Pacing Online Learning: The Impact of Video Segmentation and Active Learning on Conceptual Understanding in STEM

Lauren Dalton, Ph.D.

Katie Jager

Jeremy M. Rose, Ph.D.

Johanna Caitlin Smyth, Ph.D.

Oregon State University

Authors are listed in alphabetical order and contributed equally to this project.

Oregon State University

Abstract

Many online learning environments use pre-recorded video lectures as a primary mode for disseminating learning content. Despite the commonality of this lecture-based video format, it is not clear to what degree video length and the incorporation of active learning elements influence learner success. We investigated the efficacy of segmenting pre-recorded lecture videos and interspersing elements of active learning in an asynchronous online introductory biology course at Oregon State University. In our experiment, biology students were exposed to three "lecture styles": 1) a single long-form lecture video with formative topical questions at the lecture's conclusion, 2) a single long-form lecture video with formative topical questions interspersed throughout the lecture, and 3) a series of shorter-form lecture video segments (i.e., "chunked" versions of the long-form lecture) that had formative topical questions at each video's conclusion. We gauged student performance by assessing exam scores and learning gains on evidence-based Concept Inventory questions. Our findings indicate that all video-lecture styles allowed students to improve their knowledge of biology concepts and none were significantly better than another. These results suggest that students are able to effectively learn regardless of the online video lecture approach.

Introduction

Determining the most effective practices for teaching and learning in face-to face (F2F) and online modalities of undergraduate education is vital for student success. Active learning is one well-known pedagogical approach in F2F settings that can include students being asked to work on problem sets, examine real-world case studies, and/or participate in peer-to-peer discussions of course-relevant topics. These active elements can range from being interspersed among periods where the instructor speaks (i.e., "lectures") to completely replacing instructor-led lecture components. Incorporating such active learning elements has been demonstrated to help students retain concepts in STEM courses (Freeman et al., 014). These learning gains are accomplished even with a relatively small percentage of active-learning time compared to passive (traditional) lecture in each class period (Freeman et al., 2014). Today, active-learning pedagogy is discussed at conferences and implemented in classrooms world-wide.

Active learning in an online asynchronous classroom may take on a different look as we reimagine and leverage the technological tools to create engagement opportunities with faculty and students within the classroom (Riggs and Linder, 2016). One area we, the authors, still consider problematic is the multimedia lecture video. For many online courses, content is disseminated in part using multimedia videos. While the lecture video affords many useful features such as the ability to disseminate information with auditory and visual features as well as the ability to be watched repeatedly (Chen and Wu, 2015; Dinmore, 2019), it is still a "passive" medium for receiving information. In these videos, the instructor provides explanations of concepts often with accompanying visuals. These videos range in sophistication. However, at their base, these videos mimic closely how an instructor might present material in a F2F classroom. Due to the similarity of the video lecture with the F2F counterpart, we explored how to modulate these videos to provide the highest impact to students in an asynchronous online course.

A few studies have examined lecture videos as a tool for student learning. In one case, Varao-Sousa and Kingstone (2015) randomly assigned students in a F2F course to experience lectures in-person or a pre-recorded video of the same lectures. Students who viewed the recordings performed less well on memory tests than students who attended the live lectures, despite a lack of any differences in mind-wandering reported by students. Results of this study suggest that the mode of the material delivery is important for

learning; furthermore, delivering the same content intended for a live audience via video might not always be effective. So, how do we change the delivery of video lectures to help students learn in the online classroom?

There are a variety of suggestions to increase the effectiveness of online asynchronous lecture videos. One suggestion is to ensure a polished finish production (Guo, et al., 2014). Chen and Wu (2015) found that instructor presence in the video enhances engagement in videos, and Dinmore (2019) found a strong student preference for instructor presence over voice-only lecture videos as well. By adding simple interaction to a video lecture (i.e., the ability to skip to different lecture portions and repeat portions of the lecture content), Zhang et al. (2006) found that students achieved greater learning gains and reported higher satisfaction with their perceived learning. Another technique, segmentation, the process of subdividing lecture content into smaller sections rather than one long lecture per topic, has also been shown to enhance student engagement with course materials by allowing for learner-based pacing of content (Fiorella & Mayer, 2018). These studies have informed basic guidelines for delivering quality video content to online learners, but further research is necessary to better describe how and why these elements contribute to learning gains.

