



Active learning with online video: The impact of learning context on engagement

Kyoungwon Seo^{a,*}, Samuel Dodson^b, Negar M. Harandi^c, Nathan Roberson^d, Sidney Fels^e, Ido Roll^f

^a Department of Applied Artificial Intelligence, Seoul National University of Science and Technology, Seoul, South Korea

^b School of Information, The University of British Columbia, Vancouver, Canada

^c School of Biomedical Engineering, The University of British Columbia, Vancouver, Canada

^d Centre for Educational Excellence, Simon Fraser University, Burnaby, Canada

^e Department of Electrical and Computer Engineering, The University of British Columbia, Vancouver, Canada

^f Faculty of Education in Science and Technology, Technion — Israel Institute of Technology, Haifa, Israel

ARTICLE INFO

Keywords:

Video-based learning
Learning analytics
Engagement
Learning context
Online/blended courses

ABSTRACT

Learning with online video is pervasive in higher education. Recent research has explored the importance of student engagement when learning with video in online and blended courses. However, little is known about students' goals and intents when engaging with video. Furthermore, there is limited empirical evidence on the impact of learning context on engagement with video, which limits our understanding of how students learn from video. To address this gap, we identify a set of engagement goals for learning with video, and study associated student activity in relation to learning context (course week, exam, and rewatch). In Study 1, we conducted a survey ($n = 116$) that maps students' video viewing activities to their engagement goals and intents. We identified a variety of engagement goals, specifically *Reflect*, *Flag*, *Remember*, *Clarify*, *Skim*, *Search*, *Orient*, and *Take a break*. In Study 2, we analyzed clickstream data generated by 387 students enrolled in three semester-long courses. We examined the impact of learning context on students' engagement with video. A multilevel model showed different patterns for online and blended courses. Students in the online course showed much more strategic and adaptive use of video. As the semester progressed, students in the online courses performed fewer *Reflect* and *Search*. During exam weeks and when rewatching videos, online students performed more *Search* within the video. The only trend that was found for blended learning students was an increase in *Skim* with course week. These findings have implications for video players that adapt to context, such as helping students easily locate important in-video information during the exam week or when rewatching previously watched videos.

1. Introduction

Learning with online video in fully online or blended courses is increasingly popular in higher education. For students, productive engagement with instructional video is instrumental to learning outcomes and academic success (Hockings, Cooke, Yamashita,

* Corresponding author. 232 Gongneung-ro, Gongneung-dong, Nowon-gu, Seoul, 01811, South Korea.

E-mail addresses: kwseo@seoultech.ac.kr (K. Seo), dodsons@mail.ubc.ca (S. Dodson), negarm@ece.ubc.ca (N.M. Harandi), nroberso@sfu.ca (N. Roberson), ssfels@ece.ubc.ca (S. Fels), roll@technion.ac.il (I. Roll).

<https://doi.org/10.1016/j.compedu.2021.104132>

Received 14 December 2019; Received in revised form 7 January 2021; Accepted 9 January 2021

Available online 12 January 2021

0360-1315/© 2021 Elsevier Ltd. All rights reserved.

McGinty, & Bowl, 2008; Michael, 2006; Soffer & Cohen, 2019). Recent studies measured engagement by analyzing students' interaction with video players (Roll & Winne, 2015; van der Sluis, Van der Zee, & Ginn, 2017; Vytasek, Patzak, & Winne, 2020), showing that students use video strategically, and often in nonlinear ways (Sauli, Cattaneo, & van der Meij, 2018). Studies have further demonstrated that passive engagement with video is insufficient for learning (Koedinger, Kim, Jia, McLaughlin, & Bier, 2015). To learn, students need to actively engage with the video content (Lai & Hwang, 2016; van der Meij & Dunkel, 2020; Yousef, Chatti, & Schroeder, 2014). Student engagement can be enhanced by well-designed instructional support (Dumford & Miller, 2018; Halverson & Graham, 2019; Xia & Wilson, 2018), resulting in calls for more research into how to encourage students' deep engagement with video through the design of video content and video players (Guo, Kim, & Rubin, 2014; Mitrovic, Gordon, Piotrkowicz, & Dimitrova, 2019, pp. 320–332; Zhu, Pei, & Shang, 2017).

One approach for promoting more meaningful engagement with videos is through the use of hypervideo players (Dorn, Schroeder, & Stankiewicz, 2015; Fong, Dodson, Zhang, Roll, & Fels, 2018; Liu, Yang, Williams, & Wang, 2019). Linear video provides students with information in a fixed manner. That is, there is a single video track that advances at a fixed speed. In contrast, hypervideo allows students to navigate between linked videos in a nonlinear way (Evi-Colombo, Cattaneo, & Bétrancourt, 2020). That is, hypervideo is essentially a collection of videos that can be viewed dynamically and in various arrangements. One way to enhance the viewing experience of linear videos, without changing their structure, is via hypervideo players. Hypervideo players provide novel affordances, such as advanced navigation controls, AI-based interactions, hyperlinking to supplementary materials, and collaborative communication options (Cattaneo, van der Meij, & Sauli, 2018; Cattaneo, van der Meij, Aprea, Sauli, & Zahn, 2019; Evi-Colombo et al., 2020; Seo et al., 2020a).

Studies have looked into classroom-based use of hypervideo players (Dodson et al., 2018a; Dorn et al., 2015; Liu et al., 2019). However, more research is needed to understand how students' activities in these environments are directed by their goals and intents when engaging with video content (Chi & Wylie, 2014; Kim et al., 2014; Dodson et al., 2018b), as well as the impact of context on students' actions (Fredricks et al., 2011; Henrie, Halverson, & Graham, 2015).

The learning analytics community is exploring clickstream data to measure engagement when learning with video (Roll & Winne, 2015; van der Sluis, Van der Zee, & Ginn, 2017; Vytasek et al., 2020). For example, Kim, Yoon, Jo, and Branch (2018) analyzed students' clickstream data to examine engagement patterns in online courses, and successfully adopted learning analytics to classify different engagement patterns. Jovanovic, Mirriahi, Gašević, Dawson, and Pardo (2019) analyzed clickstream data to predict course performance based on student engagement in blended courses, and revealed the predictive power of student engagement with pre-class activities on course performance. Both of these studies called for more research on whether engagement patterns are affected by learning context.

Most studies of engagement with video have focused on how features of the video content, such as its genre (Chorianopoulos, 2018) and production style (Guo et al., 2014), affect viewing. The impact of video-based learning pedagogy has also been studied, including instructors' design choices (Brame, 2016) and learning scenarios (Andrist, Chepp, Dean, & Miller, 2014). For example, different studies have looked to improve engagement with video using in-video quizzes (Kovacs, 2016), concept maps (Liu, Kim, & Wang, 2018), and *in situ* annotations (Chiu et al., 2018). Less is known, however, about how learning context affects students' engagement with video. Learning is embedded in context, and thus, it is important to understand how context affects students' engagement in various learning activities (Axelson & Flick, 2010; Butler & Cartier, 2005; Deng & Tavares, 2013). For instance, whether or not it is an exam week can affect the ways students interact with people, information, and technology (Fong et al., 2019). More work is needed to understand how learning context associated with video-based learning affects student engagement (Chen & Wu, 2015; Hew, 2016).

Considering the findings in the literature and the areas for further research, the present study aimed to identify students' video engagement goals and how these inform students' activities, in relation to three contextual variables. We focus on online and blended course formats in higher education, using regular linear video through a hypervideo player. In Study 1, we mapped students' video viewing activities to their engagement goals and intents (see Section 3). In Study 2, we analyzed how students' engagement goals and video activities were affected by course format and contextual variables (see Section 4). We address the following research questions:

RQ1: *What are students' engagement goals and intents for their video viewing activities?*

RQ2: *How does learning context (and specifically, course week, exam, and rewatch) affect students' engagement with video in online and blended courses?*

2. Background

With mounting evidence for the benefits of learning with online video (van der Meij & Dunkel, 2020; Vieira, Lopes, & Soares, 2014; Yousef et al., 2014), and thanks to technological developments (Seo et al., 2020a), instructional videos are becoming increasingly popular in higher education. Two popular video-based course formats are online courses and blended courses. *Online courses* provide most or all of the learning content through online video lectures, without face-to-face lectures (Allen & Seaman, 2016). In these courses, videos are used as either a self-study medium or a tool to enhance the learning process (Vieira et al., 2014). Online learning is thought to have several advantages, including flexibility and accessibility to study anywhere, anytime, without requiring instructors and students to be co-located (Means, Toyama, Murphy, Bakia, & Jones, 2009; Roll, Russell, & Gašević, 2018; Van Doorn & Van Doorn, 2014). Thus, in online courses, video is often the primary source of information communicated by the instructor. In contrast with that, in *blended courses*, online instructional videos typically provide supplemental materials, which are used in combination with face-to-face lectures from instructors (Graham, 2006). A blended learning environment adds flexibility to traditional lecture-based courses and allows instructors to play a pivotal role in providing structure, organization, scaffolding, and time management to the learning experience for students (Aldhafeeri, 2015; Artino & Jones, 2012).

Researchers, instructors and policy makers are eager to find how online video-based teaching could support student learning (Ashby, Sadera, & McNary, 2011; Larson & Sung, 2009). Comparative studies between online and blended courses indicate that contextual variables other than the format alone influence learning outcomes (Nortvig, Petersen, & Balle, 2018). In a similar vein, existing literature suggests that students may engage differently in online and blended courses (Broadbent, 2017; Nortvig et al., 2018). Despite this, explicit tests of differences in student engagement across online and blended courses have seldom been undertaken, if at all. Such knowledge would provide greater insights to design instruction and technologies that are more suited for one context or the other.

