prefer instructor–tutee videos. While students who preferred the instructor-only videos seemed to cite both advantages of the instructor-only videos and the disadvantages of the instructor–tutee videos in their reasoning supporting their decision, students who

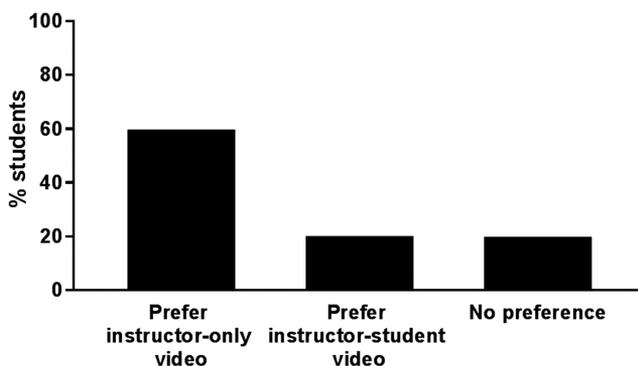


FIGURE 4. Student video preference. Nearly 60% of students preferred the instructor-only videos, 20.3% of students preferred instructor–tutee videos, and 19.8% of students had no preference. These frequencies were significantly different,  $\chi^2(2, 207) = 65.77, p < 0.001$ .

preferred the instructor–tutee videos focused exclusively on the advantages of the instructor–tutee videos.

### Finding 4: Students Perceived Higher Usefulness and Enjoyment with Regard to Watching Instructor-Only Videos Compared with Instructor–Tutee Videos

After 4 weeks of watching a set of either the instructor-only or instructor–tutee videos, students answered five Likert-scale survey questions measuring their perceived usefulness of the videos and five Likert-scale survey questions measuring their perceived enjoyment after watching the videos, each of which was rated from 1 (not at all true) to 7 (very true). The results of a Wilcoxon signed-rank test revealed that students perceived higher usefulness ( $Z = -2.61, p = 0.013$ , Cohen’s  $d = 0.18$ ) and enjoyment level ( $Z = -3.69, p < 0.001$ , Cohen’s  $d = 0.25$ ) with

TABLE 5. Reasons why students prefer instructor-only videos

| Reasons why students prefer instructor-only videos                              | Students ( $n = 124$ ) |
|---|------------------------|
| <b>Student understanding</b>  |                        |
| Instructor-only videos improve observing student understanding of content.      | 3.2%                   |
| Instructor–tutee videos are confusing to observing student.                     | 36.3%                  |
| <b>Presentation of information—directness</b>                                   |                        |
| Instructor-only videos present information in a straightforward or direct way.  | 27.4%                  |
| Instructor–tutee videos present information in an indirect or disorganized way. | 16.1%                  |
| <b>Presentation of information—correctness</b>                                  |                        |
| Instructor-only videos provide only correct information.                        | 11.3%                  |
| Instructor–tutee videos provide incorrect information.                          | 20.2%                  |
| <b>Length and time</b>  |                        |
| Instructor-only videos are short or go by quickly.                              | 5.6%                   |
| Instructor–tutee videos are long or go by slowly.                               | 6.5%                   |

**TABLE 6. Reasons why students prefer instructor–tutee videos**

| Reasons why students prefer instructor–tutee videos         | Students ( <i>n</i> = 42) |
|---|---------------------------|
| <b>Student perspective</b>                                  |                           |
| Instructor–tutee videos include a student’s perspective.    | 42.9%                     |
| Instructor-only videos are missing a student’s perspective. | 0.0%                      |
| <b>Stand-alone categories</b>                               |                           |
| Tutee is relatable in instructor–tutee video.               | 19.0%                     |
| Instructor tutors tutee in instructor–tutee video.          | 35.7%                     |

regard to watching the instructor-only videos compared with the instructor–tutee videos. However, effect sizes are small. Table 7 presents descriptive statistics and Wilcoxon signed-rank test results for the analyses of student-perceived usefulness and enjoyment of the different types of videos.

