Communications of the Association for Information Systems

Volume 53 Paper in press

2023

Learner Engagement with YouTube Videos in Informal Online Learning: An Investigation of the Effects of Segmenting, Signaling, and Weeding

Zixing Shen
New Mexico State University, zshen@nmsu.edu

Communications of the Association for Information Systems

Accepted Manuscript

Learner Engagement with YouTube Videos in Informal Online Learning: An Investigation of the Effects of Segmenting, Signaling, and Weeding

Zixing Shen
College of Business
New Mexico State University
zshen@nmsu.edu

Please cite this article as: Shen, Z. (in press). Learner engagement with YouTube videos in informal online learning: An investigation of the effects of segmenting, signaling, and weeding. *Communications of the Association for Information Systems.*

Communications of the Association for Information Systems

Education Paper ISSN: 1529-3181

Learner Engagement with YouTube Videos in Informal Online Learning: An Investigation of the Effects of Segmenting, Signaling, and Weeding

Zixing Shen
College of Business
New Mexico State University
zshen@nmsu.edu

Abstract:

Millions of educational videos available on YouTube offer unprecedented opportunities for online learning. As it invites open-ended and self-paced exploration of almost any topic, Tube has emerged as an important platform for informal online learning that occurs outside the formal classroom. A considerable number of studies have been directed toward YouTube educational videos. However, research on learner engagement with YouTube educational videos is limited, despite the central role of engagement in learning and the increasing popularity of YouTube videos in informal online learning. This paper addresses this research gap. We adopt the conceptualization that learner engagement has three dimensions – behavioral, emotional, and cognitive - and investigate how the features of segmenting, signaling, and weeding (SSW), the three multimedia learning principles, in YouTube educational video presentations collectively affect learner engagement in informal online learning. Our analysis shows that different SSW features have various associations with the three dimensions of learner engagement. These findings substantiate the empirical knowledge of learner engagement with YouTube educational videos. Our study corroborates extant video engagement research and extends its relevance to informal learning on social media. It also informs video designers and developers on adding video presentation features to optimize video engagement on social media.

Keywords: Learner Engagement, Educational Videos, YouTube, Segmenting, Signaling, Weeding, Informal Online Learning, Social Media.

[Department statements, if appropriate, will be added by the editors. Teaching cases and panel reports will have a statement, which is also added by the editors.]

[Note: this page has no footnotes.]

This manuscript underwent [editorial/peer] review. It was received xx/xx/20xx and was with the authors for XX months for XX revisions. [firstname lastname] served as Associate Editor.] **or** The Associate Editor chose to remain anonymous.]

1 Introduction

YouTube boasts a vast number of videos with educational purposes that provide scientific explanations of a phenomenon or demonstrations of an expert procedure. The staggeringly large corpus of YouTube educational videos invites self-paced exploration on almost any topic (e.g., applying makeup, computer programming, dancing, fixing cars, language, and music). By offering unprecedented learning opportunities, YouTube has arguably become a viable venue for learning (Juhasz, 2011; Zhou et al., 2020).

A considerable amount of empirical research has been directed toward YouTube educational videos. Earlier research has investigated the quality (e.g., Azer et al., 2013) and acceptance of YouTube videos (e.g., Lee & Lehto, 2013) as learning tools. More recent studies have probed into learning outcomes, such as learner satisfaction (e.g., Hong et al., 2016), perceived learning (e.g., Wang & Chen, 2020), and knowledge construction (e.g., Dubovi & Tabak, 2020). However, research on learner engagement with YouTube videos is scarce.

Understanding learner engagement with YouTube videos is critical. Engagement is an essential aspect of any learning context and a prerequisite for learning (Bulger et al., 2008). Growing evidence shows that engagement is pivotal in video-based learning (Seo et al., 2021) and online learning (Martin et al., 2020). A higher level of engagement is associated with positive learning experiences (Mello, 2016). In contrast, disengagement has been linked to adverse effects on learning outcomes (Hershcovits et al., 2020). Given the central role of engagement in learning (Fredricks et al., 2004) and increasing interest in video-based learning in online settings (Sablić et al., 2021), systematic research on learner engagement with YouTube videos is urgently needed.

This study answers this call. Specifically, we look into learner engagement with YouTube videos in informal online learning. While formal learning occurs in educational institutions that award credentials, is instructor-led, and cover an organized curriculum, informal learning is not externally mandated but is learner-controlled, exploratory, and self-directed (Greenhow & Lewin, 2016). YouTube videos not only complement and supplement student learning in the formal classroom (Clifton & Mann, 2011; Buzzetto-More, 2015) but also inform people who seek information and knowledge informally (Allgaier, 2019). They are prevalent in informal online learning (Tan, 2013, Lee et al., 2017), unstructured learning that takes place in a virtual space outside the formal classroom, as they meet the need for just-in-time, self-regulated access to distributed information as well as for flexible use of this information according to the individual's interests and needs (Hong et al., 2016). YouTube can also facilitate informal online learning by finding people with similar interests, sharing information and idea, and learning from each other (Vizcaíno-Verdú et al., 2019). As such, people are increasingly turning to YouTube for informal online learning (Lange, 2019). Yet, a recent review indicates that most research is dedicated to YouTube in formal learning, especially at the college level (Shoufan & Mohamed, 2022).

In this study, we look into informal online learning on YouTube. Drawing up the multimedia learning literature, we investigate the effects of segmenting, signaling, and weeding (SSW), three principles of multimedia learning (Mayer 2009), on learner engagement with YouTube videos in informal online learning. Segmenting breaks the learning materials into smaller pieces; signaling adds cues to highlight the essential elements in the learning material; weeding eliminates distracting and irrelevant information that does not contribute to the learning goal (Mayer, 2009). We choose to study SSW because of their particular relevance to informal online learning. Compared with formal learning, informal learning lacks support from instructors, which increases the demand to control the information load and identify relevant information on the learners. We argue that SSW can support engagement by managing information load, directing attention, and making more cognitive resources available for processing information. We identify video features that embody SSW, theorize their effects on learner engagement, and test the hypotheses.

Our study fills a gap in research on learner engagement with YouTube educational videos in informal online learning. By revealing the effects of video features, this study contributes to understanding engagement with social media videos in the learning context. In addition, it extends the application of SSW from learning outcomes to engagement, the antecedent of learning outcomes. Our findings can also inform educational video designers and developers to optimize engagement with the appropriate use of video features, create more conducing learning tools, and make more positive learning experiences on social media. They can help better understand and anticipate the reaction to the educational content and manage social media presence more effectively.

The rest of the paper is organized as follows. Section 2 reviews related work on learner engagement, educational videos, and multimedia learning and SSW. Section 3 proposes the research hypotheses. Section 4 describes the method, including sampling and measurement. Section 5 presents the analysis and results. Section 6 discusses the paper's research contributions, practice implications, limitations, and future research. Section 7 concludes the paper.

2 Related Work

2.1 Learner Engagement

Engagement is a key construct in the learning literature (Fredricks et al., 2004). It has been extensively studied since the 1980s (Appleton et al., 2008). Various conceptualizations have emerged, and engagement in learning has been defined in many different ways. While earlier work conceives engagement as a unidimensional construct (McIntyre et al., 1983), more recent research views it as multi-dimensional (Furlong et al., 2003). Engagement dimensions are often categorized into behavioral, emotional, and cognitive (Fredricks et al., 2004).

The behavioral dimension is tied to participation and involvement in educational activities and includes attending classes, listening to lectures, and completing assignments (Fredricks et al., 2004). The emotional dimension encompasses positive and negative reactions to learning content and context (Xu et al., 2020). The former refers to learners' interest, enjoyment, boredom, and anxiety for learning activities, and the latter involves their feelings towards the instructor, peers, course, and the learning environment (Fredricks & McColskey, 2012). Cognitive engagement is often interpreted as learners' mental investment in the learning experience (Furlong et al., 2003). It describes the extent to which learners are willing to undertake the learning task and how long they persist in learning (Skinner & Pitzer, 2012) and is, therefore, linked directly to achievement (Greene, 2015). When cognitively engaged, learners are attentive, work hard, and strive to master the intended knowledge and skills (Fredricks et al., 2004).