2.1. Theoretical framework

In online and blended learning courses, student's use of online video lectures calls for new ways to measure engagement. Roll & Winne (2015) proposed using learning analytics techniques to study real-time trace to better assess and understand engagement "as it unfolds." For example, studies measured engagement by looking at the duration of students' viewing patterns (Guo et al., 2014), or whether they navigate away from a video before completion (Kim et al., 2014). Notably, these measures focus on whether students accessed different parts of the video. However, they cannot capture whether a student is actively paying attention to the video or just playing it in the background while multitasking (Guo et al., 2014). An alternative way to understand student engagement with videos is to look at students' actions while watching video. Analysis of students' clickstream (such as pausing or changing playback speed) can help us better understand their engagement with video. Such analysis includes existing metrics (for instance, number of pauses is identical to number of playbacks), and could add information about the type and frequency of different forms of activity.

The Interactive, Constructive, Active, and Passive (ICAP) framework offers a useful lens through which student actions can be interpreted. The ICAP framework describes learning activities as a process that involves four different modes of engagement, from most to least productive (see Table 1; Chi & Wylie, 2014). Mapping learning activities with video to different modes of engagement can help us interpret and evaluate these. Dodson et al. (2018) put forward an Active Viewing Framework that mapped video-specific learning activities in the ICAP framework. Mitrovic, Dimitrova, Lau, Weerasinghe, and Mathews (2017) have also incorporated the ICAP framework to the design and evaluation of educational video playing systems. In their work, they found that when students engage in annotating videos and rating others' comments, their conceptual understanding of the target soft skills increases.

When interpreting students' activity, it is important to consider their goals. For instance, as shown in Table 1, fast-forwarding and rewinding video are both examples of active learning with video. However, the goals of such activities likely differ. While fast-forwarding reduces the time with the content and is used to move on, rewinding is used to repeat earlier content. Engagement is a "meta-construct," so it is important to have a comprehensive understanding of observable activities and goals (Fredricks et al., 2011). Current view of engagement focuses on different aspects of engagement. Its affective component refers to students' motivation and reactions while engaging; its cognitive component refers to investment and mental effort; and its behavioral component refers to the actions taken by students within and outside of the learning environment.

Students may engage in the same activity in order to support their learning at a number of different ICAP levels. This makes it difficult to identify at which level a student is learning. For example, a student may pause a video to take a break (Active), to write a synthesizing note on the content (Constructive), or to collaborate with a peer (Interactive). In Study 1, we attempted to more clearly map activities to ICAP levels by identifying what the most common goals and intents are for a number of common video-based learning activities. Obviously, such mapping cannot always be accurate. However, looking at students' goals and intents help us by grouping activities that may otherwise seem unrelated. For example, seeking forward in the video or playing at a higher rate are different activities, but they are often related to the same goal: skimming the current content. In RQ1, we explore how students' video viewing activities map onto different engagement goals and intents.

2.2. The impact of learning context

Understanding the impact of learning context is essential to interpreting students' activities. Learning context represents all the factors surrounding students within a learning environment that provide meaning for the messages they receive (Tessmer & Richey, 1997). These are the factors that influence and define what, when, where, and how individual students learn from instruction. Understanding learning contextual factors is important to unveil the meaning of students' learning activities (Deng & Tavares, 2013). Since context is a multi-faceted concept (Barnett & Ceci, 2002; Courtright, 2007), there can be many variables to consider.

Recent research has identified some of the contextual variables that can impact students' engagement when learning with video. Broadbent (2017) showed that students utilize different learning goals and strategies in different course formats (*online vs. blended*

Table 1

Examples of video-specific learning activities by each mode of engagement identified in the ICAP framework, adapted from Chi and Wylie (2014).

Mode of engagement	Examples of video-specific learning activities
Interactive	Debating with a peer about the justifications; Discussing similarities & differences
Constructive	Explaining concepts in the video; Comparing and contrasting to prior knowledge or other materials
Active	Manipulating the video by pausing, playing, fast-forward, rewind
Passive	Watching the video without doing anything else

course design). Harandi et al. (2018) showed that students' engagement with video may change from course week to week. In a similar vein, Harandi and her colleagues (2018) reported that students watch the videos extensively before midterm and final exams. Further, whether students are watching a video for the first time or are rewatching it, could affect students' engagement patterns (Sinha, Jermann, Li, & Dillenbourg, 2014). When analyzing students' learning activities, an in-depth insight into how students behave can be lost if contextual variables are not fully considered (Axelson & Flick, 2010; Deng & Tavares, 2013). Alas, research considering the relationship between contextual variables and engagement with video is limited, and Chen and Wu (2015) and Hew (2016) have called for more work on the impact of context.

Further elaborating on the concept of engagement, Lalmas, O'Brien, and Yom-Tov (2014) point out the important role of time and emphasize the importance of looking at engagement across as well as within sessions. Subsequently, they define user engagement as: "the emotional, cognitive, and behavioral experience of a user with a technological resource that exists, at any point in time and over time" (Lalmas, O'Brien, & Yom-Tov, 2014, p. 3). Seeking to better understand the changes to students' engagement over time, in RQ2, we examine the relationship between temporal context (and specifically, course week, exam, and rewatch) and students' engagement with video in different course formats (i.e., online and blended courses).

In what follows, we describe Study 1 (Section 3): a survey that maps in-video activities to engagement goals and intents. We then use these findings in Study 2 (Section 4), to interpret students' clickstream data from three courses. Lastly, we discuss the main findings and implications for theory and practice of video-based teaching and learning (Section 5).

3. Study 1: Mapping video viewing activities to engagement goals and intents

To associate students' video viewing activities with their different engagement goals (RQ1), we surveyed students regarding their intents for different activities when learning with video. These intent–activity pairings were then synthesized into broader engagement goals using affinity diagramming (Hanington & Martin, 2012), as described below. Interactive engagement, as described in the ICAP framework, occurs during social discussions with peers, and thus leaves no traces in systems that do not support such interactions. Thus, in this study, we focused on Constructive, Active and Passive modes of engagement.

As the study asked students to describe their intents for different video viewing activities, it relies on self-reports. While such an approach has its limitations, surveying students for their goals is relatively common, and students can report these reliably (cf. Butler, Schnellert, & Perry, 2017; Kizilcec & Cohen, 2017).

3.1. Methods

3.1.1. Participants and procedure

Ninety-six undergraduate students enrolled in one of three courses participated in the survey: forty-eight from an online, third-year psychology course and 48 from two blended first- and second-year engineering courses. Recruitment advertisements were posted on the course learning management system by one of the researchers who was not affiliated with teaching the courses. All participants completed the survey at the end of taking a video-intensive course that made use of ViDeX (Fong, Dodson, Zhang, Roll, & Fels, 2018), an experimental hypervideo player that supports *in situ* annotating and highlighting. ViDeX was also used in Study 2, and is described in Section 4.1.2.2. Respondents were entered in a raffle for a \$50 gift card.

3.1.2. Materials

The survey asked students regarding their intents for engaging in different video-based learning activities (see Appendix A). The survey focused on key activities that are supported by the video player used in their class: pausing, rewinding, fast-forwarding, highlighting, scrubbing, adjusting the playback speed, and annotating the video. For example, one of the seven questions asked, "When watching an instructional video, I typically pause because ...". The order of the seven questions was randomized to minimize ordering effects.

In order to construct meaningful options for the survey, a pre-survey with open-ended version of the questions was administered to 20 additional volunteers from other courses, mainly in Education. Multiple studies have shown that open-response questions offer more insight into student reasoning than do multiple-choice (Meir et al., 2019; Prevost, Smith, & Knight, 2016). These open-responses were analyzed with inductive coding (Schreier, 2012). The outcome of this process was the emergence of four to five intents for each activity. These intents were used as the options from which participants in the main survey could answer. This pre-survey was used solely to identify plausible intents for each activity.

Recognizing that each activity can serve multiple intents, survey respondents could choose all intents that apply. In addition, respondents were invited to write-in additional intents using an "Other, please specify" option. For example, one of the seven questions asked, "When watching an instructional video, I navigate forwards because ... a) I find the content to be irrelevant, boring, or already known to me; b) I prefer to study the video at my own pace by jumping around, instead of watching linearly; c) I want to get the gist of the content before I watch the entire video, to know what to expect and decide whether to watch it; d) Other." For all items, respondents were also reminded that "If you almost never engage in one of these activities, please skip that question." Indeed, we saw that students skipped activities that were less common. For example, for highlighting and for annotating videos, only 61 and 68 students, respectively, provided their rationale from a total of 96 students.

3.1.3. Data analysis

Data from the survey was organized in intent-activity pairs. These pairs were then grouped by the research team into broader

engagement goals using the affinity diagram approach. Affinity diagram is a technique for organizing and grouping ideas and discovering common themes (Hanington & Martin, 2012). During the process, all intent-activity pairs were written on cards. Participants then grouped them iteratively until clusters of pairs with similar meanings emerged and were named. For example, the intent for navigating forwards because “I find the content to be irrelevant, boring, or already known to me,” and the intent for increasing the playback speed because “I am familiar with the content” were grouped as a same engagement goal labelled *Skim*. The affinity diagram process is informed both by data and by theory. For example, the *Skim* goal matches a parallel strategy that has been documented while reading (Pavel, Reed, Hartmann, & Agrawala, 2014; Shin, Berthouzoz, & Durand, 2015). Affinity diagramming looks for emerging clusters. It ends when all elements are mapped onto clusters in the diagram, and when the research team together agrees that this representation captures well the diversity of responses.

3.2. Results: Students' goals, intents, and activities while learning with video

Table 2 maps engagement goals, intents, and the associated video viewing activities for each multiple-choice question in the study. The frequency of each response is also presented. Each activity is annotated with the appropriate level according to the ICAP framework (Chi & Wylie, 2014). For several activities, the relevant ICAP level can be attributed with a high degree of certainty. For example, fast-forwarding is Active, as also noted by Chi and Wylie (2014). However, without additional data, it is hard to associate between other activities and specific levels. For example, while the act of pausing is Active, students reported that they often pause in order to summarize in their notebooks (a Constructive activity). Similarly, when annotating videos, students can create verbatim notes (Active) or their own summaries (Constructive). Because students can create a word-for-word copy of the spoken content of a video by highlighting the transcript, we assume that writing a comment is a Constructive activity. However, lacking semantic information, we classified these annotations as either Active or Constructive. We identified a total of eight unique engagement goals that students used while learning with video.