#### **Finding 5: Students with a GPA below the Median Perform Better after Watching Instructor-Only Videos Compared with Instructor–Tutee Videos**

We measured student performance differences using weekly physiology quizzes. The results of a Wilcoxon signed-rank test revealed no significant difference in students’ quiz scores after they watched instructor-only videos ( $M = 7.95$ ,  $SD = 1.04$ ,  $N = 217$ ) and instructor–tutee videos ( $M = 7.89$ ,  $SD = 1.00$ ,  $N = 217$ ). However, results of further analyses after disaggregating students into a lower-achieving group ( $GPA < 3.49$ ,  $N = 107$ ) and a higher-achieving group ( $GPA \geq 3.49$ ,  $N = 110$ ) using the median GPA of the students in the course revealed that students in the lower-achieving group performed significantly better after watching instructor-only videos ( $M = 7.58$ ,  $SD = 1.03$ ) compared with instructor–tutee

videos ( $M = 7.36$ ,  $SD = 0.99$ ), with a small effect size (Cohen’s  $d = 0.25$ ). No significant difference in student performance after watching the instructor-only or instructor–tutee videos was found within the higher-achieving group of students. Table 8 shows descriptive statistics and Wilcoxon signed-rank test results for the analyses conducted on student performance data.

We conducted post hoc Z-tests to explore whether there were differences between the proportion of higher-achieving and lower-achieving students in the advantages and disadvantages that they reported for each type of video and what reasons they reported for preferring instructor-only or instructor–tutee videos. We conducted a Z-test for each category of each question. We did find a significant difference in why students appreciated instructor-generated videos: lower-achieving students reported that the videos improved their understanding of physiology significantly more frequently than higher-achieving students ( $Z = 3.04$ ,  $p < 0.01$ ). However, we found no significant differences between the two groups of students in the advantages and disadvantages of either video or why they preferred either the instructor-only or instructor–tutee videos.

#### **DISCUSSION**

In this study we used a randomized crossover design to explore the impact of watching instructor-only videos and instructor–tutee videos over the course of 8 weeks in the context of a college physiology course. We found that, even though students could articulate advantages and disadvantages of both types of videos, students were more likely to prefer, value, and enjoy instructor-only videos. In contrast with previous lab-based cognitive science studies, we found that lower-performing students performed better on physiology quizzes after watching instructor-only videos compared with instructor–tutee videos.

**TABLE 7. Descriptive statistics and Wilcoxon signed-rank test results for survey data<sup>a</sup>**

|                      | <i>N</i> | Instructor-only videos |           | Instructor–tutee videos |           | <i>Z</i> | <i>p</i>           | <i>df</i> |
|----------------------|----------|------------------------|-----------|-------------------------|-----------|----------|--------------------|-----------|
|                      |          | <i>M</i>               | <i>SD</i> | <i>M</i>                | <i>SD</i> |          |                    |           |
| Perceived usefulness | 217      | 5.48                   | 1.26      | 5.21                    | 1.26      | –2.61    | 0.013 <sup>b</sup> | 216       |
| Perceived enjoyment  | 217      | 4.62                   | 1.31      | 4.27                    | 1.36      | –3.69    | 0.000 <sup>b</sup> | 216       |

<sup>a</sup>Students found the instructor-only videos significantly more enjoyable and more useful than the instructor–tutee videos.

<sup>b</sup>Holm-Bonferroni adjustment applied on the  $\alpha$  level.

**TABLE 8. Descriptive statistics and Wilcoxon signed-rank test results for student performance<sup>a</sup>**

|                        | <i>N</i> | Instructor-only videos |           | Instructor–tutee videos |           | <i>Z</i> | <i>p</i>           | <i>df</i> |
|------------------------|----------|------------------------|-----------|-------------------------|-----------|----------|--------------------|-----------|
|                        |          | <i>M</i>               | <i>SD</i> | <i>M</i>                | <i>SD</i> |          |                    |           |
| Overall comparison     | 217      | 7.95                   | 1.04      | 7.89                    | 1.00      | –1.26    | 0.214              | 216       |
| Higher-achieving group | 110      | 8.32                   | 0.71      | 8.41                    | 0.65      | –0.97    | 0.348              | 109       |
| Lower-achieving group  | 107      | 7.58                   | 1.03      | 7.34                    | 0.99      | –2.75    | 0.008 <sup>b</sup> | 106       |