Research from past decades in formal learning shows that learner engagement has short-term and long-term impacts on students (Hew, 2016). In the short term, learner engagement is predictive of grades and conduct in school (Voelkl, 1997). In the long run, learner engagement can be linked to individual achievement and self-esteem (Maddox & Prinz, 2003). Learner engagement is also found to be affected by factors related to the learner (e.g., personal interest or curiosity), the instructor (e.g., instructor accessibility and passion), and the course design (e.g., opportunities for peer interaction) (Breslow et al., 2013; Fredricks et al., 2004).

Based on the literature, this study construes learner engagement as the amount of attention, interest, emotional energy, and mental resources invested in learning (Jensen et al., 2016). We consider informal online learning with YouTube educational videos as intentional, as learners have the purpose of learning something even before the learning process begins and as conscious as the learner is aware that they have learned something. We view engagement in informal online learning with YouTube educational videos as behavioral, emotional, and cognitive connections with the video and peers on YouTube.

2.2 Educational Videos

The advances in information technology, particularly the affordance of mobile phones to phones to record, disseminate, and access video, have made videos very popular. The explosive growth of social media platforms explicitly formed around videos (e.g., YouTube) has also given rise to videos on social media. Videos nowadays permeate the increasingly digital world. They are of growing relevance in many areas, including but not limited to digital marketing (Tafesse, 2020), e-commerce (López-Nores et al., 2013), and gaming (Huang et al., 2019).

Educational videos have been widely studied in the 1970s and 1980s and remain a topic of interest in certain communities, like vocational training and professional development (Bétrancourta & Benetosa, 2018). Thanks to advances in digital technologies in the past two decades, video has made a remarkable comeback in education. Educational videos are now used as a content resource and play a pivotal role in the advent of the flipped classroom and Massive Open Online Courses (MOOCs).

The popularity of videos in learning has reignited research interest in educational videos. The dominant dependent variables are related to the learning outcomes (e.g., perceived learning, information recall, knowledge transfer, and learner satisfaction). Most studies have investigated how one particular video feature, such as social interaction (Andel et al., 2020), instructor presence (van Wermeskerken et al., 2018), and camera viewpoint (Boucheix et al., 2018), affects learning outcomes measured by test performance. They are conducted primarily in the laboratory using short multimedia examples and with college students as research participants. Research on YouTube educational videos mainly adopts the field study method to understand the relationships between learner factors (e.g., interest, self-efficacy) and perceived learning outcomes by learners via surveys (e.g., Hong et al., 2016) and interviews (e.g., Wang & Chen, 2020).

A growing but limited number of studies have taken learner engagement as a focal variable in online learning (e.g., Guo et al., 2014; Hew, 2016; Albó et al., 2019; Shao & Chen, 2021). They examine the factors contributing to learner engagement with videos in different courses in formal learning. Their findings corroborate that factors related to the learner, instructor, and course are associated positively with learner engagement with videos in online learning (Hew, 2016; Shao & Chen, 2021). More importantly, they also reveal that video features can affect learner engagement. For example, Guo and colleagues (2014) observe that informal talking-head and Khan-style videos are more engaging than pre-recorded classroom lectures, PowerPoint slides, or code screencasts in MOOCs. They also discover that engagement increases when the video is short (Guo et al., 2014). We continue this line of research on video features by probing into SSW features in the multimedia learning literature.

2.3 Multimedia Learning and Segmenting, Signaling, and Weeding

Educational videos fall under the umbrella of multimedia learning, which combines different media (e.g., text, sound, graphics, and amination) into learning. In multimedia learning, learners are exposed to three types of cognitive load - intrinsic, extraneous, and germane (Sweller, 1988; Sweller, Van Merrienboer, & Paas, 1998). Intrinsic load (aka essential processing) is related to the difficulty of the subject under study (i.e., the learning content). Extraneous load (aka incidental processing) is evoked by the learning material but does not directly contribute to learning. Germane load (aka generative processing) is the level of cognitive activity necessary to reach the desired learning outcome imposed by learning processes. Moreover, learners actively carry out a coordinated set of cognitive processes during multimedia learning. They engage in three different cognitive processes: (1) selecting - paying attention to important elements in the learning materials for further processing; (2) organizing - mentally incorporating the new information into a coherent cognitive structure; (3) integrating - mentally connecting the new information with existing relevant knowledge (Mayer, 2003; 2009).

Multimedia material is more likely to create a meaningful learning experience if the material is developed with the above cognitive loads and processes in mind. In multimedia learning, information comes from the verbal (auditory) channel and the visual (ocular) channel concurrently. The dual-channel information processing can incur a heavy cognitive load (Mayer, 2009). The continuous flow of information in the video also generates a heavy cognitive load due to the lack of information permanence (Leahy & Sweller, 2005). As humans can only process a limited amount of information at any given moment, the heavy cognitive load needs to be managed to facilitate learning. The intrinsic cognitive load should be carefully structured, especially when the material has a high intrinsic load; the extraneous load should be minimized, and the germane load should be optimized. The cognitive processes of selecting, organizing, and integrating should be supported to maximize learning effectively. To this end, various multimedia learning principles are distilled to help design multimedia materials better aligned with the human cognitive structure and process (Mayer 2003, 2009). In this study, we tap into three of them - segmenting, signaling, and weeding (SSW).

Segmenting can reduce the cognitive load by allowing learners to access information presented in segments rather than in one long continuous stream. This principle can be implemented by giving learners control over how they interact with the video. For example, learners can modify the pace of video presentations by selecting the content (Scheiter & Gerjets, 2007) and clicking a pause button (Biard et al., 2018). Segmenting can also be implemented by system-paced segmentation of the information flow (Wouters et al., 2007) imposed by the video. System-paced segmentation can be achieved by breaking the video into shorter segments with a break in-between (Ibrahim et al., 2012). Both learner control and system-paced segmentation aim to regulate the amount of information exposed to learners.

The signaling principle stipulates that humans learn better when shown precisely what to pay attention to on the screen. It is implemented by adding cues that signal the main ideas and concepts of the learning material (Mayer, 2003). Signaling does not provide new information or change the content of the learning materials. Instead, it manipulates the visuospatial characteristics of the learning materials to draw attention to essential elements of the visual representation. Generally, signaling can be text- or picture-based (Mayer, 2009). Text-based signals consist of texts that are inserted into learning materials. They include but are not limited to headings, summaries, and typographical cues (Lorch, 1989). Picture-based signals, also known as visual signals, include explicit symbols (like arrows) and techniques (like flashes) overlaid onto a graphic. They can be shapes, dynamic lines, highlight coloring, or moving dots added in the video.

The weeding principle, also known as the coherence principle, states that humans learn better when extraneous, distracting material is not included. It suggests that all visual and audio elements must directly support learning goals. Any information irrelevant to the learning goal can negatively affect cognitive processing and should be excluded from the materials. Extraneous information should be eliminated, even if it contains interesting and potentially motivating elements, such as illustrations, music, or sounds (Moreno & Mayer, 2007). Mayer and Moreno (2003) use the term "weeding" as the need to uproot video, graphics, words, or sounds that are not central to the learning objective.

SSW have received extensive research over the years in educational videos. Empirical research has demonstrated that when applied individually (e.g., Biard et al., 2018; Wang et al., 2020) or collectively (e.g., Ibrahim et al., 2012), SSW can effectively improve learning outcomes. Like other research on educational videos, the research on SSW is exclusively interested in learning outcomes and is carried out in experiments with college students in formal learning settings. In fact, the multimedia learning literature generally focuses on learning outcomes with the experiment research method (Mayer, 2009). This study investigates the effects of SSW on learner engagement, the antecedent of learning outcomes in informal online learning on YouTube. We pose the following research question: How do SSW affect learner engagement with YouTube educational videos in informal online learning?

3 Hypothesis Development

Learners execute the three cognitive processes of selecting, organizing, and integrating by sharing attention between them (Barrouillet & Camos, 2007). They also need to pay attention to processing new information. They do so by quickly and repeatedly switching their attention back and forth between the three different cognitive processes. Segmenting limits the incoming information, minimizes the need to switch attention constantly, and enables learners to devote their attention to maintaining the information shown in the video. That is, segmenting facilitates more attention in both maintaining and processing the learning material. As attention is central to engagement (O'Brien & Toms, 2008), segmenting supports engagement.