While segmenting appears to be beneficial, the appropriate video-segment length has not been fully explored. Some research suggests 10-15 minutes is the optimal length, although this is based on one article published in 1978 that assessed waning student attention after the start of an in-person lecture (as cited in Bradbury, 2016) and may not be applicable for an online environment. Another study involving massive open online courses (MOOCs), suggests that the median watch time for a video was less than six minutes (Guo et al., 2014). This study indicated that students are unlikely to watch and, by extension, learn from videos that are longer than

six minutes in duration. However, these researchers did not measure retention of course concepts in this study. In addition, given these were large open online courses, the student demography may not align with those enrolled in online courses offered by universities. At institutions like Oregon State University (OSU), students take courses for credit and could therefore have additional extrinsic motivation to learn the concepts. There may also be some significant drawbacks to segmenting videos into small time periods, such as reduced continuity of a complex topic. Additionally, the potential for an overwhelming number of short videos needed to cover a concept may be perceived as daunting to a learner. Thus, a "sweet spot" is needed, where videos are: 1) short enough to allow students to easily dedicate time to each lesson segment; 2) actively engage students in the content for the full lesson duration; and 3) preserve the continuity of the lesson topic.

One possible solution is to add engagement points mid-video. This would increase the length of the video, and preserve continuity without losing learners' attention. This idea mimics the incorporation of active learning problems mid-lecture in a F2F classroom, which has been shown to boost students' retention of ideas (Freeman et al., 2014). In this study, we tested this idea, examining the efficacy of segmenting video lectures, with and without interspersed interactive content, in a large asynchronous online biology course at Oregon State University. Our research questions were:

- Do segmented lecture videos increase conceptual learning compared to single continuous lecture videos?
- How is student learning impacted by interspersing active learning elements among either a single continuous video or segmented videos?

Methodology

The BI 204: Introductory Biology I class at OSU Ecampus participated in this study. As an 11-week asynchronous online course, Ecampus students do not attend lectures, but rather watch lecture videos throughout the week among other class activities. The high enrollment of this course allowed for a sample size conducive to meaningful quantitative analyses. This course is taught solely through Ecampus and is only offered to students enrolled in exclusively Ecampus degrees at OSU. This is a population that is already acquainted with using video-based learning materials, and are more likely to benefit from any positive changes to video delivery formats that may come from this study. We used the Ecampus Video Driven Learning (VDL) system to host the lectures. This is a software platform that supports incorporating active learning elements and tracks student engagement in the videos. Student learning was evaluated through the use of research-validated concept inventories (Baum et al., 2005; *Q4B Concept Inventories | Questions For Biology*, n.d.; Smith et al., 2008) and exam grades.

Study Participants

The study took place in the BI 204 course during both fall 2020 (4 sections) and winter 2021 (3 sections) terms. Student enrollment was capped at 80 students for each section. All enrolled students were informed of the project. Research participants completed a consent form indicating their willingness to be in the study. In total, 109 students consented to participate. To reduce the chance of confounding between the treatment outcome and individual differences in student performance, all students experienced each treatment during the term. To promote participation, students were compensated with a $10 Amazon gift card. All instructors were unaware of which students were participating in the study. Principal Investigators uninvolved in the delivery of course content removed identifiable markers and removed non-participants from the data prior to analysis by the full research team. All research protocols were approved by the OSU IRB prior to the start of the term.