<sup>a</sup>There was no significant difference in students’ quiz scores after they watched instructor-only videos and instructor–tutee videos. However, when students were split into lower- and higher-achieving groups, the lower-achieving group performed significantly better after watching instructor-only videos compared with instructor–tutee videos. No significant difference in student performance after watching the instructor-only or instructor–tutee videos was found within the higher-achieving group of students.

<sup>b</sup>Holm-Bonferroni adjustment applied on the  $\alpha$  level.

### Differences between Lab-Based Cognitive Science Studies and Classroom-Based, Discipline-Based Education Research Studies

There has been a recent push to incorporate lab-based cognitive science findings into college science classrooms and recommendations for discipline-based education research to become more theoretically grounded (McDaniel *et al.*, 2017; Mestre *et al.*, 2018). Specifically, there have been suggestions that, when designing multimedia instruction such as instructional videos, developers need to consider both the science of learning, or the research focused on how people learn, and the science of instruction, or the research focused on how to present material in ways that maximize student learning (Mayer, 2008). This study is aligned with these recommendations; we explored whether the results of lab-based cognitive science experiments would replicate when embedded in the context of an undergraduate physiology course. What we found is that they did not. While the previous lab-based studies showed that students performed better after watching a video of an instructor tutoring another student compared with watching a video of only an instructor, we did not find that pattern in this study. When we examined the impact on all students, we found no performance differences between the instructor-only and the instructor-tutee videos. Only when we disaggregated students into higher-achieving and lower-achieving groups did we see a difference, and we found that lower-achieving students performed significantly better after watching the instructor-only videos compared with the instructor-tutee videos. Why did we see an opposite effect for these students compared with what was previously observed in lab-based studies?

One interpretation for the conflicting results is that problem solving in physiology may be fundamentally different than problem solving in physics or chemistry (Chi *et al.*, 2008; Muldner *et al.*, 2014), so the dialogue between the instructor and tutee was not as essential for observing students' learning. Alternatively, in the previous lab-based studies, the only way students received information about a particular topic was through the videos, but in the context of this course, students also were exposed to content and problem solving in the active-learning classes and other non-video preclass assignments. We tried as much as possible to constrain the information presented in the videos to new information, but it was often impossible to completely disentangle the content from other components of the course. Thus, student scores on quizzes, although primarily based on the videos, may have been influenced by other aspects of the course, which may have diluted the potential impact of the instructor-tutee videos.

Another interpretation is that the cognitive science lab-based studies did not account for some of the factors that influence students in a real college course. While lab-based studies have the advantage of being reductionist and often being able to control for many factors, they often lack the complexities of a real classroom setting. One assumption of this study is that students watched and fully engaged with the videos. In the context of a lab-based study, students have no other distractions that may prevent them from fully engaging with the video. However, in the context of a real course, students have competing demands on their time, which may influence their engagement with the videos.

Students highlighted that a benefit of the instructor-only videos was the short length (averaging 17 minutes) and a disadvantage of the instructor-tutee videos was the long length (averaging 21 minutes). It could be possible that something as simple as the difference in video length—as opposed to whether the instructor was tutoring a student—could result in a difference in engagement and may partially explain why lower-performing students performed better after watching the shorter instructor-only videos compared with the longer instructor-tutee videos (Guo *et al.*, 2014; Brame, 2016). Students who watch the videos in a controlled lab space likely experience fewer distractions or competing demands on the time that they have already committed to participating in the study. Thus, the length of the video may only be a concern in the context of a real course, when students experience distractions and have choice over what to do with their time, as they do when completing an outside-class preclass assignment.