Segmenting can have additional effects on engagement. As suggested by prior research (Jacques et al., 1995), if learners do not feel the video is navigable or uncomfortable with its structure, they will not be sufficiently engaged to continue to watch it. Learner control allows learners to select the content. Enabling learners to skip certain parts they already understand and jump to another part they want to work on more, learner control allows learners to adjust the presentation pace to their individual needs. As such, learner control supports more mental resources for processing and consolidating the learner-perceived important information. As they allocate more mental resources, a key aspect of engagement (Jensen et al., 2016), learners are more engaged in learning.

Furthermore, segmenting has an impact on cognitive load. Segmenting divides the learning content into more intellectually manageable chunks and lowers the intrinsic load. As intrinsic load is reduced, more mental resources are available for generative processing (germane load). Taking all the above together, we expect that segmenting enhances learner engagement.

Hypothesis 1: Segmenting is associated positively with learner engagement.

Properly using signals can guide learners' attention to facilitate the selection and extraction of essential information (de Koning et al., 2009). Studies have suggested that signals capture and hold attention (MacInnis et al., 1991). Textual signals (e.g., typographical signals such as italics and boldface) direct learners' visual attention (i.e., viewing pattern) to relevant words, making the words and the concepts they describe visually distinguishable from other text (Mautone & Mayer, 2001). Likewise, visual signals purposefully lead the learners' eyes to the prominent features on the screen (Wang et al., 2020). Furthermore, textual and visual signals can both strengthen the tendency to maintain attention to the selected elements, as research shows that learners who studied the signaled materials attended longer and more frequently to relevant parts than learners who studied the non-signaled materials (Ozcelik et al., 2010; Jamet, 2014).

Signaling emphasizes particular aspects of the visual display and simplifies learners' decisions about which information is relevant. It reduces visual search and the unnecessary load associated with locating relevant information in the selecting cognitive process. By lowering extraneous processing (extraneous load) in the perceptual processing (i.e., selecting), signaling frees more cognitive resources for generative processing (germane load) in the conceptual processing (i.e., organizing and integrating).

Moreover, signaling can support the cognitive process of organizing and integrating directly. Textual signals can cue the structure of the materials and help with information organization. For example, headings are helpful with topic shifts and connections (Lorch & Lorch, 1996). They may also label the main concepts and contribute to constructing a globally coherent representation of the materials (Mautone & Mayer, 2001). Similarly, visual signals can highlight an objective's (sub)parts and their spatial connection (Tversky et al., 2002) and link to the corresponding textual information (Huk & Steinke, 2007).

Furthermore, visual signals can represent unclear or overly implicit relations in the presentation. Research has shown that visual cues can explicate temporal relations (Lowe & Boucheix, 2007), illustrate causal relations (Tversky et al., 2000), and demonstrate associations (Shah et al., 1999). In other words, visual signals can effectively indicate links between related elements and help build an integrated mental representation, thus supporting the integrating cognitive process. As the on-screen use of signals can induce more mental resources for the conceptual processing of learning content, we predict that signaling increases engagement.

Hypothesis 2: Signaling is associated positively with learner engagement.

Including irrelevant information competes for limited cognitive resources and often primes learners for incidental processing (extraneous load). Weeding can reduce the extraneous load by improving relevance and free up more mental resources for essential and generative processing (intrinsic and germane loads). Thus, we propose the following hypothesis:

Hypothesis 3: Weeding is associated positively with learner engagement.

4 Methods

4.1 Data Collection

We chose to study learner engagement with YouTube science videos because YouTube has emerged as a fertile ground for informal online learning of conceptually rich domains such as science (Dubovi & Tabak, 2020). YouTube has indeed become a viable tool for distributing scientific knowledge to the general public (Kohler & Dietrich, 2021). We sampled YouTube videos on 20 physics and astronomy topics, such as the big bang, black holes, dark energy, and quantum gravity. Learners of the sampled YouTube videos voluntarily wanted to educate themselves and enhance their knowledge. They had different levels of domain knowledge and pursued a common interest in the physics and astronomy topic covered in the YouTube video.

We searched the YouTube site with the topic name (e.g., dark energy) as a search term and followed the research precedence (e.g., Cranwell et al., 2015) to include the first 40 English-speaking videos with an educational focus rated by the YouTube ranking algorithm. While YouTube's search algorithm exhibits enough video variation within the top 20 search results (Rieder et al., 2018), the first 40 videos would further the variation and better represent the educational videos.

Next, each video was assigned a number on the seven-Likert scale (one stands for least popular and seven for most popular) based on its ranking from the YouTube search. We put the top five or six search results on the Likert scale of 7 (most popular), the next five or six on the Likert scale of 6 (more popular), so on and forth, and finally, the bottom five or six on the Likert scale of 1 (least popular). Then we randomly picked one video on each Likert scale. Seven videos were chosen for one topic to represent the 40 videos at different points of the popularity spectrum. This process was conducted across all 20 topics. In total, 140 videos were included for data analysis.

4.2 Measurement

Learner engagement is most often gauged by learners' perceptions of their learning experiences. Scales have been developed to elicit self-reports (e.g., Appleton et al., 2006). Some have also observed learners' overt learning behaviors to infer engagement (Chi & Wylie, 2014). Fewer studies have examined engagement through physiological measures such as eye-tracking (e.g., Jensen et al., 2016). We measured engagement by coding observable social media activities, such as viewing, voting, and commenting, as social media analytics are widely used in research on engagement on social media (e.g., Liikkanen & Salovaara, 2015; Khan, 2017; Oh et al., 2017; Xu et al., 2019).

The number of views shows how many times a video has been watched and is suitable to proxy the behavioral engagement that focuses on attendance. On YouTube, the voting control, which includes thumbs-up (likes) and thumbs-down (dislikes), allows viewers to express their feelings toward viewing experiences. Like indicates positive feelings and dislike negative feelings. The number of votes (the sum of the number of likes and dislikes) was used to operationalize emotional engagement.

In formal online learning, online discussions can detect cognitive engagement, as they indicate learners' attention to and analysis and synthesis of the learning material (Zhu, 2006). However, online discussions on YouTube (YouTube comments) do not all reflect the mental resources invested in learning as cognitive engagement is conceptualized in this study. Some of the different categories of YouTube comments (e.g., information, opinion, response, general conversation, and site process) (Madden et al., 2013) are unrelated to cognitive engagement. For example, the general conversation category includes greetings, thanking, giving, and requesting personal information.

To better capture cognitive engagement, we tapped into Reddit comments. Reddit is one of the most popular platforms for content sharing and discussion. Users can post news, questions, and other information in the form of text, images, links, and videos. Posts are organized by subject into user-created boards called "communities" or "subreddits". Users can also comment on posts. Comments consist of different perspectives and interpretations from users and help them analyze and engage with the content. Comments are a vital and valuable feature of Reddit.

Though some Reddit comments have content related to non-topic socializing (e.g., "great feedback") and subreddit rules and norms (e.g., "this post doesn't belong here"), the vast majority of Reddit comments focus on seeking information, commenting on previous comments, adding new ideas or facts to the discussion, and supporting answers through providing references. A content analysis of the AskScience subreddits shows that all these comments had content related to explanation with disagreement, explanation with agreement, explanation with neutral presentation, information seeking, or providing resources (Haythornthwaite et al., 2018). As viewers use Reddit comments to explain, discuss, explore, and debate video content, Reddit comments indicate cognitive engagement.

Moreover, Reddit has a much higher contributory level than YouTube. While YouTube is a social networking site for video viewing, the uptake by viewers has been mainly as consumers, simply watching the videos posted rather than contributing as registered account holders (Chau, 2010). Reddit is a discussion forum with a high contributory level. Compared with YouTube comments (Figure 1), Reddit comments (Figure 2) are more extended and in-depth exchanges of information and ideas. The lengthy and vibrant discussions on Reddit demonstrate the learners' mental efforts in learning, thus indicating cognitive engagement.