Over half of the students in the study chose the following intent-activity pairs: (i) *Pausing* in order to write a note (48% also noted that they pause to think and reflect); (ii) *Rewinding* to make sure that they did not miss anything, or because they did not understand the video first time around; and (iii) *Scrubbing* (that is, hovering over the filmstrip) in order to better understand what is being said in the video.

Several activities can serve more than one goal, and most goals can be accomplished using more than one type of activity. Thus, results show a many-to-many mapping. For example, students may pause to take a break, reflect, or use external tools such as notebooks and a browser. Similarly, a similar goal can be served by various activities. For example, students who want to save time because the content is familiar to them either increase the playback speed or skim forward.

Table 2
Engagement goals, intents, associated viewing activities, and frequencies of responses.

Engagement goal	Intent	Video viewing activity	N
Reflect	I want to write a note.	Pause (A/C)	63
	I need time to think and reflect on what I just watched.		46
	I search for additional information using the web, textbook, and so on.	Annotate videos (A/C)	16
	I summarize the video to save study time.		30
Flag	Later, it is easier to review my annotations than to rewatch the entire video.	Highlight (A)	27
	I bookmark content that I want to return to later.	Highlight (A)	25
	I mark the first time that important concepts appear.		17
Remember	I want to remember something important for an assignment or test.	Annotate videos (A/C)	39
	It helps me remember the content.	Rewind (A)	28
	It helps me remember what is said in the video.	Highlight (A)	25
Clarify	I zoned out or got distracted, and want to make sure I did not miss anything.	Rewind (A)	58
	I did not get the explanation the first time.		56
	The speaker is talking too slow or fast.	Reduce playback speed (A)	36
Skim	Speeding up the playback saves me time.	Increase playback speed (A)	39
	The speaker is talking too slow or fast.		36
	I am familiar with the content.		34
Search	I find the content to be irrelevant, boring, or already known to me.	Fast-forward (A)	32
	To better understand what is being said in the video.	Scrub (A)	50
	I like to read the transcript while watching the video.		42
	To locate a specific piece of information to navigate to.		36
Orient	I am looking for a keyword.		19
	I prefer to study the video at my own pace by jumping around, instead of watching linearly.	Fast-forward (A)	25
	I want to get the gist of the content before I watch the entire video, to know what to expect and decide whether to watch it.		21
Take a break	Something else is grabbing my attention or I need a break.	Pause (A)	36

Note: Each video viewing activity is labelled with its associated ICAP level(s): “A” for Active and “A/C” for Active or Constructive, depending on the situation.

3.3. Discussion

Study 1 produced a mapping between activities, intents, and engagement goals for eight video viewing activities. The relationship between intents and activities was captured directly from students' surveys. The grouping of intents to goals was the outcome of the affinity diagramming process. Fig. 1 summarizes these results. When the same engagement goal was related to the same activity, we added up the different intents (e.g., students reported three different intents when pausing to *Reflect*, so we summed the counts for each activity associated with a particular goal: $63 + 46 + 16 = 125$). Fig. 1 highlights the most common goal for each activity.

Study 1 identifies a set of goals that students hold while engaging with online video, and captures the way in which students achieve these different goals. Notably, the goals identified in this work resemble previous results. For example, when describing the ICAP framework, Chi et al. (2018) describe the Active level as one in which students manipulate their environment without providing any new information. The authors provide examples such as pausing, rewinding, or repeating videos. The authors further describe how these manipulations serve the purpose of driving attention to (or away from) certain parts of the learning material, as is the case with goals such as *Skim* and *Search* which were identified here. Chi et al. (2018) go on to say that major goals for the manipulation of learning objects is to store them in ones' memory, and link and activate with prior knowledge. Goals such as *Flag* and *Remember*, from this study, match these descriptions.

The *Reflect* goal, and its manifestation in pausing and annotation activities, also echoes similar findings from the literature (Evi-Colombo et al., 2020). In their review, the authors describe how the act of summarizing video content, and revisiting these notes later, serves to drive elaboration, justification, and reflection. Students in our study reported similar intents.

The goals described here also resemble many reading strategies (cf. McNamara, 2007). This is not surprising, as both activities are designed to learn from provided information, at the control of the student. Thus, goals such as *Skim*, *Clarify*, or *Search* describe strategic use of actions. These, in their turn, serve bigger goals of learning.

This study also demonstrates the challenge of using learning analytics techniques for capturing these goals. In addition to the lack of video-viewing trace data for Interactive activities (such as conversations), clickstream data is also lacking for pausing, which is reported by students to be a key viewing activity. Furthermore, pausing can serve opposite engagement goals — whether *Take a break* or

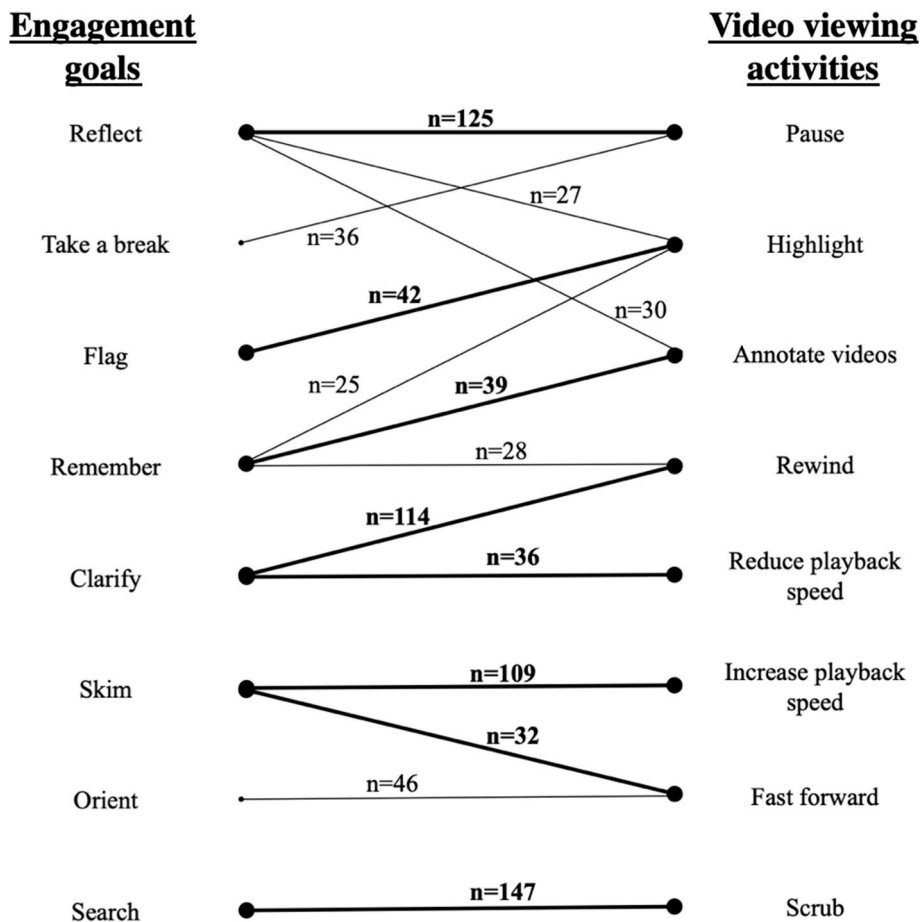


Fig. 1. Engagement goals and video viewing activities. Bold lines indicate the primary goal for each activity. For example, pause is primarily associated with *Reflect* ($n = 125$) but also *Take a break* ($n = 36$).

Reflect. Some of this ambiguity may be resolved with other methods, such as sequence mining, which evaluates these pauses as part of longer sequences, or using other machine learning techniques (Perez et al., 2017; Sinha et al., 2014).

The narrow bandwidth of students' clickstream also makes it hard to attribute ICAP levels to some activities, such as annotating. In general, with regard to ICAP levels, two observations can be made. The first is that students can use a variety of activities across a variety of levels to achieve the same engagement goal. The second is that clickstream data can be used predominantly to capture Active level activities.

In the next section, we apply the mappings between activities and goals when looking at student trace data to evaluate how students' goals and activities are affected by learning context in the lifespan of a course.

4. Study 2: Analyzing the impact of learning context on students' engagement goals with video

The goal of Study 2 was to examine the impact of learning context on students' goals and activities during their video learning (RQ2). We analyzed clickstream data from 387 undergraduate students in three semester-long courses. Students' interactions with online video were recorded and analyzed based on the goals identified in Study 1. We then evaluated the relationship between various contextual variables and students' activity using multilevel modeling.

4.1. Methods

4.1.1. Participants and procedure

Data were collected from 387 students enrolled in three different courses as they watched instructional video with an experimental hypervideo player: Psych ($n = 117$; a third-year psychology course), Eng1 ($n = 137$; a second-year engineering course), and Eng2 ($n = 133$; a first-year engineering course). Psych was a fully online learning course, while Eng1 and Eng2 were blended courses. Students watched online videos as part of their weekly course work throughout the semester. At the beginning of the semester, students consented to having their trace data captured and analyzed anonymously. Students were thanked for their time but not otherwise compensated for their participation.

4.1.2. Materials

4.1.2.1. Online and blended courses. Table 3 and Fig. 2 provide detailed information about the three courses in Study 2. In the analysis, we merged data from the two blended courses (Eng1 and Eng2). This choice was motivated by both data-driven and conceptual reasons. From a conceptual perspective, the two engineering courses are similar in content and population (Engineering courses at the same university), modality (blended courses, where online video supplements face-to-face lectures), and in their type of videos (PowerPoint slides with voice-over). Thus, merging them is in line with our research questions. From a statistical point of view, merging the courses has several benefits. First, it increases the statistical power of the multilevel approach, since it increases the number of observations per cell. Second, by running fewer models, we reduce the risk of Type II errors. To evaluate the decision to merge data from both Eng courses, we ran a multilevel model with a categorical variable of "Course ID" (coded "1" if the student was in the Eng1 course and "0" if the student was in the Eng2 course). If the two courses exhibit different viewing patterns, we would have observed significant loading on this variable. There was no significant effect for the course variable ($p > .05$), indicating that there were no statistical differences between the two blended courses.