We chose to include the same amount of content in both types of videos, which meant that the instructor-tutee videos had to be longer because of the time needed for the instructor to interact with students and correct their misconceptions. While we could have made the length of the videos equivalent, we would have been unable to cover the same amount of content in both videos; we would have covered less content in the instructor-tutee videos due to the additional dialogue between the student and the instructor. The decision to cover the same amount of content in both videos, and thus make the videos different lengths, was made because these videos covered course content on which students would be tested, and we did not think it fair that some students would be exposed to content that other students would not. However, this is a limitation of the study design. It would be interesting to compare videos of similar length in the future to try to establish whether the length of the video was an important factor.

On the basis of these conflicting studies and alternative explanations, we argue that we need more lab-based studies and classroom-based studies. There are advantages and disadvantages to each approach, and a greater degree of bidirectional communication between classroom-based research and lab-based research will help us make decisions about what the most appropriate recommendations may be for instructors.

### Lower-Achieving Students Scored Higher on Quizzes after Watching the Instructor-Only Videos

Although we found no differences between the effect of video type when we examined the scores of all students, when we disaggregated higher-achieving and lower-achieving students, we found that lower-achieving students scored higher on quizzes after watching the instructor-only videos compared with the instructor-tutee videos. Previous lab-based studies have not explored the impact of videos on students of different academic achievement levels, and the small numbers of students in these lab-based studies limits the ability to do this. However, the large number of students in this current study allowed us to disaggregate students into two groups.

Why might lower-achieving students perform better after watching instructor-only videos? In Mayer's theory of multimedia learning, he outlines three learning principles that are most relevant to student multimedia learning: 1) humans possess separate channels for processing visual and verbal material,

2) each channel can process only a small amount of material at one time, and 3) deep learning depends on the learner's cognitive processing (e.g., selecting, organizing, and integrating) during learning. Our previous rationale for including tutees in the video to maximize student learning primarily aligned with the third principle; we based this study on previous studies that suggested that including a tutee in instructional videos may enhance students' abilities to cognitively process the information presented in the video. However, our results suggest that Mayer's second principle for multimedia learning may warrant additional consideration: our working memories have a very limited capacity, and students therefore have a limited capacity to process information (Miller, 1956; Cowan, 2001; Mayer, 2008). Mayer suggests that a key to successful multimedia instruction is to reduce students' extraneous processing or the cognitive processing that does not directly support the instructional goal. In this study, students highlighted that a disadvantage of the instructor-tutee videos was that the tutees sometimes provided incorrect information that the instructor would correct later in the video, which caused some observing students to feel confused. In previous lab studies showing that students learned better after watching instructor-tutee videos, Chi and colleagues hypothesized that a tutee's expression of an incorrect statement followed by a tutor's correction or feedback may enhance the observing students' learning, because when students experience such conflict between the tutor's and tutee's thought processes, they tend to engage in a knowledge-building type of constructive activity when trying to explain the conflict (Chan *et al.*, 1997; Chi *et al.*, 2017). However, some students in this study reported that, when both an incorrect and a correct idea were provided during the video, they felt confused, because they struggled to remember which idea was correct. This may indicate that the presentation of incorrect information, in addition to correct information about physiology, required extraneous processing that did not directly support students' learning and could be the underlying reason for why the lower-achieving students performed worse after watching the instructor-tutee videos. We hypothesize that either lower-achieving students have less working memory capacity, or perhaps higher-achieving students spent extra time reviewing the videos to clarify what information was correct, even if their working memory capacity was exceeded when watching the video for the first time.

### Considering Student Preference in Making Instructional Decisions

This is the first study, to our knowledge, that explored student preferences and attitudes toward instructor-generated videos in biology. Students preferred instructor-only videos to instructor-tutee videos. Students also enjoyed instructor-only videos more and found them more useful than instructor-tutee videos. Does it matter what students prefer or enjoy? We would argue that it may. The results of this study indicate that students did not learn more from instructor-tutee videos. So, if instructor-only videos do not harm student learning, then we would argue that listening to student preference could increase their engagement with a video and may, in turn, increase their learning. Using the lens of expectancy value theory to explore student motivation in the classroom (Wigfield and Eccles, 2000; Eccles and Wigfield, 2002; Cooper *et al.*, 2017), we would pre-

dict that, if students perceive a greater value in instructor-only videos, perceive that instructor-only videos will enhance their learning (students reported this as an advantage of the videos), and see relatively little cost (students report the short time of the instructor-only videos as an advantage), then they will be more motivated to fully engage with the videos.