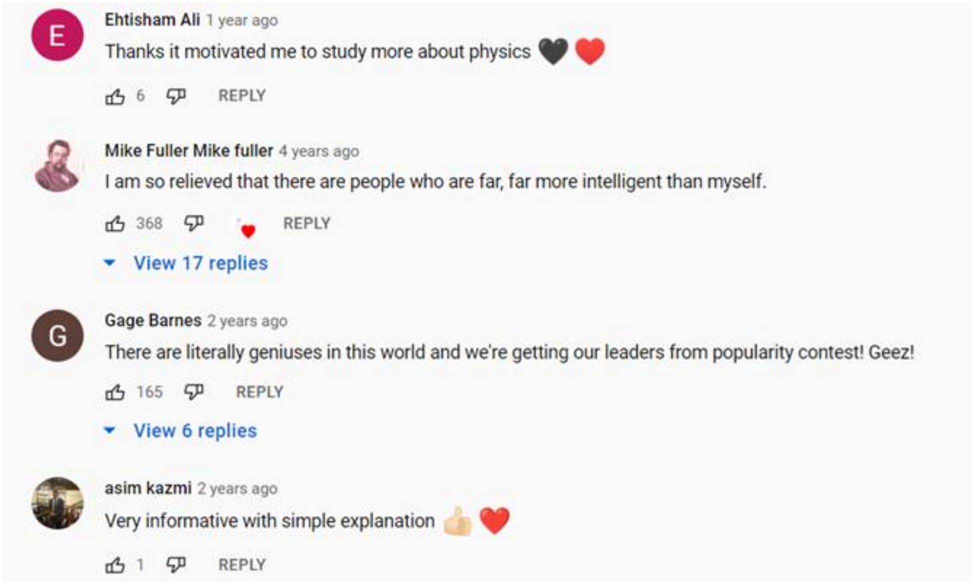

Figure 1. YouTube Comments Screenshot

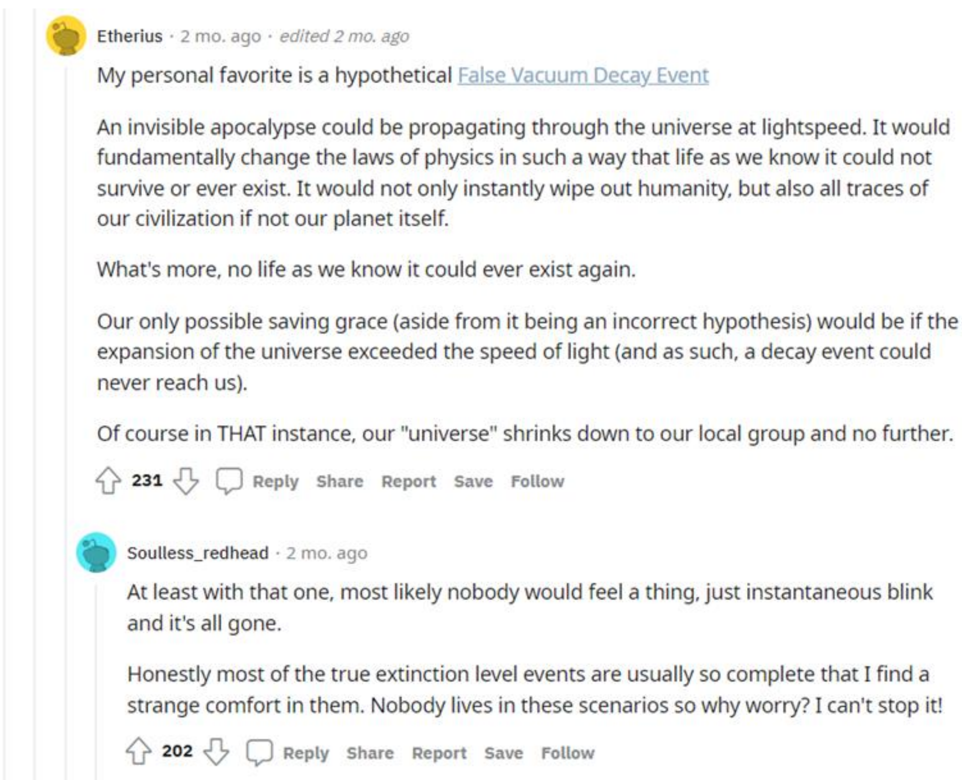

Figure 2. Reddit Comments Screenshot

The number of Reddit comments was used as a proxy for the cognitive dimension of learner engagement. The data on Reddit comments were captured via vidIQ (Purcariu, 2015). VidIQ is a certified YouTube service provider that offers tools for analytics, keyword research, video optimization, comment management, thumbnail generation, bulk updates, data updates, etc. We used the analytics tool through the VidIQ browser extension to get the number of Reddit comments for each sampled video. The browser extension identifies each video by its URL and displays the video metrics with the vidIQ Scorecard. Of the various sections on the vidIQ Scorecard, the Social section shows the video's engagement statistics on YouTube, Facebook, and Reddit (see Figure 3 for an illustration). We obtained the number of Reddit comments from the Social section of the vidIQ Scorecard. Our analysis of YouTube content on another social networking platform is not uncommon for research in YouTube educational videos. For example, YouTube educational videos are examined via student Facebook activities (Tan, 2013).

Figure 3. The Social Section of vidIQ Scorecard (Source: https://vidiq.com/features/scorecard/)

All the sampled YouTube videos flowed continuously, and we did not detect system-paced segmentation. YouTube supports the playback function, which allows the viewers to pause the video at any time to focus on or review specific segments of the video or jump to a different point in time. The statistics of the usage of the playback function were unavailable, so we could not measure this type of learner-paced presentation. Instead, we measured learner control with the number of key moments.

Key moments highlight the chapters of the videos. Google algorithm is now trained to break up a video into key moments. Key moments are displayed in Google Search for the video, showing a timestamp indicating where that moment starts and a label describing what happens at that moment (see Figure 4 for an illustration). Viewers can scan key moments and locate and access specific information from the video. A simple click on any key moments in the video will take the viewer to the specific section. We searched the video link of each sampled video on Google. Out of the 140 sampled videos, 29 did not display key moments. We captured the number of key moments for 111 videos.

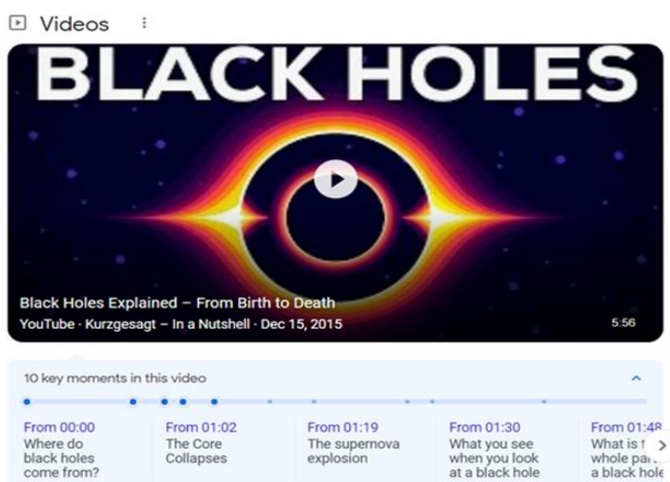

Figure 4. Key Moments Screen Shot

Signaling was measured by the presence of textual and visual signals. Figure 5 and Figure 6 show an example of textual and visual signals, respectively. Each video was checked manually for the presence of signals. A video was coded as 0 if no textual signal was present and 1 otherwise. The same procedure was applied to visual signals as well. The manual tracking of signaling was time-consuming but not especially difficult due to the binary nature of the signaling feature.