4.1.2.2. ViDeX hypervideo player. Students watched instructional videos using the ViDeX, an experimental hypervideo player designed for video-based online and blended learning (Fong, Dodson, Zhang, Roll, & Fels, 2018). ViDeX consists of three interface elements: the player, the filmstrip, and the transcript (see Fig. 3). The player is similar to other conventional video players in that it allows students to play, fast-forward and rewind, and adjust playback speed, closed captioning, and volume. The filmstrip is located at the bottom of the player and provides a visual overview of the video through thumbnails. ViDeX also includes hypervideo capabilities. The transcript combines thumbnails and the spoken content of the video. Students can highlight or annotate videos by clicking and dragging content in the transcript or on the filmstrip and then selecting an annotation type from the popup window.

4.1.3. Data collection and processing

Students' clickstream data (e.g., pauses, rewinds, fast-forwards, scrubs, and playback speed adjustments) were collected over a semester (105 days). Our analysis focuses on students' activities within viewing sessions. We defined a session as a single student-video-date triplet, that is, playing a certain video on a certain day by a certain student. Watching the same video on two different days was considered two sessions, as was watching two different videos by the same student on the same day. We further split sessions

Table 3

Basic information about the courses. For the blended courses, we report data from two engineering courses.

Course format	Course topic	Number of weeks	Number of videos	Video lecture style	Video duration in minutes (std)
Online course	Psychology	14	90	Slides with voice-over	8.56 (4.28)
Blended course	Engineering	8; 5	33; 13	Slides with voice-over	18.85 (13.10); 7.75 (3.33)

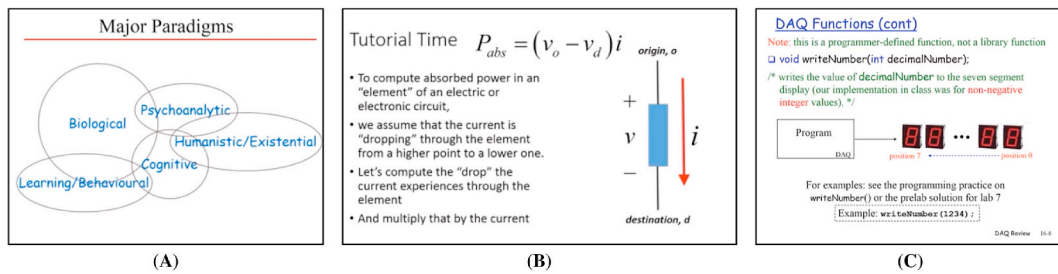


Fig. 2. Video screenshots from three courses: (A) Psych, an online third-year psychology course; (B) Eng1, a blended second-year engineering course; and (C) Eng2, a blended first-year engineering course. The video can be viewed at: <https://www.youtube.com/playlist?list=PLjnfG16LWxcoruPi71JK0Gp90Ev1Y0IQz>.

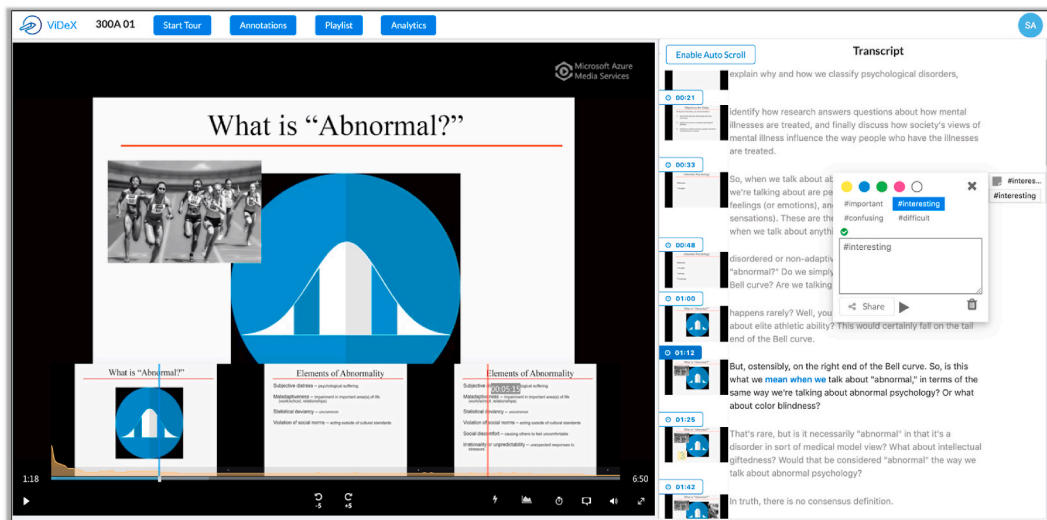


Fig. 3. A screenshot of ViDeX, the hypervideo player. The visual content player is on the upper-left, the navigation filmstrip is on the lower-left, and the transcript is displayed on the right. The user can highlight or annotate videos on the popup window.

that included a period inactivity longer than 10 min (no video playing and no user interaction). A within-session analysis, as used here, can look into students' activities once they began watching a video. Their decision to watch a certain video is outside the scope of this work. For each session, we included the following data: student ID, video ID, date, the number of activities associated with each engagement goal (see Section 4.1.3.1), and learning context variables (see Section 4.1.3.2).

4.1.3.1. Engagement goals. We used the mapping identified in Study 1 (see Table 2 and Fig. 1) to count the number of activities associated with each engagement goal for each session. When the same goal is served by multiple activities, we summed up their frequencies. For example, the frequency of *Clarify* was calculated as the sum of the number of rewinding and reducing playback speed.

When an activity is associated with multiple goals, we identified the most common goal (see bold lines in Fig. 1). For example, we treated rewinding as an instance of *Clarify* (114 votes), rather than *Remember* (28 votes). The goal of *Remember* was left with no associated activities, as all of its associated activities were more common with other goals. The goal of *Orient* was also removed for the same reason. The goal of *Take a break* also had no associated activities, because long periods of inactivity between actions were treated as a separation between two viewing sessions. This process is further described in Section 4.1.3.

Unlike most video players, ViDeX supports highlighting video. Alas, when preparing the data for analysis, it became clear that students rarely used the highlighting and annotating features. For example, more than 88% of the annotations (2,458 of 2,787) were created by nearly 5% of the students (20 out of 387) a phenomenon which has been observed in previous studies of video-based annotation in the classroom (Dodson et al., 2018a). The fact that such a small percentage of the overall students created so many annotations requires further attention. However, for the purpose of evaluating adaptive viewing patterns, the goal of *Flag* was removed. Consequently, the data analysis focused on four engagement goals, intents, and their associated video viewing activities (see Table 4).

4.1.3.2. Learning context variables. To evaluate the impact of learning context, we examined the following variables: Course Week (first to last), Exam (exam week vs. non-exam week, where an exam week represents the one-week period before students take the

exam), and Rewatch (rewatch vs. first-watch). The models also included Video Duration (video length in minutes).

4.1.4. Data analysis

The trace data was summarized with descriptive analysis. Then, we fitted a three-level mixed-model analysis, consisting of session data nested within video, nested within student. Multilevel modeling can be used to predict dependent variables using multiple independent variables coming from different levels (Raudenbush & Bryk, 2002). Multilevel modeling has the advantage of dealing with problems in traditional regression analysis, such as aggregation bias, dependency between points, and under-estimated standard errors (Lee, 2000). Multilevel analysis was conducted using Stata 14.

We ran a total of eight models: the two course formats (online or blended) by each of the four engagement goals (*Reflect*, *Clarify*, *Skim*, and *Search*). We analyzed the course formats separately to avoid three-way interactions. We took an iterative approach to model-building by including each level and associated controls to examine the fit of each model. We began with a “null” model which included only the dependent variable and the nesting variables (Model 1). Then we introduced the main effects in Model 2. Then we tested for interaction effects (Model 3). We selected the model with the best fit using fit parameters and a likelihood-ratio test; comparing nested models to more complex models (Model 1 vs. Model 2, then Model 2 vs. Model 3). If a more complex model with additional terms did not show a statistically significant improvement in terms of model fit using the likelihood-ratio test, we adopted the simpler model for the sake of parsimony and increased statistical power. For ease of interpretation (and brevity), we only report our final models. A primary interest is the estimated main effects for course week, exam, rewatch, and video duration. To account for inflated type I error rates with eight different models, we adopt the Bonferroni correction to the critical p-value ($0.05/8 = 0.00625$).

4.2. Results

4.2.1. Descriptive analysis

A total of 14,262 video sessions (10,025 in the online course and 4,237 in the blended courses) were recorded over a period of 105 days. As shown in Table 5, students in the blended courses (Eng1 and Eng2) showed a higher average number of activities per session for all engagement goals, except *Search*, compared to the online course.

4.2.2. Online course multilevel modeling

We first analyzed the online course data. The three-level model, including the interaction between exam and rewatch, explains more total variance and has a higher improved fit than the other models (see Table 6). Course week is negatively associated with *Search* instances ($\beta = -0.280$), that is, for every four weeks students did roughly one less search per session. Course week also predicts a decrease in activities associated with *Reflect* ($\beta = -0.086$). Exam week predicts less *Reflect* ($\beta = -1.456$) and more *Search* ($\beta = 2.005$). Rewatch predicts less *Reflect* ($\beta = -2.189$) and more *Search* ($\beta = 4.906$). Video duration (measured in minutes) shows a statistically significant positive relationship with all four engagement goals, that is, longer videos have more activities registered, though this increase is not linear. A two-way interaction between exam and rewatch was significant, meaning that rewatch during exam week predicts less *Clarify* ($\beta = -0.336$) and less *Search* ($\beta = -3.711$). These results are revisited in the Discussion.