Treatment Groups

All BI 204 students were assigned to one of three treatment groups based on the student's enrolled section. Enrollment in the course sections was randomized by a timed release of seats to ensure even filling of sections. Every student in the course experienced all treatments throughout the term; however, only data from consenting study participants were analyzed (see Table 1). Lecture length and question placement within each lecture video varied by treatment. Each treatment group was tested on the same set of questions. The three treatments were:

1. **Continuous:** Single full-content lecture videos (each 15 to 30 minutes long) with the interactive content questions at the end of the lecture
2. **Interspersed:** Single full-content lecture videos with pauses for interspersed interactive questions occurring approximately every 7–9 minutes during the video duration
3. **Segmented:** Same lecture content as the continuous group, but the lecture videos were divided into discrete 7–9-minute segments with interactive questions at the end of each video segment

Lectures were delivered through Video Driven Learning (VDL) software. The course material for BI 204 was divided into three content blocks, each framed by an exam. Block 1 covered material from the start of the quarter to the first midterm, Block 2 included material between the first and second midterms, and Block 3 covered material between the second midterm and the final exam. Each section of BI 204 received a different treatment in each of the three content blocks. Each content block had all three treatment groups to account for differences in topic difficulty throughout the course (see Table 1 on page 5).

Table 1. Video Treatment by BI 204 Content Block

Content Block	Section 1 (n=47)	Section 2 (n=40)	Section 3 (n=22)
1	Continuous	Segmented	Interspersed
2	Segmented	Interspersed	Continuous
3	Interspersed	Continuous	Segmented

Quantitative Measurements of Student Learning

Learning gains were operationalized in two ways:

1. **Concept Inventory performance:** The first involved collecting responses to research-validated Concept Inventory (CI) questions (D'Avanzo, 2008). These questions were developed via a research-validated process to ensure that each question tests for the intended biological concepts (Baum et al., 2005; *Q4B Concept Inventories | Questions For Biology*, n.d.; Smith et al., 2008). The CI questions used in this study were chosen specifically for their alignment with the corresponding BI 204 content blocks. Before gaining access to any BI 204 learning materials, students first completed a pre-test assignment via Canvas to determine a baseline for knowledge of course concepts. This pre-test assignment encapsulated 45 CI questions that spanned the full breadth of BI 204 topics. As the term progressed, students then completed separate "quiz" assignments at the end of each content block (i.e., major-exam checkpoint), with each quiz containing a subset of the original 45 CI questions. The question subsets corresponded to the topics covered in that content block and thus were designed to serve as measures of knowledge gains from the course learning materials. For each CI-based assignment, we recorded student performance on the CI questions and awarded full assignment credit toward their overall course grade based solely on their participation in the quizzes. This participation-grading approach gave students the opportunity to test their knowledge in a risk-reduced format and mitigated students using external resources to answer the questions. All students enrolled in BI 204, regardless of their participation in the study, were assigned these CI-based quizzes, in case the activity provided learning gains.

2. **Exam performance:** We used exam scores as a proxy for course learning outcome achievement. We compared the average scores in each of the three treatments to determine if there was a significant effect of treatment on student learning.

Results

Our participants (N = 109) matched our target population. More than 98% (n = 107) of the participants had taken at least one other Ecampus class, and the majority (55%; n = 60) had taken 6 or more Ecampus classes (See Figure 1 on page 6). Thus, we assumed that most of the students understood how to navigate the online learning environment, and our results may be generalizable to other OSU Ecampus courses.
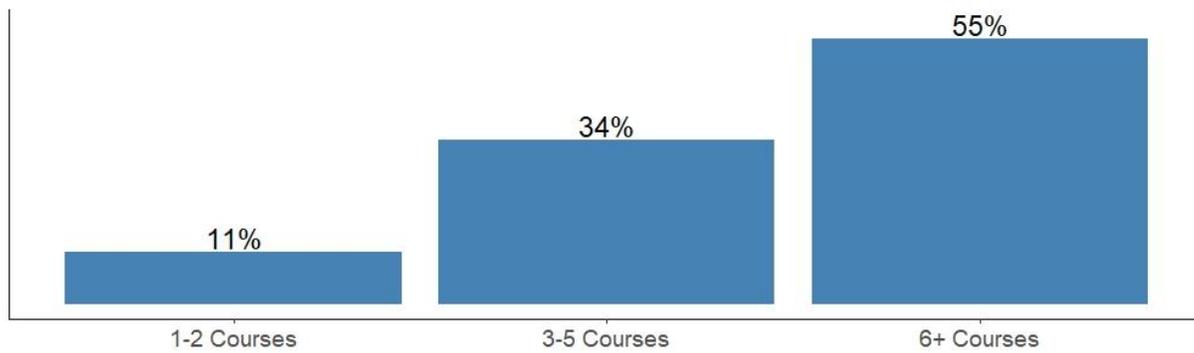
Figure 1. Percent of Participants' Experience Taking Ecampus Courses