Figure 5. Textual Signal Screenshot

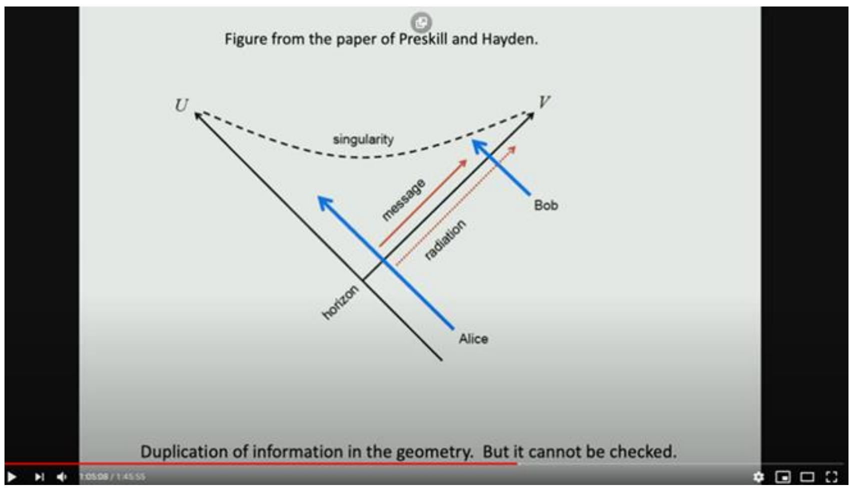

Figure 6. Visual Signal Screenshot

For weeding, we identified one video feature – card - that is not directly relevant to the content covered in the video. YouTube cards are clickable elements that can be placed at any point in the YouTube video. They can be used to promote a channel (channel cards), add links to associated websites, merchandise sites, or crowdfunding pages (link cards), or guide your viewers to any other videos or playlists that are publicly available on YouTube (video or playlist cards). They appear on the top right-hand corner of the screen while the video is streaming. When viewers click on the YouTube card, it will expand to showcase links to other videos, playlists, channels, or external web pages. Up to five cards can be added to each video and they can be used in any video. Figure 7 shows the screenshot of a card that features a video. YouTube cards are not essential to the video content, only functioning to make the video more interactive by adding interactive components. We watched each video to check the use of cards. A video was coded as 1 if no card was present and 0 otherwise.

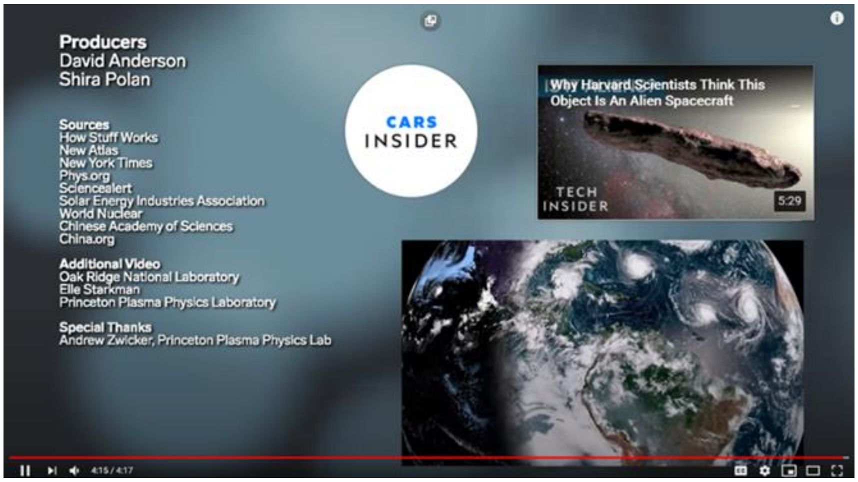

Figure 7. Card Screenshot

We also controlled several variables. Video length (minutes) and narration speed (words per minute) were controlled as prior research shows that they can affect learner engagement with educational videos (Guo et al., 2014). We also controlled video resolution and loudness because they impinge upon the popularity of YouTube educational videos (ten Hove & van der Meij, 2015). We used the default decibel to indicate loudness. The YouTube default decibel is measured by decibels relative to full scale (dBFS), where 0 dBFS is the maximum digital signal level. In addition, given that older videos have had a more extended period to accumulate responses, the number of days since posted (the time interval between the posting date of a video and the day the video was sampled) was also used as a control variable.

Summarized in Table 1 are the variables and their measurements.

Table 1. Variables and Measurements

Variable		Measurement	Data Type	Source
Learner Engagement	behavior engagement	# YouTube views	discrete	YouTube
	emotional engagement	# YouTube votes	discrete	YouTube
	cognitive engagement	# Reddit comments	discrete	vidIQ
SSW	segmenting	# key moments	discrete	manual coding
	textual signaling	presence of headings, summaries, typographical cues, etc.	binary	manual coding
	visual signaling	presence of shapes, dynamic lines, highlight coloring, moving dots, etc.	binary	manual coding
	weeding	presence of YouTube cards	binary	manual coding
Control	days posted	# days between video posted and sampled	discrete	manual coding
	video length	# minutes	continuous	YouTube
	resolution	# pixels	continuous	YouTube
	loudness	default decibel	continuous	YouTube
	narration speed	# words/minute	continuous	vidIQ

5 Methods

5.1 Assumption Check

We ran three multiple linear regression models to test our hypotheses. The number of YouTube views (behavioral engagement), the number of YouTube votes (emotional engagement), and the number of Reddit comments (cognitive engagement) are the dependent variables (DVs). The DVs measured at a ratio scale can take on an infinite number of values. It was reasonable to treat their discrete values as continuous. The normal distribution can be assumed if skewness is between -2 to +2 and kurtosis is between -3 to +3 (Hair et al., 1998). Descriptive analysis showed that the four DVs had skewness and kurtosis values greater than the suggested normality thresholds. Therefore, a log transformation was performed on the three DVs. The summary statistics of the transformed DVs and other variables are presented in Table 2.

Table 2. Descriptives

	Variable	Mean	St.dev	Max	Min	Skewness	Kurtosis
IVs	# key moments	7.89	2.83	16	1	-.68	-.11
	presence of textual signals	.80	.40	1	0	-1.52	.31
	presence of visual signals	.76	.42	1	0	-1.25	-.42
	presence of cards	.75	.48	1	0	.61	-1.65
DVs	# YouTube views (log)	11.32	2.97	16.17	4.62	-.34	-.82
	# YouTube votes (log)	7.22	3.10	12.46	1.09	-.26	-1.03
	# Reddit comments (log)	2.12	1.52	7.93	.69	1.78	3.10
Control	days posted	49.14	29.57	100	1	-.071	-1.26
	video length	28.37	30.01	105.55	2.23	1.15	-.182
	resolution	1193.13	228.91	1280.72	320.24	-2.48	2.87
	default decibel	-2.94	4.47	4.80	-21.70	-1.27	2.42
	narration speed	147.43	37.83	214.40	0	-2.05	3.98

The three models took the number of key moments, presence of textual signals, presence of visual signals, and presence of cards as independent variables (IVs). Table 3 presents the study variables' bivariate correlations (Pearson's r). All correlations between IVs were well below the 0.7 thresholds (Hair et al., 2010), indicating that multicollinearity is not a cause for concern. The variance inflating factors (VIF) of IVs were less than two, indicating the absence of multicollinearity.

Table 3. Correlations

		1	2	3	4	5	6	7	8	9	10	11	12
1	# key moments	-											
2	presence of textual signals	.150	-										
3	presence of visual signals	-.097	.043	-									
4	presence of cards	-.025	-.010	.144	-								
5	# YouTube views (log)	-.019	-.022	.491**	.396**	-							
6	# YouTube votes (log)	-.141	.012	.508**	.466**	.953**	-						
7	# Reddit Comments (log)	.058	.014	.344**	.281**	.667**	.704**	-					
8	days posted	.010	-.151	-.125	-.204*	.117	.005	.009	-				
9	video length	.484**	.094	-.122	-.390**	-.384**	-.475**	-.283**	.071	-			
10	resolution	-.099	.106	.329**	.285**	.216*	.300**	.188*	-.382**	-.354**	-		
11	default decibel	-.175	-.039	.119	-.015	.113	.142	.204*	.119	-.246**	-.090	-	
12	narration speed	.029	.007	.310**	.040	.138	.180	.096	-.121	-.047	.126	.195*	-

Notes:
*. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant at the 0.01 level (2-tailed).