4.2.3. Blended course multilevel modeling

We repeated the analysis for data from the blended courses. In the multilevel model, interaction effects were non-significant, and their inclusion did not improve the fit of the model. As such we used the two-level model since it contributed the most to explaining the total variance of the blended courses data. Results show little impact of learning context on student engagement (see Table 7). Course week was found to predict more *Skim* ($\beta = 0.310$). Exam and rewatch had no relationship with engagement goals. Video duration

Table 4

Four main engagement goals, intents, and their associated video viewing activities, as identified in Study 1.

Engagement goal	Intent	Video viewing activity
Reflect	<ul style="list-style-type: none"> I want to write a note. I need time to think and reflect on what I just watched. 	Pause (A/C)
Clarify	<ul style="list-style-type: none"> I search for additional information using the web, textbook, and so on. I zoned out or got distracted, and want to make sure I did not miss anything. I did not get the explanation the first time. 	Rewind (A)
Skim	<ul style="list-style-type: none"> The speaker is talking too slow or fast. I find the content to be irrelevant, boring, or already known to me. Speeding up the playback saves me time. The speaker is talking too slow or fast. 	Reduce playback speed (A) Fast-forward (A) Increase playback speed (A)
Search	<ul style="list-style-type: none"> I am familiar with the content. To better understand what is being said in the video. I like to read the transcript while watching the video. To locate a specific piece of information to navigate to. I am looking for a keyword. 	Scrub (A)

Note: Each video viewing activity is labelled with its associated ICAP level(s): “A” for Active and “A/C” for Active or Constructive, depending on the situation.

Table 5

The average number (and standard deviation) of activities associated with engagement goals per session in the online and blended courses.

Engagement goals	Online course	Blended courses
Reflect	4.58 (6.77)	6.61 (12.81)
Clarify	1.00 (2.20)	1.61 (2.95)
Skim	2.01 (7.60)	2.43 (5.22)
Search	13.86 (23.16)	12.38 (17.02)

Table 6

Online course multilevel model for engagement goals in learning context.

Learning context	Engagement goals			
	Reflect	Clarify	Skim	Search
Main effects				
Course week	-0.086*** (0.025)	-0.010 (0.007)	-0.039 (0.020)	-0.280*** (0.069)
Exam	-1.456*** (0.181)	0.138 (0.057)	0.408 (0.193)	2.005*** (0.592)
Rewatch	-2.189*** (0.247)	0.134 (0.080)	0.212 (0.272)	4.906*** (0.833)
Video duration	0.003*** (0.001)	0.001*** (0.000)	0.002*** (0.000)	0.018*** (0.001)
Constant	5.347*** (0.366)	0.745*** (0.084)	1.036*** (0.246)	5.276*** (0.869)
Interaction effects				
Exam × Rewatch	0.941** (0.302)	-0.336*** (0.099)	-0.510 (0.336)	-3.711*** (1.027)

Note: Values are beta coefficients (standard errors); boldface indicates statistical significance after bonferroni correction; **p < .01; ***p < .001.

showed a statistically significant, but small, positive relationship with all four goals (beta <0.004).

4.3. Discussion

In this study, we examined the impact of learning context on students' engagement goals with online video. In the online course, course week, exam, and rewatch (i.e., watching previously watched videos), were associated with students' engagement with videos. However, this was not the case in the blended course. While the blended course showed an overall higher level of activity (that is, a higher number of activities associated with engagement goals), students in these courses showed less adaptive use of video. One possible explanation for this is the role of instructional videos in the different course formats. In the online course, videos are the primary source of information. Thus, students use video more strategically. It may be that students in the blended courses used the online videos only for specific purposes, like preparing for class or doing homework. This may explain why these students showed more activities per session overall, but their use of video was less sensitive to changes in the learning context. Another possible explanation is that students who choose an online course did so because they were more accustomed and skilled at learning from instructional video. Thus, they showed more strategic learning activity, one that adapts to contextual variables.

One variable showed a very consistent outcome, across all engagement goal types and course formats: video length was positively associated with the number of activities. However, the magnitude of this relationship was smaller than all other learning context variables. Even after multiplying by typical video duration (such as 20 min), the effect of video length remains a magnitude below the effect of other factors. That is, the increase in activity level did not grow linearly with video length. This may have several reasons. For one, students may be dropping from long videos, thus, essentially, turning them into much shorter videos (Guo et al., 2014). Second, it may be that students attend to specific information in video, and the amount of information does not grow linearly with length.

4.3.1. Video learning in online courses

Table 8 summarizes the results of the engagement goal-related activities in the online course. Overall, as course week progresses, we saw a decrease in students' level of activity. Specifically, activities that are associated with *Reflect* decreased as the term continued.

Table 7

Blended course multilevel model for engagement goals in learning context.

Learning context	Engagement goals			
	Reflect	Clarify	Skim	Search
Main effects				
Course week	0.159 (0.178)	0.116** (0.042)	0.310*** (0.075)	-0.306 (0.253)
Exam	-0.621 (0.446)	-0.121 (0.106)	0.095 (0.192)	-0.866 (0.627)
Rewatch	-1.129** (0.409)	0.031 (0.098)	0.288 (0.176)	-0.375 (0.574)
Video duration	0.002*** (0.000)	0.000*** (0.000)	0.001*** (0.000)	0.003*** (0.001)
Constant	3.289*** (0.724)	0.452** (0.169)	0.147 (0.297)	8.475*** (1.034)

Note: Values are beta coefficients (standard errors); boldface indicates statistical significance after bonferroni correction; **p < .01; ***p < .001.

It may be that students had more load and thus stopped less to reflect. Alternatively, students improved their learning habits, and were able to summarize what they needed from the video without stopping it. Students also *Search* less as the course week progresses, possibly because they can easily locate a specific piece of information to navigate to. It is likely that as students adapt to the expectations of the instructor, the role of the video in the class, and the affordances of the video player, they may become more efficient and focused, and consequently need fewer actions to achieve their goals.

As to be expected, during exam week, students *Search* more and *Reflect* less, probably in an attempt to be more efficient and mindful of their time. Online learning students *Search* more during the exam week when first watching videos, possibly because they spend more time and attention seeking video segments that they perceive are valuable for the exam. This is not the case when students first watch videos during weeks where there is no exam. They may also take fewer notes during the exam week, and thus *Reflect* less. This is interesting because this strategic engagement appears only in online courses and not in blended courses.

Students *Search* more and *Reflect* less when they rewatch videos. This is another sign of strategic use of video — they are familiar with these videos and return to them for a specific purpose. This can be explained by previous findings that suggest that students selectively pick parts to rewatch (Harandi et al., 2018; Kim et al., 2014). They also *Reflect* less, possibly since they take fewer notes when rewatching videos.

Students *Clarify* less during exam weeks when rewatching videos. *Clarify* activities, such as rewinding the video and reducing playback speed, increase engagement time with the video. One common intent for *Clarify* is that students were unable to understand the information in the video. Thus, students may *Clarify* less during exam weeks for previously watched videos because they already know the content. Students may also be mindful of their time in a week before the exam, which reduces the time that they choose to devote to clarifying video segments.

4.3.2. Video learning in blended courses

Students in the blended model showed an increase of *Skim* as the course week progressed. This may be an instance of increased time pressure — skimming ahead is a simple way to finish a video faster. It may also be an artifact of the video being a secondary source, hence students skim forward to find its added value on top of the lecture — and they learn to do so more efficiently as the term continues (hence more skimming). As noted above, overall, students' engagement patterns in the blended courses showed little adaptivity.

5. Overall Discussion and Conclusion

Our first research question focused on students' goals and intents when learning with online video. Although the growing body of engagement research has been adopting a learning analytics approach (e.g., Jovanovic et al., 2019; Kim et al., 2018; Vytasek et al., 2020), little has been done to map students' video viewing activities to video engagement goals and intents. We found complex and multi-dimensional links between engagement goals, intents, and associated video viewing activities. Our mixed-method approach allowed us to observe the many-to-many mappings between video viewing activities and engagement goals (see Fig. 1 and Table 2).

Our second research question sought to understand the relationship between learning context and students' engagement with video. We found that when students watch a video for the first time, they *Clarify* more, while on repeated watches, they *Search* more. During exam weeks, students *Clarify* less. Students' level of engagement drops with course week, possibly due to increased efficiency. That is, the reduction in activity level may reflect a better understanding of the utility of the video, and how it can be used efficiently. One clear finding is that students in online courses use video much more adaptively and strategically than students in blended courses.

Students in the study worked with a hypervideo player that supported advanced interactions such as annotation and highlighting. However, despite instructions and demonstrations, the vast majority of the students chose not to use these features. Interviews that we have conducted with students suggested that students still summarized videos — however, they chose to do so outside the dedicated interface (Dodson, Roll, Harandi, Fels, & Yoon, 2019). This may have been caused by usability issues with the (ViDeX) hypervideo player, or because most conventional video players (e.g., Netflix, Vimeo, and YouTube) do not support video annotations, and thus students may not be used to annotating videos. However, a more significant explanation is one of information management. Students preferred to have the video summaries available to them in their personal notebook, together with other sources of information, rather than being left on the video player (Dodson et al., 2019). The current study, which was limited to clickstream data from the video interface, falls short of analyzing these constructive activities. Similarly, the study did not evaluate interactive engagement (e.g., debating with a peer about the justifications; discussing similarities & differences). This kind of engagement is clearly beneficial for learning. However, its support by current technologies, pedagogies, and workflows is lacking; and when hypervideo features, such as

Table 8
Summary of online course multilevel modeling results.