Students learned with each lecture style
Overall, we expected that students in all sections, regardless of lecture style, should learn course concepts and by extension improve their concept inventory scores after engaging with the lesson material. To test student learning, concept inventories were given prior to learning class material and then again afterwards. As expected, the mean percent score increased with all the treatment groups (see Table 2). We found that students learned course concepts, as indicated by improved concept inventory scores, across all treatment groups, and the learning gains were not significantly different amongst the three "lecture styles." This finding demonstrates that students were learning throughout the term, irrespective of video modality. This positive result indicates that none of the lecture styles were *harmful* to student learning.

Table 2. Average Percentage Point Improvement of Concept Inventory Questions between Treatments

Treatment (n = 109 for each)	Average Improvement (95% Confidence Interval)
Continuous	13.6 ± 0.5 points
Interspersed	11.6 ± 4.3 points
Segmented	13.1 ± 4.9 points

Concept Inventory learning gains were similar amongst all delivery formats
Next, we analyzed whether there was a significant difference in student improvement on concept inventory questions based on the lecture delivery type. In order to measure learning gains, the difference was calculated between the percent of concept inventory questions correctly answered at the end of the unit minus the percent of unit specific concept inventory questions correctly answered at the start of the term. Some possible issues with this measure are that the concept inventory questions were the same exact questions, meaning some students