5.2 Hypothesis Testing

Table 4 reports the standardized coefficients of segmenting, signaling, weeding, and control variables from the three models with behavioral, emotional, and cognitive engagement as the DV. Model 1 has the number of YouTube views (behavioral engagement) as DV. H1 is rejected as the number of key moments (segmenting) does not significantly influence the number of YouTube views. The coefficient of visual signaling is 0.465 ($p < .001$), and textual signaling has an insignificant coefficient. Therefore, H2 is partially supported. H3 is supported as card (weeding) has a significant coefficient of -0.274 ($p < 0.01$). Of the control variables, days posted and video length are significant, with the coefficient of 0.158 ($p < 0.05$) and -.292 ($p < 0.01$), respectively.

Table 4. Results from Multiple Linear Regression Models

		Model 1		Model 2		Model 3	
		Behavioral Engagement (# YouTube Views)		Emotional Engagement (# YouTube Votes)		Cognitive Engagement (# Reddit Comments)	
Variable		Beta	p	Beta	p	Beta	p
Segmenting	# key moments	.069	.429	-.070	.393	.251	.015
Signaling	presence of textual signals	.016	.830	.042	.554	.007	.938
	presence of visual signals	.465	.000	.437	.000	.310	.002
Weeding	presence of cards	-.274	.002	-.296	.000	-.151	.127
Control	days posted	.251	.003	.158	.041	.083	.387
	video length (minutes)	-.292	.005	-.358	.000	-.276	.021
	resolution	-.022	.810	.002	.979	.020	.852
	default decibel	-.045	.574	-.019	.807	.137	.148
	narration speed (words/minute)	.008	.915	.036	.627	-.046	.621
Adjusted R2		.462		.562		.269	

Model 2 has the number of YouTube votes (emotional engagement) as DV. The number of key moments (segmenting) does not have a significant coefficient, and thus, H1 is rejected. Visual signaling significantly correlates with the number of votes (beta = 0.437, $p < 0.001$). Textual signaling is insignificant in predicting the number of votes. As such, H2 is partially supported. H3 is supported as card (weeding) has a significant coefficient of -0.296 ($p < 0.001$). Of the control variables, days posted and video length are significant, with coefficoefficients0.251 ($p < 0.01$) and -.358 ($p < 0.001$), respectively.

Model 3 has the number of Reddit comments as DV. H1 is supported with a significant coefficient of the number of key moments (segmenting) (beta = 0.251, $p < 0.05$). Again, the results of signaling are mixed. Visual signaling is a significant predictor (beta =0.310, $p < 0.01$), but textual signaling is not. H3 is not supported as card has an insignificant coefficient. Of the control variables, only video length is significant (beta = -0.276, $p < 0.05$).

5.3 Summary of Results

Table 5 summarizes the results of the hypotheses testing. It is clear from the results that the various SSW features are associated with different dimensions of learner engagement with YouTube educational videos. Segmenting contributes only to cognitive engagement. The presence of visual signaling effectively induces all three dimensions of learner engagement. Surprisingly, the appearance of textual signaling does not affect learner engagement. One possible explanation is that textual signals fail to capture attention and facilitate information organization. Video allows the simultaneous presentation of visual and auditory information and supports multiple channels of information acquisition and processing. All these different external representations compete for limited attention. Visual signals have the visual salience that forcefully directs attention and produces engagement. In comparison, textual signals may not be strong enough to counter the more powerful directions of attention from the video's dynamics. As a result, it may not free up cognitive resources. It is also possible that the freed-up cognitive resources are not sizable enough to affect learner engagement.

Table 5. Summary of Hypotheses Testing Results

	Behavioral engagement (# YouTube views)	Emotional engagement (# YouTube votes)	Cognitive engagement (# Reddit comments)
H1 segmenting - # key moments			+*
H2 signaling – textual signal			
H3 signaling – visual signal	+***	+***	+**
H4 weeding - card	-**	-***	
Note: + significant positive effect; - significant negative effect * significant at the 0.05 level; ** significant at the 0.01 level; *** significant at the 0.001 level			

Weeding is significantly associated with behavioral and emotional engagement, but the direction of association is opposite to what we have hypothesized. This difference may be explained in terms of learners' characteristics. The effects of the coherence principle on learning outcomes (e.g., knowledge retention and transfer) have been empirically verified in controlled laboratory studies with learners with little prior knowledge and limited interest in the domain of instruction (e.g., Mayer et al., 2001; Ibrahim et al., 2012). However, empirical research in authentic online learning settings reveals extra interesting but irrelevant information on educational videos does not result in lower learning achievement as the learners' prior knowledge and interest level may mitigate the effects of the coherence principle (Muller et al., 2008). The learners in our study are primarily physics and astronomy enthusiasts with domain knowledge and intrinsic motivation. YouTube cards offer them opportunities to explore their interest more on the Internet and, as a result, contribute positively to learner engagement.

Contrary to prior research (Guo et al., 2014; ten Hove & van der Meij, 2015), our study does not find significant effects of video resolution, loudness, and narration speed on learner engagement. This result may be attributed to the sampled videos' reasonable resolution, default decibel, and narration speed, which do not distract learners. Our study confirms the earlier finding (Guo et al., 2014) that learners are more engaged with shorter videos. Finally, though it exhibits significant associations with behavioral and emotional engagement, the number of days since posted does not affect cognitive engagement.

6 Discussion

6.1 Research Contributions

Our study improves the understanding of learner engagement and multimedia learning principles. It makes several contributions to the literature on learner engagement with educational videos. First, our study expands the perspectives of video engagement in learning. Extant research approaches learner engagement with educational videos mainly from a behavioral standpoint. For example, engagement is indicated by the number of videos watched (Albó et al., 2019) and the time students spend watching videos (Guo et al., 2014). Our study goes beyond the one-dimension view and captures emotional and cognitive engagement along with behavioral engagement. The multiple dimension perspective is valuable for future research in engagement with educational videos from different platforms in various educational contexts.

Second, our study furthers the measurement of learner engagement in online settings. Engagement in learning on social media is often measured by surveying learners' learning experiences (e.g., Hong & Gardner, 2019) or analyzing online discussions and postings (e.g., Xu et al., 2020). We assess engagement via social media analytics. As learners' digital tracks (e.g., viewing and discussion behaviors) on social media are captured and stored, our analytics approach can be adapted to measure learner engagement on other social media platforms (e.g., Facebook and WeChat).

Third, our study advances the knowledge of the factors that affect engagement with educational videos in informal online learning. Prior research of educational videos in the online video-based learning environment reports that student factors (e.g., personal interest or curiosity), instructor factors (e.g., instructor accessibility and passion), course design factors (e.g., opportunities for peer interaction) contribute to learner engagement in course setting (Breslow et al., 2013; Hew, 2016; Hattingh, 2017). Our study demonstrates the impacts of video presentation on engagement with videos in informal learning. It provides empirical evidence that segmenting can enhance cognitive engagement. It shows that when both textual and visual signals are present in videos, visual signals are more effective than textual signals in engendering behavioral, emotional, and cognitive engagement. Our study also reveals that weeding contributes to learner engagement in informal learning, thus inviting further investigations on the effects of the coherence principle of multimedia design.

Our study also enriches the literature on multimedia learning by applying multimedia learning principles to learner engagement. Our findings may help understand why some multimedia presentation features benefit learning outcomes whereas others are not. For example, the inability of textual signals to improve learner engagement may explain why some studies (e.g., Mautone & Mayer, 2001) have found no effects of textual signals on knowledge retention and transfer. Engagement could help reconcile the mixed effects of signaling on learning outcomes. A better understanding of engagement can produce more comprehensive knowledge of the impact of SSW on learning outcomes.

In addition, our study extends the research on multimedia learning principles from laboratory settings to the actual learning environment. The findings from laboratory experiments where researchers control how students learn may not directly apply to the more learner-controlled context. For instance, signaling effects are not verified when the treatment is applied to a classroom-based environment (Tabbers et al., 2004). Similarly, the effects of weeding on learning outcomes cannot be generalized to learning in the classroom setting (Muller et al., 2008). This study goes beyond formal learning in the classroom. Our examination of SSW in informal online learning sheds new light on their effects on learning and directs more research attention to the generalizability of multimedia learning principles to the authentic learning environment.