Learning context	Engagement goals			
	Reflect	Clarify	Skim	Search
As Course Week goes on During Exam week	(-) A decrease	(-) A decrease during exam weeks for previously watched videos		(-) A decrease (+) An increase, mainly for first watched videos
When Rewatch previously watched videos				(++) A sharp increase

annotation, are made available, their frequency of use is often low (e.g., [Dodson et al., 2018a](#); [Dorn et al., 2015](#)). As shown by this set of studies, within the ICAP framework, most engagement (or at least most clickstream data) focuses on the Active level. More needs to be done to support and assess more involved forms of learning with video. The design of hypervideo environments targets this gap ([Evi-Colombo et al., 2020](#); [Pardo et al., 2015](#); [Risko, Foulsham, Dawson, & Kingstone, 2012](#)). The benefits of these technologies depend on students overcoming their old habits, as demonstrated by the minimal use of constructive affordances in the current study. Future studies could investigate instructional designs that promote interactive engagement, for example, the value of intentionally designing and integrating questions for discussions on the video content.

Several recommendations for instructional design can be identified based on the presented results. At their core, these recommendations recognize that students use video differently depending on their context. For example, the findings show that students search much more during exam weeks. During these times, instructors can post video summaries of the full-length videos. These summaries can be based on the instructors' priorities, or based on students' earlier viewing activities. Conversely, students sought to clarify content more often on first-watches. Instructors could support that in different ways, such as providing more detailed versions of videos, or directing students' attention to video segments that are crucial for understanding. For video systems that include detailed instructor-facing analytics, instructors could identify areas of confusion and address these in other interactions with students, such as during lectures. In addition, instructors could help their students better navigate their videos. Two goals were associated with navigating video: Orienting oneself to the structure of the video during first watches, and searching for specific content during rewatches. Both of these activities impose extraneous cognitive load that can distract students from the video content itself. Instructors could provide a table of contents and/or an index that describes the video content by time. Such tools could help students orient and search (and find) information more effectively. Our final recommendation for instructors is based on Study 1. Conventional video players (such as YouTube) support Active learning, yet fall short of supporting Constructive and Interactive learning. For courses that make heavy use of video, it is recommended to use hypervideo players that can support these types of activities, such as annotation and sharing of annotations ([Evi-Colombo et al., 2020](#)). Some of this learning may also occur offline, for example, by encouraging students to share, debate, and contrast their annotations. All of these suggestions are especially pertinent in online courses, for which video is the main learning resource and in which students are much more strategic in using it.

The study has several limitations. First, our data do not involve other meaningful activities that could not be traced by our hypervideo player (e.g., personal note-taking, debating or discussing with peers). The interplay between what happens online and offline is an important dimension to take into account for a holistic picture of students' learning with online video ([Dodson et al., 2019](#)). Future studies should look into activity that occurs outside of the hypervideo player.

The challenge of interpreting data may be broader than this study alone. Nearly all activities that are captured using learning analytics techniques are at the Active level of the ICAP framework. As the difference between the Active and Constructive levels is often manifested in semantically (e.g., the content of students' notes), inferring which activities are supporting Constructive learning is difficult. Lacking a technique for semantic analysis, one way to infer meaningful learning may be to move from analyzing frequencies of activities to sequences. Such analysis can shed light on complex mental processes, such as meaningful learning (cf. [Perez et al., 2017](#); [Sinha et al., 2014](#)).

Another limitation has to do with the qualitative nature of Study 1. As with all qualitative research, interpretation depends heavily on the perspective of the researchers. Repeating a similar process, by different researchers in a different set of courses, is of interest. The study seeks to associate between observable activities and engagement goals. Such inference is always limited. While some associations are clearer (e.g., between skipping forward and skimming), others would benefit from corroboration, possibly through interviews and observations. Drawing on previous work on engagement, especially theoretically-informed means of measuring engagement (e.g., [O'Brien & Toms, 2008](#); [2010](#)) could be a fruitful line of inquiry for better understanding the different video-based learning activities identified in Study 1. One artifact of our approach is that we chose to focus on the single most common goal for each activity. This, naturally, added noise to the interpretation process. Further, two goals (*Orient* and *Remember*) were left without associated activities, and thus removed from Study 2 analysis. While these goals were found less common in Study 1, their removal from Study 2 is a limitation of the study. Lastly, the study was done with a specific player, style of videos, and use in class. It is of interest to evaluate the dependence of the results on these important factors. Like any system, different video players have different affordances. Also, a well-designed video may be enough to encourage students' productive engagement without complicated video viewing activities. The impact of the video player and content on students' activities, in order to achieve their goals, need to be further studied.

Improved engagement is one key goal of intervention ([Dumford & Miller, 2018](#); [Halverson & Graham, 2019](#); [Seo, Fels, Kang, Jung, & Ryu, 2020b](#)). Results from this study suggest ways in which engagement may be supported. For example, online learning students *Search* more during exam week or when rewatching previously watched videos, possibly because they want to revisit video segments of interest. Thus, future video systems could emphasize certain affordances based on contextual variables. For example, on a first watch, they can provide an overview of the video, as students report that they scrub to achieve that. On a rewatch, they can ask the students what information they came to observe, turning the overview into a navigation aid (e.g., previously watched content by the same user, previously watched by the class in the aggregate, or content that students previously returned to). Hypervideo players with novel affordances have been put forward, such as hyperlinked markers that allow students to jump directly to the video segment of interest (e.g., [Cattaneo et al., 2018](#); [2019](#)), and research suggests these features may help students navigate videos in a cognitively more efficient way (e.g., [Sauli et al., 2018](#)). The findings in this study add to the knowledge base to enhance understanding of how students engage with video, and how educational technologies should be designed to assist engagement goals in a different learning context.

Finally, this study has implications for instructors. Unlike face-to-face lectures, where instructors can see their students and get instant feedback on their instruction, there are still many challenges for instructors in understanding students' video viewing activities ([Fong et al., 2019](#); [Giannakos, Chorianopoulos, & Chrisochoides, 2015](#); [Kim et al., 2014](#)). In teaching with video, there is a distinct

technology gap for instructors to gather information necessary to make informed decisions about how to structure their classes around video (Fong et al., 2019; Seo et al., 2020a). For example, a video in which students apply many *Skim* may be perceived as irrelevant by them. Similarly, a video that has more activities that relate to engagement goals such as *Clarify* may suggest that students struggle to understand the content of the video. The ability to unveil this kind of insight into how and why students are interacting with instructional video could then be used as a guideline for instructors to make adjustments to their teaching. For instance, an instructor could edit video segments where many students *Skim*, because that is a part of the video that students feel is irrelevant, boring, or already known to them. Also, an instructor could add more slides or explanations wherever students *Clarify* a lot, because that is an interval where students are struggling to understand. Overall, to better support student learning, not only we need to describe their activities, but it is also important to understand the rationale for the specific patterns.

Author contributions aection

Kyoungwon Seo: Conceptualization, Methodology, Investigation, Writing – original draft, Visualization, Project administration; Samuel Dodson: Conceptualization, Methodology, Investigation, Data curation, Writing – original draft; Negar M. Harandi: Resources, Data curation; Nathan Roberson: Validation, Formal analysis; Sidney Fels: Writing – review & editing, Supervision, Project administration, Funding acquisition; Ido Roll: Writing – review & editing, Conceptualization, Writing, Supervision, Funding acquisition.

Acknowledgements

The support of instructors was important in developing this publication. The authors declare that there is no conflict of interest regarding the publication of this article. The terms of this arrangement have been reviewed and approved by the University of British Columbia in accordance with its policy on objectivity in research.

Appendix A. Survey Prompt & Questions

Below are typical activities that students perform with ViDeX. For each question, please choose one or more answers that best explain your intent(s) for doing this activity. Please choose all answers that apply. If you almost never engage in one of these activities, please skip that question.

1. When watching an instructional video, I typically **pause** because ...
 - a. I need time to think and reflect on what I just watched.
 - b. I want to write a note.
 - c. Something else is grabbing my attention or I need a break.
 - d. I search for additional information using the web, textbook, and so on.
 - e. Other, please specify here:
2. When watching an instructional video, I typically **navigate forwards** because ...
 - a. I find the content to be irrelevant, boring, or already known to me.
 - b. I want to get the gist of the content before I watch the entire video, to know what to expect and decide whether to watch it.
 - c. I prefer to study the video at my own pace by jumping around, instead of watching linearly.
 - d. Other, please specify here:
3. When watching an instructional video, I typically **navigate backwards** because ...
 - a. I zoned out or got distracted, and want to make sure I did not miss anything.
 - b. I did not get the explanation the first time.
 - c. It helps me remember the content.
 - d. I can start from the end goal and understand the path that was taken to get there.
 - e. Other, please specify here:
4. When watching an instructional video, I typically **highlight sections of the video** because ...
 - a. It helps me remember what is said in the video.
 - b. I bookmark content that I want to return to later.
 - c. Later, it is easier to review my annotations than to rewatch the entire video.
 - d. I mark the first time that important concepts appear.
 - e. Other, please specify here:
5. When watching an instructional video, I typically **annotate on the video itself** because ...
 - a. I want to add relevant information that is not on the video.
 - b. I want to remember something important for an assignment or test.
 - c. I summarize the video to save study time.
 - d. Notes on (ViDeX) help me find the original videos.
 - e. Other, please specify here:
6. When watching an instructional video, I typically **search the filmstrip and/or transcript** because ...
 - a. To locate a specific piece of information to navigate to.
 - b. To better understand what is being said in the video.