may perform better simply because they had experienced the question before. However, we did not see gains for all students. Another issue was that the number of questions per unit were imbalanced. Content block 2 in particular had only 5 questions, therefore the possibility for improvement was more limited. We found a large overlap in scores among treatment groups. Figure 2 on page 7 shows the difference scores for the three lecture delivery treatments. The mean scores ± the standard deviations for each treatment are: Continuous = 13.62 ± 19.26; Interspersed = 11.56 ± 18.89; Segmented = 13.09 ± 21.02. The plot indicates there is general improvement in the concept inventory in each lecture delivery treatment with at least 75% percent of students improving their score after engaging in the lecture content.
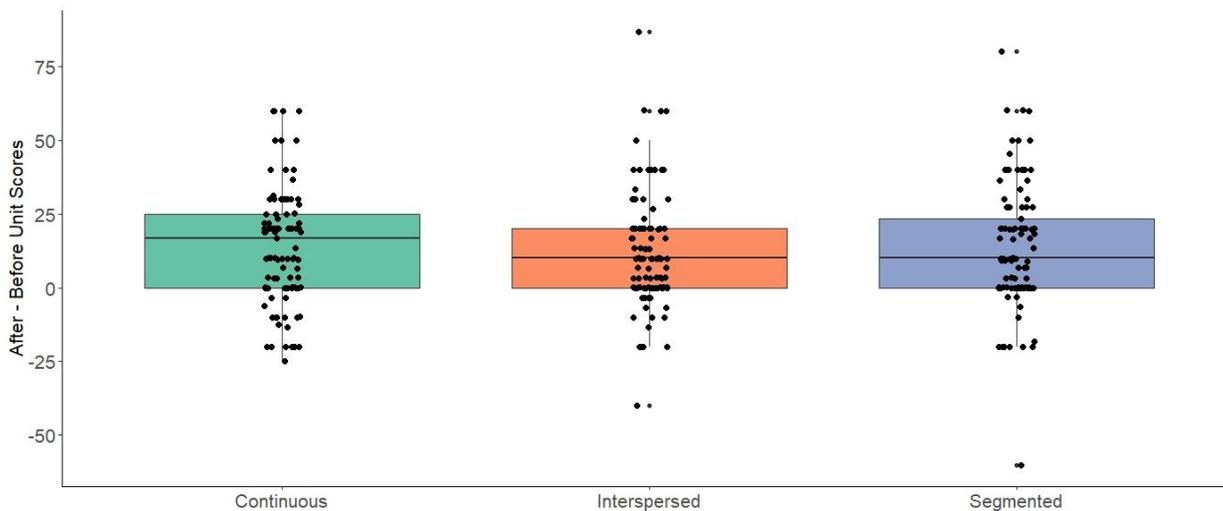
Figure 2. Percent Difference of Concept Inventory Score by Lecture Delivery Method
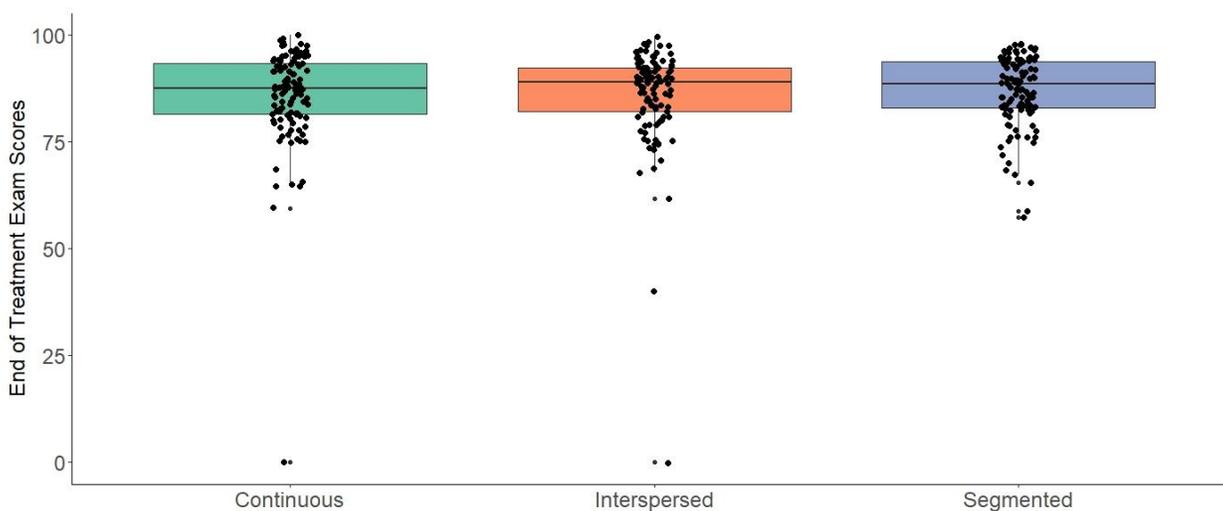

Figure 3. Exam Performance by Lecture Delivery Method

Based on a single-factor repeated measures ANOVA, there were no significant effects of treatment on concept inventory average improvement scores $F(1, 216) = 0.34$, $p = 0.72$. Thus, the concept inventory score differences did not vary based on the type of lecture delivery.

Lecture delivery format did not impact exam performance.
In theory, exams should test course knowledge. In this study, exams can be a secondary measure of a student's learning gains during the course. Therefore, we tested whether the type of delivery treatment significantly influenced exam performance. Because the order of treatments varied depending on the student's enrolled section of BI 204, we first examined if treatment order influenced exam scores.