6.2 Practical Implications

This study encourages YouTube video producers to make more rigorous decisions based on data rather than just intuitions. The findings provide some guidelines for designing engaging YouTube educational videos. First, the use of signaling is highly recommended. In designing signals, preference should be given to visual signals over textual signals. Textual signals are usually easier to implement but are less effective than visual signals in engaging learners. In addition, we recommend incorporating cards into YouTube educational videos to increase viewer interaction with the video for more learner engagement. Adding timestamps in YouTube videos can also facilitate viewers navigating the video content. Finally, video length matters. If possible, keep the video short to engage viewers.

Though based on our findings from informal learning on social media, the design guidelines mentioned above are also relevant to other learning topic areas and contexts. Therefore, our research is valuable to video producers, platform providers, and instructors towards high-quality production and organization of educational video content. In addition, the practical significance of our study goes beyond learning, given the increasing popularity of videos in everyday life and work. What we have found in educational video engagement can benefit practitioners who want to use videos to engage with their users and customers on websites and social media.

6.3 Limitations and Future Research

This research has limitations. First, we did not include all the possible SSW features. We were not able to analyze the YouTube feature of playback. Future research can collect data on how this segmenting feature is used and explore its effects on engagement. Future research can also look into other signals like anchors. Anchors are the few key data elements that learners use to construct the explanatory structure (i.e., a frame) that guides collecting data and organizing information (Klein et al., 2007). They have been empirically proven helpful in locating pertinent patterns and connections in the data stream (Pontis & Blandford, 2016). It would be interesting to see how anchors affect engagement.

Second, our findings may not be generalized to other online video content. The analyzed videos are on physics and astronomy. They relate to declarative knowledge. Declarative knowledge, also characterized as “knowing that”, includes factual and conceptual information. YouTube educational videos can also be tutorials (e.g., a problem-solving walkthrough) related to procedural knowledge. Procedural learning, also known as “knowing how”, involves memorizing an organized and discrete sequence of behaviors. The findings from this study may not apply or apply well to tutorials about how-tos. Future research can sample tutorial videos to examine the differences in the effects of video features on learner engagement.

The audience in the sampled videos also limits the generalizability of our study. The viewers of the educational videos in this study were general knowledge seekers who learned physics and astronomy voluntarily outside the formal classroom. YouTube is also a learning platform for other science learners, such as school-age children in informal learning (Dyosi & Hattingh, 2018) and college students in formal learning (Clifton & Mann, 2011). This study's video features that engage the viewers interested in science may not engage school children and other learner populations. Also, the videos targeted at children and other learners can differ significantly in features. So, future research can validate our findings in different groups of learners.

7 Conclusion

The last decade has witnessed a tremendous expansion of YouTube educational videos in various learning settings. This study addresses the lack of research on learner engagement with YouTube educational videos in informal learning. Drawing upon multimedia design research, we propose and test the hypotheses regarding the effects of SSW on learner engagement. The findings of this study enrich the understanding of engagement with YouTube educational videos from the cognitive perspective. They can also help inform instructors and video producers about making online videos more engaging to support learning. It is hoped that this research provides a good starting point for further research on engagement with educational videos on social media.

Acknowledgments

I thank Dr. Michael J. Pritchard for collecting data and Dr. Songxin Tan for helping with data collection and analysis.

References

Albó, L., Hernández-Leo, D., & Oliver, V. M. (2019). Smartphones or laptops in the collaborative classroom? A study of video-based learning in higher education. *Behaviour & Information Technology*, *38*(6), 637-649.

Allgaier, J. (2019). Science and environmental communication on YouTube: Strategically distorted communications in online videos on climate change and climate engineering. *Frontier in Communication*, 4, 36.

Andel, S. A., de Vreede, T., Spector, P. E., Padmanabhan, B., Singh, V. K., & de Vreede, G. (2020). Do social features help in video-centric online learning platforms? A social presence perspective. *Computers in Human Behavior, 113*, 106505.

Appleton, J. J., Christenson, S. L., Kim, D., & Reschly, A. L. (2006). Measuring cognitive and psychological engagement: Validation of the student engagement instrument. *Journal of School Psychology, 44,* 427-445.

Azer, S. A., Al Grain, H. A., Al Khelaif, R. A., & Al Eshaiw, S.M. (2013). Evaluation of the educational value of YouTube videos about physical examination of the cardiovascular and respiratory systems. *Journal of Medical Internet Research, 15*(11), e.241.

Barrouillet, P., & Camos, V. (2007). The time-based resource sharing model of working memory. In N. Osaka, R. H. Logie, & M. D'Esposito (Eds.), *The Cognitive Neuroscience of Working Memory* (pp. 59-80). Oxford, England: Oxford University Press.

Bétrancourta, M. & Benetosa, K. (2018). Why and when does instructional video facilitate learning? A commentary to the special issue "developments and trends in learning with instructional video." *Computers in Human Behavior, 89*, 471-475.

Biard, N., Cojean, S., & Jamet, E. (2018). Effects of segmentation and pacing on procedural learning by video. *Computers in Human Behavior, 89*, 411-417.

Boucheix, J., Gauthier, P., Fontaine, J., & Jaffeux, S. (2018). Mixed camera viewpoints improve learning medical hand procedure from video in nurse training? *Computers in Human Behavior, 89*, 418-429.

Breslow, L., Pritchard, D. E., DeBoer, J., Stump, G. S., Ho, A. D. & Seaton, D. T. (2013). Studying learning in the worldwide classroom. Research into edX's first MOOC. *Research & Practice in Assessment, 8,* 13-25.

Bulger, M. E., Mayer, R. E., Almeroth, K. C., & Blau, S. D. (2008). Measuring learner engagement in computer-equipped college classrooms. *Journal of Educational Multimedia and Hypermedia, 17*(2), 129-143.

Buzzetto-More, N. (2015). Student attitudes towards the integration of YouTube in online, hybrid, and web-assisted courses: An examination of the impact of course modality on perception. *MERLOT Journal of Online Learning and Teaching, 11*(1), 55-73.

Chau, C. 2010. YouTube as a participatory culture. New Directions for Youth Development, 128, 65-74.

Chi, M. T. H., & Wylie, R. (2014). The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist, 49*(4), 219-243.

Clifton, A., & Mann, C. (2011). Can YouTube enhance student nurse learning? *Nurse Education Today*, 31(4), 311-313.

Cranwell, J., Murray, R., Lewis, S., Leonardi-Bee, J., Dockrell, M., & Britton, J. (2015). Adolescents' exposure to tobacco and alcohol content in YouTube music videos. *Addiction, 110*(4), 703-711.

de Koning, B., Tabbers, H., Rikers, R., & Paas, F. (2009). Towards a framework for attention cueing in instructional animations: Guidelines for research and design. *Educational Psychology Review, 21*, 113-140.

Dubovi, L., & Tabak, I. (2020). An empirical analysis of knowledge co-construction in YouTube comments. *Computer & Education, 156*, 103939.

Dyosi, N., & Hattingh, M. (2018). Understanding the extent of and factors involved in the use of YouTube as an informal learning tool by 11- to 13-year-old children. *Innovative Technologies and Learning, 11003*, 351-361.

Fredricks, J. A., Blumenfeld, P. C. & Paris, A. (2004). School engagement: Potential of the concept: State of the evidence. *Review of Educational Research, 74*, 59-119.

Fredricks, J. A., & McColskey, W. (2012). The measurement of student engagement: A comparative analysis of various methods and student self-report instruments. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of Research on Student Engagement* (pp. 763–782). Springer Science + Business Media.

Furlong, M. J., Whipple, A. D., St. Jean, G., Simental, J., Soliz, A., & Punthuna, S. (2003). Multiple contexts of school engagement: Moving toward a unifying framework for educational research and practice. *The California School Psychologist, 8*(1), 99-113.

Greene, B. A. (2015). Measuring cognitive engagement with self-report scales: Reflections from over 20 years of research. *Educational Psychologist, 50*(1), 14-30.

Greenhow, C., & Lewin, C. (2016). Social media and education: Reconceptualizing the boundaries of formal and informal learning. *Learning, Media and Technology*, 41(1), 6-30.

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), New York: ACM Press.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L., (1998). *Multivariate Data Analysis (5th ed.)*. Upper Saddle River, NJ, USA: Prentice-Hall.