- c. I like to read the transcript while watching the video.
 - d. I am looking for a keyword.
 - e. Other, please specify here:
7. When watching an instructional video, I typically **change the playback speed** (faster and/or slower) because ...
- a. The speaker is talking too slow or fast.
 - b. Speeding up the playback saves me time.
 - c. I am familiar with the content.
 - d. I am unfamiliar with the content.
 - e. Other, please specify here:

References

- Aldhafeeri, F. M. (2015). Blended learning in higher education. In *International handbook of E-learning, volume 2* pp. 121–132. Routledge.
- Allen, I. E., & Seaman, J. (2016). *Online report card: Tracking online Education in the United States. Babson survey research group*. Babson Park, MA: Babson College, 231 Forest Street, 02457.
- Andrist, L., Chepp, V., Dean, P., & Miller, M. V. (2014). Toward a video pedagogy: A teaching typology with learning goals. *Teaching Sociology*, 42(3), 196–206. <https://doi.org/10.1177/0092055X14524962>
- Artino, A. R., Jr., & Jones, K. D., II (2012). Exploring the complex relations between achievement emotions and self-regulated learning behaviors in online learning. *The Internet and Higher Education*, 15(3), 170–175. <https://doi.org/10.1016/j.iheduc.2012.01.006>
- Ashby, J., Sadara, W. A., & McNary, S. W. (2011). Comparing student success between developmental math courses offered online, blended, and face-to-face. *The Journal of Interactive Online Learning*, 10(3), 128–140. Retrieved from <https://www.learntechlib.org/p/111204/>.
- Axelson, R. D., & Flick, A. (2010). Defining student engagement. *Change: The Magazine of Higher Learning*, 43(1), 38–43. <https://doi.org/10.1080/00091383.2011.533096>
- Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn?: A taxonomy of far transfer. *Psychological Bulletin*, 128(4), 612–637. <https://doi.org/10.1037/0033-2909.128.4.612>
- Brame, C. J. (2016). Effective educational videos: Principles and guidelines for maximizing student learning from video content. *CBE-Life Sciences Education*, 15(4), es6. <https://doi.org/10.1187/cbe.16-03-0125>
- Broadbent, J. (2017). Comparing online and blended learner's self-regulated learning strategies and academic performance. *The Internet and Higher Education*, 33, 24–32. <https://doi.org/10.1016/j.iheduc.2017.01.004>
- Butler, D. L., & Cartier, S. C. (2005). Multiple complementary methods for understanding self-regulated learning as situated in context. Paper presented at the annual *Meetings of the American educational research association*. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/versions?doi=10.1.1.700.868>.
- Butler, D. L., Schnellert, L., & Perry, N. E. (2017). *Developing self-regulating learners*. New Jersey: Pearson Canada Incorporated.
- Cattaneo, A. A., van der Meij, H., Aprea, C., Sauli, F., & Zahn, C. (2019). A model for designing hypervideo-based instructional scenarios. *Interactive Learning Environments*, 27(4), 508–529. <https://doi.org/10.1080/10494820.2018.1486860>
- Cattaneo, A. A., van der Meij, H., & Sauli, F. (2018). An empirical test of three instructional scenarios for hypervideo use in a vocational education lesson. *Computers in the Schools*, 35(4), 249–267. <https://doi.org/10.1080/07380569.2018.1531597>
- 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>
- Chi, M. T., Adams, J., Bogusch, E. B., Bruchok, C., Kang, S., Lancaster, M., & Wylie, R. (2018). Translating the ICAP theory of cognitive engagement into practice. *Cognitive Science*, 42(6), 1777–1832. <https://doi.org/10.1111/cogs.12626>
- Chiu, P. S., Chen, H. C., Huang, Y. M., Liu, C. J., Liu, M. C., & Shen, M. H. (2018). A video annotation learning approach to improve the effects of video learning. *Innovations in Education & Teaching International*, 55(4), 459–469. <https://doi.org/10.1080/14703297.2016.1213653>
- Chi, M. T., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist*, 49(4), 219–243. <https://doi.org/10.1080/00461520.2014.965823>
- Chorianopoulos, K. (2018). A taxonomy of asynchronous instructional video styles. *International Review of Research in Open and Distance Learning*, 19(1), 295–311. <https://doi.org/10.19173/irrodl.v19i1.2920>
- Courtright, C. (2007). Context in information behavior research. *Annual Review of Information Science and Technology*, 41(1), 273–306. <https://doi.org/10.1002/aris.2007.1440410113>
- Deng, L., & Tavares, N. J. (2013). From Moodle to Facebook: Exploring students' motivation and experiences in online communities. *Computers & Education*, 68, 167–176. <https://doi.org/10.1016/j.compedu.2013.04.028>
- Dorn, B., Schroeder, L. B., & Stankiewicz, A. (2015). Piloting TRACE: Exploring spatiotemporal anchored collaboration in asynchronous learning. In *Proceedings of the 18th ACM conference on computer supported cooperative work & social computing* (pp. 393–403). <https://doi.org/10.1145/2675133.2675178>
- Dodson, S., Roll, I., Fong, M., Yoon, D., Harandi, N. M., & Fels, S. (2018b). An active viewing framework for video-based learning. In *Proceedings of the fifth annual ACM conference on learning at scale* (pp. 1–4).
- Dumford, A. D., & Miller, A. L. (2018). Online learning in higher education: Exploring advantages and disadvantages for engagement. *Journal of Computing in Higher Education*, 30(3), 452–465. Retrieved from <https://eric.ed.gov/?id=EJ1218398>.
- Dodson, S., Roll, I., Fong, M., Yoon, D., Harandi, N. M., & Fels, S. (2018a). Active viewing: A study of video highlighting in the classroom. In *Proceedings of the conference on human information interaction & retrieval* (pp. 237–240).
- Dodson, S., Roll, I., Harandi, N. M., Fels, S., & Yoon, D. (2019). Weaving together media, technologies and people. *Information and Learning Sciences*, 120(7/8), 519–549. <https://doi.org/10.1108/ILS-01-2019-0011>
- Evi-Colombo, A., Cattaneo, A., & Bétrancourt, M. (2020). Technical and pedagogical affordances of video annotation: A literature review. *Journal of Educational Multimedia and Hypermedia*, 29(3), 193–226. Retrieved from <https://www.learntechlib.org/primary/p/215718/>.
- Fong, M., Dodson, S., Zhang, X., Roll, I., & Fels, S. (2018). ViDeX: A platform for personalizing educational videos. In *Proceedings of the 18th ACM/IEEE on joint conference on digital libraries* (pp. 331–332).
- Fong, M., Dodson, S., Harandi, N. M., Seo, K., Yoon, D., Roll, I., & Fels, S. (2019). Instructors Desire Student Activity, Literacy, and Video Quality Analytics to Improve Video-based Blended Courses. In *Proceedings of the sixth (2019) ACM conference on learning@ scale* (pp. 1–10).
- Fredricks, J., McColskey, W., Meli, J., Mordica, J., Montrosse, B., & Mooney, K. (2011). *Measuring student engagement in upper elementary through high school: A description of 21 instruments*. Issues and Answers Report, 098 (pp. 26–27). Retrieved from <https://eric.ed.gov/?id=ED514996>.
- Giannakos, M. N., Chorianopoulos, K., & Chrisochoides, N. (2015). Making sense of video analytics: Lessons learned from clickstream interactions, attitudes, and learning outcome in a video-assisted course. *International Review of Research in Open and Distance Learning*, 16(1), 260–283. <https://doi.org/10.19173/irrodl.v16i1.1976>
- Graham, C. R. (2006). Blended learning systems. *The handbook of blended learning*, 3–21.