### Recommendations and Limitations

This study was conducted in the context of one physiology course at one institution, and a single instructor was present in every video. As we highlight in this article, context matters, and this limits the generalizations of our findings beyond this particular context. Further, our finding that lower-achieving students performed better after watching instructor-only videos compared with instructor-tutee videos contradicts previous lab studies, and the significant difference in lower-achieving students' performance on the physiology quiz is small. Also, there were implementation differences between this study and the previous lab studies; one of the biggest differences is that all the information was presented via video in the previous lab studies, whereas the information presented in the videos in this study built upon information that was presented in an in-person class. On the basis of these contradictory results, we cannot recommend either instructor-only or instructor-tutee videos for the purpose of enhancing student performance at this time.

Future research could explore the impact of video length on student learning and connect it to student motivation and engagement in watching the videos. Given the literature that shows that students are more engaged with shorter videos (Guo *et al.*, 2014), ruling out the length of the video as the underlying reason for student preference of instructor-only videos would be an important extension of this work.

Our research findings provide instructors with advantages and disadvantages of both instructor-only and instructor-tutee videos that they can consider as they develop instructor-generated videos. For example, if an instructor is creating an instructor-generated video for the purpose of teaching students a difficult or confusing concept, he or she may want to consider the more straightforward and clearer instructor-only videos. However, if the purpose of the instructor-generated video is to build students' confidence and help them to understand that other students struggle with physiology concepts or to break down common student misconceptions, then it may be more appropriate to create instructor-tutee videos. Future research could explore whether tutees giving wrong answers is helpful for student learning by having tutees use a script in the videos so they do not make mistakes. This may also impact the disadvantages that students report about instructor-tutee videos. We encourage instructors to be thoughtful in their creation and implementation of instructor-generated videos, and we expect this study to contribute to a growing body of literature that will eventually help instructors make evidence-based decisions when using instructor-generated videos in their courses.

### CONCLUSION

In this study, we explored student perceptions of and performance on physiology quizzes after watching instructor-only videos and instructor-tutee videos. We found that students are more likely to prefer, value, and enjoy instructor-only videos. In contrast with previous lab studies, we also found that

lower-performing students performed better on physiology quizzes after watching instructor-only videos compared with instructor-tutee videos. This study illustrates the complexities of translating cognitive science lab-based studies to the real classroom and highlights the need for further research in this area.

## ACKNOWLEDGMENTS

This study was funded by National Science Foundation IUSE award #1504893. We thank the members of Arizona State University's Evidence Based Teaching in STEM journal club for their thoughtful feedback on this work and Hayley Dunlop for her help with the analyses presented in this article. We acknowledge Amy Pate, Guy Mullins, the tutees in the videos, and the instructional design team for helping create the instructor-generated videos, and Miles Orchinik and the instructional team of the course that integrated these videos. We acknowledge the help of Christiana Bruchok, David Yaghmourian, and Joshua Adams during the Fall 2016 pilot year. Finally, we thank the students in the physiology course who took the time to provide feedback so that we can use evidence to make instructional decisions to enhance their experience.