A two-factor mixed model ANOVA, with video format as the within-subjects factor and the sequence order as the between-subjects factor,

revealed no effect of sequence on exam scores, $F_{(2, 318)} = 1.72$, $p = 0.18$. After removing the treatment-order factor, the single-factor repeated measures ANOVA revealed no significant effect of treatment on average exam scores, $F_{(1.93, 208.2)} = 0.84$, $p = 0.43$.

Figure 3 on page 7 shows exam scores were similarly distributed among the lecture styles. The mean scores $\pm$ the standard deviations for each treatment are: Continuous = 85.74 $\pm$ 11.84; Interspersed = 85.81 $\pm$ 12.07; Segmented = 86.83 $\pm$ 8.40). Each were negatively skewed. The pairwise multiple comparison procedure of the means detected no significant differences among the type of delivery with all Bonferroni adjusted p-values above 0.05. Note that the ability to detect differences based on treatment condition was limited by a ceiling effect in exam scores (topping out at 100%). The non-normality of residuals was a concern; however, log-transformed exam scores did not reveal any additional insights (data not shown).

Discussion

Pre-recorded video lectures are often used to disseminate information to online learners. However, previous research has not addressed how video length or the incorporation of active learning elements influence learner success. We studied the efficacy of adding active learning elements and segmenting lecture videos in an OSU Ecampus introductory biology course.

In our study, we found that none of the video segmenting approaches resulted in a significant change in students' retention of course concepts, as shown by their exam scores and the concept inventory assessments. This is reassuring in the sense that it may indicate that students are able to learn regardless of the degree of lecture-video partitioning. The same content was delivered in each format, suggesting that there is some leeway to the format in which instructors can present the course content.

In planning the project, we had considered that there may be no difference in learning mastery among students, but that students may still express a preferred learning format. To address this, at the end of the study, we asked the students about their preference for the different lecture video formats. Student preference is an important consideration in student learning, especially when the overall learning might not be changed. In the future, we plan to analyze these qualitative responses to identify the underlying themes regarding video preference and how such preferences might influence Ecampus course instruction.

Limitations and Future Directions

There are several limitations to this study. First, we used Concept Inventory (CI) questions as one metric of student learning because those questions have been validated by prior research. This means that there has been a rigorous process to ensure that the questions developed assess learning of those concepts (Baum et al., 2005; *Q4B Concept Inventories | Questions For Biology*, n.d.; Smith et al., 2008). However, access to these questions was limited especially in some content areas. This led to an imbalance in the number of CI questions used in each content block. The middle content block in particular had only a few questions relative to the other two blocks. This made it challenging to assess a change in student percent improvement because there were few questions with which students could improve scores. In the future, we might consider ways to balance the number of questions asked across blocks to more finitely assess student learning gains.

Another limitation to our analysis was a ceiling effect for student exam scores. Exam scores in the course tended towards the upper quartile of the possible scores. This score distribution is expected but caused the average difference in student exam scores across treatments to be limited by the score ceiling. This is why we also included the concept

inventory as a secondary metric of student learning.

In addition, we used multiple-choice questions as the primary mode of formative assessment within the videos. However, as new technologies become available for delivery of asynchronous material, an increasing variety of question types could be asked that may promote deeper learning of the material. For instance, matching, labelling and/or short-answer questions may be more appropriate for certain learning outcomes. Note that any new assessments will need to take into account the extent to which new Artificial Intelligence (AI) tools can be used to respond to questions. Probing students' deeper knowledge of course material may require more creative solutions to assessment, or solutions that utilize AI tools as part of the assessment.

Finally, we used the Ecampus Video Driven Learning (VDL) platform in order to collect quantitative metrics on student watch time, as well as the number of attempts and frequency of correct answers for each student. Unfortunately, there were a few drawbacks to this system that prevented the collection of these data. First, when students rewatched a video, the record of their progress from their first attempt was erased. Second, we were unable to collect the desired granular data on watch time. We had hypothesized that students who preferred shorter videos would pause the longer videos or return to the video or a video segment after stepping away. Such information could help us to more finitely determine how Ecampus students interact with video media in a virtual classroom. The VDL system is currently not capable of performing these fine metrics. Until another system is implemented or this is improved, there is no way for this type of granular detail to be accessed and analyzed at OSU.