- Guo, P. J., Kim, J., & Rubin, R. (2014). How video production affects student engagement: An empirical study of MOOC videos. In *Proceedings of the first ACM conference on Learning@ scale conference* (pp. 41–50). <https://doi.org/10.1145/2556325.2566239>
- Halverson, L. R., & Graham, C. R. (2019). Learner engagement in blended learning environments: A conceptual framework. *Online Learning*, 23(2), 145–178. <https://doi.org/10.1007/s12528-018-9179-z>
- Harandi, N. M., Aghareparast, F., Linares, L., Dodson, S., Roll, I., Fong, M., & Fels, S. (2018). Student video-usage in introductory engineering courses. In *Proceedings of the Canadian engineering education association (CEEA)* (pp. 1–8).
- Hanington, B., & Martin, B. (2012). *Universal methods of design: 100 ways to research complex problems, develop innovative ideas, and design effective solutions*. Rockport Publishers.
- Henrie, C. R., Halverson, L. R., & Graham, C. R. (2015). Measuring student engagement in technology-mediated learning: A review. *Computers & Education*, 90, 36–53. <https://doi.org/10.1016/j.compedu.2015.09.005>
- Hew, K. F. (2016). Promoting engagement in online courses: What strategies can we learn from three highly rated MOOCs. *British Journal of Educational Technology*, 47(2), 320–341. <https://doi.org/10.1111/bjet.12235>
- Hockings, C., Cooke, S., Yamashita, H., McGinty, S., & Bowl, M. (2008). Switched off? A study of disengagement among computing students at two universities. *Research Papers in Education*, 23(2), 191–201. <https://doi.org/10.1080/02671520802048729>
- Jovanovic, J., Mirriahi, N., Gašević, D., Dawson, S., & Pardo, A. (2019). Predictive power of regularity of pre-class activities in a flipped classroom. *Computers & Education*, 134, 156–168. <https://doi.org/10.1016/j.compedu.2019.02.011>
- Kim, J., Guo, P. J., Seaton, D. T., Mitros, P., Gajos, K. Z., & Miller, R. C. (2014). Understanding in-video dropouts and interaction peaks in online lecture videos. In *Proceedings of the first ACM conference on Learning@ scale conference* (pp. 31–40). <https://doi.org/10.1145/2556325.2566237>
- Kim, D., Yoon, M., Jo, I. H., & Branch, R. M. (2018). Learning analytics to support self-regulated learning in asynchronous online courses: A case study at a women's university in South Korea. *Computers & Education*, 127, 233–251. <https://doi.org/10.1016/j.compedu.2018.08.023>
- Kizilcec, R. F., & Cohen, G. L. (2017). Eight-minute self-regulation intervention raises educational attainment at scale in individualist but not collectivist cultures. *Proceedings of the National Academy of Sciences*, 114(17), 4348–4353. <https://doi.org/10.1073/pnas.1611898114>
- Koedinger, K. R., Kim, J., Jia, J. Z., McLaughlin, E. A., & Bier, N. L. (2015). Learning is not a spectator sport: Doing is better than watching for learning from a MOOC. In *Proceedings of the second (2015) ACM conference on Learning@ scale* (pp. 111–120). <https://doi.org/10.1145/2724660.2724681>
- Kovacs, G. (2016, April). Effects of in-video quizzes on MOOC lecture viewing. In *Proceedings of the third (2016) ACM conference on learning@ scale* (pp. 31–40). <https://doi.org/10.1145/2876034.2876041>
- Lai, C. L., & Hwang, G. J. (2016). A self-regulated flipped classroom approach to improving students' learning performance in a mathematics course. *Computers & Education*, 100, 126–140. <https://doi.org/10.1016/j.compedu.2016.05.006>
- Lalmas, M., O'Brien, H., & Yom-Tov, E. (2014). Measuring user engagement. *Synthesis Lectures on Information Concepts, Retrieval, and Services*, 6(4), 1–132. <https://doi.org/10.2200/S00605ED1V01Y201410ICR038>
- Larson, D. K., & Sung, C. H. (2009). Comparing student performance: Online versus blended versus face-to-face. *Journal of Asynchronous Learning Networks*, 13(1), 31–42. Retrieved from <https://eric.ed.gov/?id=EJ837556>.
- Lee, V. E. (2000). Using hierarchical linear modeling to study social contexts: The case of school effects. *Educational Psychologist*, 35(2), 125–141. https://doi.org/10.1207/S15326985EP3502_6
- Liu, C., Kim, J., & Wang, H. C. (2018, April). ConceptScape: Collaborative concept mapping for video learning. In *Proceedings of the 2018 CHI conference on human factors in computing systems* (pp. 1–12). <https://doi.org/10.1145/3173574.3173961>
- Liu, C. J., Yang, C. L., Williams, J. J., & Wang, H. C. (2019, April). NoteStruct: Scaffolding note-taking while learning from online videos. *Extended abstracts of the 2019 CHI conference on human factors in computing systems* (p. LBW1116). <https://doi.org/10.1145/3290607.3312878>
- McNamara, D. S. (Ed.). (2007). *Reading comprehension strategies: Theories, interventions, and technologies*. Psychology Press.
- Means, B., Toyama, Y., Murphy, R., Bakia, M., & Jones, K. (2009). *Evaluation of evidence-based practices in online learning: A meta-analysis and review of online learning studies*. Project Report. Centre for Learning Technology.
- van der Meij, H., & Dunkel, P. (2020). Effects of a review video and practice in video-based statistics training. *Computers & Education*, 143, 103665. <https://doi.org/10.1016/j.compedu.2019.103665>
- Meir, E., Wendel, D., Pope, D. S., Hsiao, L., Chen, D., & Kim, K. J. (2019). Are intermediate constraint question formats useful for evaluating student thinking and promoting learning in formative assessments? *Computers & Education*, 141, 103606. <https://doi.org/10.1016/j.compedu.2019.103606>
- Michael, J. (2006). Where's the evidence that active learning works? *Advances in Physiology Education*, 30(4), 159–167. <https://doi.org/10.1152/advan.00053.2006>
- Mitrovic, A., Dimitrova, V., Lau, L., Weerasinghe, A., & Mathews, M. (2017). Supporting constructive video-based learning: Requirements elicitation from exploratory studies. In *International conference on artificial intelligence in education* (pp. 224–237). Springer. https://doi.org/10.1007/978-3-319-61425-0_19. Retrieved from.
- Mitrovic, A., Gordon, M., Piotrkowicz, A., & Dimitrova, V. (2019). Investigating the effect of adding nudges to increase engagement in active video watching. *International conference on artificial intelligence in education*. https://doi.org/10.1007/978-3-030-23204-7_27
- Nortvig, A. M., Petersen, A. K., & Balle, S. H. (2018). A literature review of the factors influencing E-learning and blended learning in relation to learning outcome, student satisfaction and engagement. *Electronic Journal of E-Learning*, 16(1), 46–55. Retrieved from <https://eric.ed.gov/?id=EJ1175336>.
- O'Brien, H. L., & Toms, E. G. (2008). What is user engagement? A conceptual framework for defining user engagement with technology. *Journal of the American Society for Information Science and Technology*, 59(6), 938–955. <https://doi.org/10.1002/asi.20801>
- O'Brien, H. L., & Toms, E. G. (2010). The development and evaluation of a survey to measure user engagement. *Journal of the American Society for Information Science and Technology*, 61(1), 50–69. <https://doi.org/10.1002/asi.21229>
- Pardo, A., Mirriahi, N., Dawson, S., Zhao, Y., Zhao, A., & Gašević, D. (2015). Identifying learning strategies associated with active use of video annotation software. In *Proceedings of the fifth international conference on learning analytics and knowledge* (pp. 255–259). <https://doi.org/10.1145/2723576.2723611>
- Pavel, A., Reed, C., Hartmann, B., & Agrawala, M. (2014). Video digests: A browsable, skimmable format for informational lecture videos. *UIST*, 10, 2642918–2647400. <https://doi.org/10.1145/2642918.2647400>
- Perez, S., Massey-Allard, J., Butler, D., Ives, J., Bonn, D., Yee, N., et al. (2017). Identifying productive inquiry in virtual labs using sequence mining. *International conference on artificial intelligence in education*. https://doi.org/10.1007/978-3-319-61425-0_24
- Prevost, L. B., Smith, M. K., & Knight, J. K. (2016). Using student writing and lexical analysis to reveal student thinking about the role of stop codons in the central dogma. *CBE-Life Sciences Education*, 15(4), ar65. <https://doi.org/10.1187/cbe.15-12-0267>
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Vol. 1). sage.
- Roll, I., Russell, D. M., & Gašević, D. (2018). Learning at scale. *International Journal of Artificial Intelligence in Education*, 28(4), 471–477. <https://doi.org/10.1007/s40593-018-0170-7>
- Risko, E. F., Foulsham, T., Dawson, S., & Kingstone, A. (2012). The collaborative lecture annotation system (CLAS): A new tool for distributed learning. *IEEE Transactions on Learning Technologies*, 6(1), 4–13. <https://doi.org/10.1109/TLT.2012.15>
- Roll, I., & Winne, P. H. (2015). Understanding, evaluating, and supporting self-regulated learning using learning analytics. *Journal of Learning Analytics*, 2(1), 7–12. <https://doi.org/10.18608/jla.2015.21.2>
- Sauli, F., Cattaneo, A., & van der Meij, H. (2018). Hypervideo for educational purposes: A literature review on a multifaceted technological tool. *Technology, Pedagogy and Education*, 27(1), 115–134. <https://doi.org/10.1080/1475939X.2017.1407357>
- Seo, K., Fels, S., Yoon, D., Roll, I., Dodson, S., & Fong, M. (2020a). Artificial intelligence for video-based learning at scale. In *Proceedings of the ACM conference on learning @ scale* (pp. 215–217).
- Seo, K., Fels, S., Kang, M., Jung, C., & Ryu, H. (2020b). Goldilocks conditions for workplace gamification: how narrative persuasion helps manufacturing workers create self-directed behaviors. *Human-Computer Interaction*, 1–38. <https://doi.org/10.1080/07370024.2020.1744145>
- Schreier, M. (2012). *Qualitative content analysis in practice*. Sage publications.

- Shin, H. V., Berthouzoz, F., Li, W., & Durand, F. (2015). Visual transcripts: Lecture notes from blackboard-style lecture videos. *ACM Transactions on Graphics*, 34(6), 1–10. <https://doi.org/10.1145/2816795.2818123>
- Sinha, T., Jermann, P., Li, N., & Dillenbourg, P. (2014). *Your click decides your fate: Inferring information processing and attrition behavior from mooc video clickstream interactions*. arXiv preprint arXiv:1407.7131. Retrieved from <https://arxiv.org/abs/1407.7131>.
- van der Sluis, F., Van der Zee, T., & Ginn, J. (2017). Learning about learning at scale: Methodological challenges and recommendations. In *Proceedings of the fourth (2017) ACM conference on learning@ scale* (pp. 131–140).
- Soffer, T., & Cohen, A. (2019). Students' engagement characteristics predict success and completion of online courses. *Journal of Computer Assisted Learning*, 35(3), 378–389. <https://doi.org/10.1111/jcal.12340>
- Tessmer, M., & Richey, R. C. (1997). The role of context in learning and instructional design. *Educational Technology Research and Development*, 45(2), 85–115. <https://doi.org/10.1007/BF02299526>
- Van Doorn, J. R., & Van Doorn, J. D. (2014). The quest for knowledge transfer efficacy: Blended teaching, online and in-class, with consideration of learning typologies for non-traditional and traditional students. *Frontiers in Psychology*, 5, 324. <https://doi.org/10.3389/fpsyg.2014.00324>
- Vieira, I., Lopes, A. P., & Soares, F. (2014). The potential benefits of using videos in higher education. In *Proceedings of EDULEARN14 conference* (pp. 750–756). IATED Publications. Retrieved from <http://hdl.handle.net/10400.22/4853>.
- Vytasek, J. M., Patzak, A., & Winne, P. H. (2020). Analytics for student engagement. In *Machine learning paradigms* (pp. 23–48). Cham: Springer. https://doi.org/10.1007/978-3-030-13743-4_3.
- Xia, J., & Wilson, D. C. (2018, February). Instructor perspectives on comparative heatmap visualizations of student engagement with lecture video. In *Proceedings of the 49th ACM technical symposium on computer science education* (pp. 251–256). <https://doi.org/10.1145/3159450.3159487>
- Yousef, A. M. F., Chatti, M. A., & Schroeder, U. (2014). The state of video-based learning: A review and future perspectives. *Int. J. Adv. Life Sci*, 6(3/4), 122–135.
- Zhu, Y., Pei, L., & Shang, J. (2017). Improving video engagement by gamification: A proposed design of MOOC videos. In *International conference on blended learning* (pp. 433–444). https://doi.org/10.1007/978-3-319-59360-9_38