## REFERENCES

- American Association for the Advancement of Science. (2011). *Vision and change in undergraduate biology education: A call to action*. Washington, DC: Retrieved November 28, 2017, from <https://visionandchange.org/files/2013/11/aaas-VISchange-web1113.pdf>
- Bloom, B. S. (1984). The 2 sigma problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher*, 13(6), 4–16.
- Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). *Taxonomy of educational objectives, Handbook I: The cognitive domain* (Vol. 19, p. 56). New York: David McKay.
- Bogdan, R. C., & Biklen, S. K. (1998). Foundations of qualitative research in education. *Qualitative research in education: An introduction to theory and methods* (pp. 1–48). Boston, MA: Allyn and Bacon.
- Brame, C. J. (2013). Flipping the classroom. Retrieved February 2, 2018, from <https://cft.vanderbilt.edu/wp-content/uploads/sites/59/Flipping-the-classroom.pdf>
- Brame, C. J. (2016). Effective educational videos: Principles and guidelines for maximizing student learning from video content. *CBE—Life Sciences Education*, 15(4), es6.
- Carini, R. M., Kuh, G. D., & Klein, S. P. (2006). Student engagement and student learning: Testing the linkages. *Research in Higher Education*, 47(1), 1–32.
- Chan, C., Burtis, J., & Bereiter, C. (1997). Knowledge building as a mediator of conflict in conceptual change. *Cognition and Instruction*, 15(1), 1–40.
- Chi, M. T. H., Kang, S., & Yaghmourian, D. L. (2017). Why students learn more from dialogue-than monologue-videos: Analyses of peer interactions. *Journal of the Learning Sciences*, 26(1), 10–50.
- Chi, M. T. H., Roy, M., & Hausmann, R. G. (2008). Observing tutorial dialogues collaboratively: Insights about human tutoring effectiveness from vicarious learning. *Cognitive Science*, 32(2), 301–341.
- Chi, M. T. H., Siler, S. A., & Jeong, H. (2004). Can tutors monitor students' understanding accurately? *Cognition and Instruction*, 22(3), 363–387.
- Cooper, K. M., Ashley, M., & Brownell, S. E. (2017). Using expectancy value theory as a framework to reduce student resistance to active learning: A proof of concept. *Journal of Microbiology & Biology Education*, 18(2), doi: 10.1128/jmbe.v18i2.1289
- Cowan, N. (2001). Metatheory of storage capacity limits. *Behavioral and Brain Sciences*, 24(1), 154–176.
- Craig, S. D., Gholson, B., Ventura, M., & Graesser, A. C. (2000). Overhearing dialogues and monologues in virtual tutoring sessions: Effects on questioning and vicarious learning. *International Journal of Artificial Intelligence in Education*, 11(3), 242–253.
- Crowe, A., Dirks, C., & Wenderoth, M. P. (2008). Biology in Bloom: Implementing Bloom's taxonomy to enhance student learning in biology. *CBE—Life Sciences Education*, 7(4), 368–381.
- Driscoll, D. M., Craig, S. D., Gholson, B., Ventura, M., Hu, X., & Graesser, A. C. (2003). Vicarious learning: Effects of overhearing dialog and monologue-like discourse in a virtual tutoring session. *Journal of Educational Computing Research*, 29(4), 431–450.
- Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, 53(1), 109–132.
- Eddy, S. L., & Hogan, K. A. (2014). Getting under the hood: How and for whom does increasing course structure work? *CBE—Life Sciences Education*, 13(3), 453–468.
- Enders, C. K. (2003). Using the expectation maximization algorithm to estimate coefficient alpha for scales with item-level missing data. *Psychological Methods*, 8(3), 322.
- Fox Tree, J. E. (1999). Listening in on monologues and dialogues. *Discourse Processes*, 27(1), 35–53.
- 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 USA*, 111(23), 8410–8415.
- Freeman, S., Haak, D., & Wenderoth, M. P. (2011). Increased course structure improves performance in introductory biology. *CBE—Life Sciences Education*, 10(2), 175–186.
- Freeman, S., O'Connor, E., Parks, J. W., Cunningham, M., Hurley, D., Haak, D., ... Wenderoth, M. P. (2007). Prescribed active learning increases performance in introductory biology. *CBE—Life Sciences Education*, 6(2), 132–139.
- Glesne, C., & Peshkin, A. (1992). *Becoming qualitative research: An introduction*. White Plains, NY: Longman.
- Gross, D., Pietri, E. S., Anderson, G., Moyano-Camihort, K., & Graham, M. J. (2015). Increased preclass preparation underlies student outcome improvement in the flipped classroom. *CBE—Life Sciences Education*, 14(4), ar36.
- Guo, P. J., Kim, J., & Rubin, R. (2014). How video production affects student engagement: An empirical study of mooc videos. *Proceedings of the first ACM conference on Learning@ scale* (pp. 41–50).
- Jensen, J. L., Kummer, T. A., & Godoy, P. D. D. M. (2015). Improvements from a flipped classroom may simply be the fruits of active learning. *CBE—Life Sciences Education*, 14(1), ar5.
- Little, R. J. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, 83(404), 1198–1202.
- Mayer, R. E. (2008). Applying the science of learning: Evidence-based principles for the design of multimedia instruction. *American Psychologist*, 63(8), 760.
- McAuley, E., Duncan, T., & Tammen, V. V. (1989). Psychometric properties of the Intrinsic Motivation Inventory in a competitive sport setting: A confirmatory factor analysis. *Research Quarterly for Exercise and Sport*, 60(1), 48–58.
- McDaniel, M. A., Mestre, J. P., Frey, R. F., Gouravajhala, R., Hilborn, R. C., Miyatsu, T., ... Yuan, H. (2017). Maximizing undergraduate STEM learning: Promoting research at the intersection of cognitive psychology and discipline-based education research. <https://circle.wustl.edu/white-paper-maximizing-undergraduate-stem-learning/>
- Mestre, J. P., Chevillat, A., & Herman, G. L. (2018). Promoting DBER-cognitive psychology collaborations in STEM education. *Journal of Engineering Education*, doi: 10.1002/jee.20188
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63(2), 81.
- Monaghan, P., & Stenning, K. (1998). Effects of representational modality and thinking style on learning to solve reasoning problems. In Gernsbacher, M. A., & Derry, S. J. (Eds.), *Proceedings of the 20th Annual Meeting of the Cognitive Science Society of America*, (pp. 716–721). Mahwah, NJ: Lawrence Erlbaum Associates.
- Muldner, K., Lam, R., & Chi, M. T. H. (2014). Comparing learning from observing and from human tutoring. *Journal of Educational Psychology*, 106(1), 69.