As more students move to higher education online, the need to investigate and support students' learning in these spaces continues to be urgent. This study addressed this need, and provides a reassuring message about student learning through video lecture. We show one way that F2F methodologies can be adapted to the virtual asynchronous classroom and that there is flexibility in the format of delivery.

References

Baum, D. A., Smith, S. D., & Donovan, S. S. S. (2005). The Tree-Thinking Challenge. *Science*, *310*(5750), 979–980. https://doi.org/10.1126/science.1117727

Bradbury, N. A. (2016). Attention span during lectures: 8 seconds, 10 minutes, or more? *Advances in Physiology Education*, *40*(4), 509–513. https://doi.org/10.1152/advan.00109.2016

Chen, C.-M., & Wu, C.-H. (2015). Effects of different video lecture types on sustained attention, emotion, cognitive load, and learning performance. *Computers & Education, 80*, 108-121. https://doi.org/10.1016/j.compedu.2014.08.015

D'Avanzo, C. (2008). Biology Concept Inventories: Overview, status, and next steps. *BioScience*, *58*(11), 1079–1085. https://doi.org/10.1641/B581111

Dinmore, S. (2019). Beyond lecture capture: Creating digital video content for online learning – a case study. *Journal of University Teaching & Learning Practice, 16*(1). https://doi.org/10.53761/1.16.1.7

Fiorella, L., & Mayer, R. E. (2018). What works and doesn't work with instructional video. *Computers in Human Behavior, 89*, 465–470. https://doi.org/10.1016/j.chb.2018.07.015

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. https://doi.org/10.1073/pnas.1319030111

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 Conference*, 41–50. https://doi.org/10.1145/2556325.2566239

Questions for Biology. (n.d.). *Q4B Concept Inventories*. University of British Columbia. https://q4b.biology.ubc.ca/concept-inventories/

Riggs, S. A., & Linder, K. E. (2016). Actively engaging student)s in asynchronous online classes. (IDEA Paper #64). IDEA Center. https://www.ideaedu.org/idea_papers/actively-engaging-students-in-asynchronous-online-classes/

Smith, M.K., Wood, W.B., & Knight, J.K. (2008). The Genetics Concept Assessment: A new concept inventory for gauging student understanding of Genetics. *CBE – Life Sciences Education*, *7*(4), 422-430. https://doi.org/10.1187/cbe.08-08-0045

Varao-Sousa, T. L., & Kingstone, A. (2015). Memory for lectures: How lecture format impacts the learning experience. *PLOS ONE*, *10*(11), e0141587. https://doi.org/10.1371/journal.pone.0141587

Zhang, D., Zhou, L., Briggs, R. O., & Nunamaker, J. F. (2006). Instructional video in e-learning: Assessing the impact of interactive video on learning effectiveness. *Information & Management*, *43*(1), 15–27. https://doi.org/10.1016/j.im.2005.01.004

About the Research Unit at Oregon State Ecampus

Vision
The Ecampus Research Unit strives to be leaders in the field of online higher education research through contributing new knowledge to the field, advancing research literacy, building researcher communities and guiding national conversations around actionable research in online teaching and learning.

Mission
The Ecampus Research Unit responds to and forecasts the needs and challenges of the online education field through conducting original research; fostering strategic collaborations; and creating evidence-based resources and tools that contribute to effective online teaching, learning and program administration.

Contact
Naomi R. Aguiar, Ph.D.
Associate Director of Research
Oregon State Ecampus
541-737-9204
naomi.aguiar@oregonstate.edu

Creative Commons License
This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.

Suggested Citation
Dalton, L., Jager, K., Rose, J. M. & Smyth, J. C. (2023). *Pacing online learning: The impact of video segmentation and active learning on conceptual understanding in STEM. White Paper*. Oregon State University Ecampus Research Unit.