- O'Flaherty, J., & Phillips, C. (2015). The use of flipped classrooms in higher education: A scoping review. *The Internet and Higher Education, 25*, 85–95.
- Rackaway, C. (2012). Video killed the textbook star? Use of multimedia supplements to enhance student learning. *Journal of Political Science Education, 8*(2), 189–200.
- Rose, K. K. (2009). Student perceptions of the use of instructor-made videos in online and face-to-face classes. *Journal of Online Learning and Teaching, 5*(3), 487.
- Rubin, D. B. (1987). *Multiple imputation for non-response in surveys*. New York: Wiley.
- Ryan, R. M. (1982). Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. *Journal of Personality and Social Psychology, 43*(3), 450.
- Ryokai, K., Vaucelle, C., & Cassell, J. (2003). Virtual peers as partners in storytelling and literacy learning. *Journal of Computer Assisted Learning, 19*(2), 195–208.
- Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods, 7*(2), 147.
- Schunk, D. H., Hanson, A. R., & Cox, P. D. (1987). Peer-model attributes and children's achievement behaviors. *Journal of Educational Psychology, 79*(1), 54.
- Stockwell, B. R., Stockwell, M. S., Cennamo, M., & Jiang, E. (2015). Blended learning improves science education. *Cell, 162*(5), 933–936.
- Strauss, A., & Corbin, J. M. (1990). *Basics of qualitative research: Grounded theory procedures and techniques*. Newbury Park, CA: Sage Publications.
- VanLehn, K. (2011). The relative effectiveness of human tutoring, intelligent tutoring systems, and other tutoring systems. *Educational Psychologist, 46*(4), 197–221.
- Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology, 25*(1), 68–81.
- Wood, W. B., & Tanner, K. D. (2012). The role of the lecturer as tutor: Doing what effective tutors do in a large lecture class. *CBE—Life Sciences Education, 11*(1), 3–9.