Hattingh, M. (2017). A preliminary investigation of the appropriateness of YouTube as an informal learning platform for pre-teens. *Advances in Web-Based Learning - ICWL 2017, 10473*,101-110.

Haythornthwaite, C., Kumar, P., Gruzd, A., Gilbert, S., Esteve del Valle, M., & Paulin, D. (2018). Learning in the wild: coding for learning and practice on Reddit. *Learning, Media and Technology, 43*(3), 219-235.

Hershcovits, H., Vilenchik, D., & Gal, K. (2020). Modeling engagement in self-directed learning systems using principal component analysis. *IEEE Transactions on Learning Technologies, 13*(1), 164-171.

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.

Hong, J.-C., Hwang, M.-Y., Szeto, E., Tsai, C.-R., Kuo, Y.-C. & Hsu, W.-Y. (2016), Internet cognitive failure relevant to self-efficacy, learning interest, and satisfaction with social media learning. *Computers in Human Behavior, 55*(A), 214-222.

Hong, Y., & Gardner, L. (2019). Undergraduates' perception and engagement in Facebook learning groups. *British Journal of Educational Technology, 50*(4), 1831-1845.

Huang, Y., Jasin, S., & Manchanda, P. (2019). "Level up": Leveraging skill and engagement to maximize player gameplay in online video games. *Information Systems Research, 30*(3), 927-947.

Huk, T. & Steinke, M. (2007). Learning cell biology with close-up views or connecting lines: Evidence for the structure mapping effect. *Computers in Human Behavior, 23*, 1089-1104.

Ibrahim, M., Antonenko, P. D., Greenwood, C. M., & Wheeler, D. (2012). Effects of segmenting, signaling, and weeding on learning from educational video. *Learning, Media and Technology, 37*(3), 220-235.

Jacques, R., Precce, J., & Carey, J. T. (1995). Engagement as a design concept for hypermedia. *Canadian Journal of Educational Communications, 24*(1), 49-59.

Jamet, E. (2014). An eye-tracking study of cueing effects in multimedia learning. *Computers in Human Behavior, 32*, 47-53.

Jensen, M., Lee, Y., Cameron W. Piercy, C. W., Dunbar, N. E., Elizondo, J., Bessarabova, E., Twyman, N. W., Burgoon, J. K., Valacich, J. S., Adame, B., Miller, C., & Wilson, S. (2016). Exploring failure and engagement in a complex digital training game: A multi-method examination. *AIS Transactions on Human-Computer Interaction, 8*(1), 1-20.

Juhasz, A. (2011). *Learning from YouTube [Video Book].* MIT Press (BK).

Khan, M. L. (2017). Social media engagement: What motivates user participation and consumption on YouTube? *Computers in Human Behavior, 66*, 236-247.

Klein, G., Phillips, J. K., Rall, E. L., & Peluso, D. A. (2007). A data-frame theory of sensemaking. In R. R. Hoffman (Ed.), *Expertise out of context: Proceedings of the Sixth International Conference on Naturalistic Decision Making* (p. 113-155). Lawrence Erlbaum Associates Publishers.

Kohler, S., & Dietrich, T. C. (2021). Potentials and limitations of educational videos on YouTube for science communication. *Frontiers in Communication*, 6, 581302.

Lange, P. G. (2019). Informal learning on YouTube. In The International Encyclopedia of Media Literacy (pp. 1–11). John Wiley & Sons, Inc.

Leahy, W., & Sweller, J. (2005). Interactions among the imagination, expertise reversal, and element interactivity effects. *Journal of Experimental Psychology: Applied, 11*, 266-276.

Lee, D. Y., & Lehto, M. R. (2013). User acceptance of YouTube for procedural learning: An extension of the Technology Acceptance Model. *Computers & Education, 61*, 193-208.

Lee, C. S., Osop, H., Goh, D. H., & Kelni, G. (2017). Making sense of comments on YouTube educational videos: a self-directed learning perspective. *Online Information Review, 41*(5), 611-625.

Liikkanen, L. A., & Salovaara, A. (2015). Music on YouTube: User engagement with traditional, user-appropriated and derivative videos. *Computers in Human Behavior, 50*, 108-124.

López-Nores, M., Blanco-Fernández, Y., Pazos-Arias, J. J. (2013). Cloud-Based Personalization of New Advertising and e-Commerce Models for Video Consumption. *The Computer Journal, 56*(5), 573–592.

Lorch, R. F. (1989). Text-signaling devices and their effects on reading and memory processes. *Educational Psychology Review, 1*, 209-234.

Lorch, R. F. & Lorch, E. P. (1996). Effects of headings on text recall and summarization. *Contemporary Educational Psychology, 21*, 261-278.

Lowe, R. K. & Boucheix, J. M. (2007). Eye-tracking as a basis for animation design. In: *12th Bi-annual Meeting of the European Association of Research on Learning and Instruction*, Budapest, August 28 - September 1.

MacInnis, D. J., Moorman, C., & Jaworski, B. J. (1991). Enhancing and measuring consumers' motivation, opportunity, and ability to process brand information from ads. *Journal of Marketing, 55*(4), 32-53.

Madden, A., Ruthven, I., & McMenemy, D. (2013). A classification scheme for content analyses of YouTube video comments. *Journal of Documentation, 69*(5), 693-714.

Maddox, S. J., & Prinz, R. J. (2003). School bonding in children and adolescents: Conceptualization, assessment, and associated variables. *Clinical Child and Family Psychology Review, 6*(1), 31-49.

Martin, F., Sun, T., & Westine, C. D. (2020). A systematic review of research on online teaching and learning from 2009 to 2018. *Computers & Education, 159*, 104009.

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

Mayer, R.E. (2003). The promise of multimedia learning: Using the same instructional design methods across different media. *Learning and Instruction, 13*(2), 125-39.

Mayer, R. E. (Ed.) (2009). *The Cambridge handbook of multimedia learning (2nd ed.)*. Cambridge, U.K.: Cambridge University Press.

Mayer, R. E., Heiser, J. & Lonn, S. (2001). Cognitive constraints on multimedia learning: When presenting more material results in less understanding. *Journal of Educational Psychology, 93*(1), 187-198.

Mayer, R. E., & Moreno, R. (2003). Nine ways to reduce cognitive load in multimedia learning. *Educational Psychologist, 38*(1), 43-52.

McIntyre, D. J., Copenhaver, R. W., Byrd, D. M., & Norris, W. R. (1983). A study of engaged student behavior within classroom activities during mathematics class. *The Journal of Educational Research, 77*(1), 55-59.

Mello, L. V. (2016). Fostering postgraduate student engagement: online resources supporting self-directed learning in a diverse cohort. *Research in Learning Technology, 24*, 1-16.

Moreno, R., & Mayer, R. (2007). Interactive multimodal learning environments. *Educational Psychology Review, 19*(3), 309-326.

Muller, D. A., Lee, K. J., & Sharma, M. D. (2008). Coherence or interest: Which is most important in online multimedia learning? *Australasian Journal of Educational Technology, 24*(2), 211-221.

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.

Oh, C., Roumani, Y., Nwankpa, J. K., & Hu, H. (2017). Beyond likes and tweets: Consumer engagement behavior and movie box office in social media. *Information & Management, 54*, 25-37.

Ozcelik, E., Arslan-Ari, I., & Cagiltay, K. (2010). Why does signaling enhance multimedia learning? Evidence from eye movements. *Computers in Human Behavior, 26*, 110-117.

Pontis, S., & Blandford, A. (2016). Understanding "Influence": An empirical test of the data-frame theory of sensemaking. *Journal of the Association for Information Science and Technology, 67*(4), 841-858.

Purcariu, M. (2015). Tagging behavior in tutorials on YouTube: A case study approach. *Studia Universitatis Babes-Bolyai-Ephemerides, 60*(2), 61-74.

Rieder, B., Matamoros-Fernández, A., & Coromina, Ò. (2018). From ranking algorithms to 'ranking cultures' Investigating the modulation of visibility in YouTube search results. *Convergence, 24*(1), 50-68.

Sablić, M., Mirosavljević, A., & Škugor, A. (2021). Video-based learning (VBL)—past, present and future: An overview of the research published from 2008 to 2019. *Technology, Knowledge and Learning, 26*(4), 1061-1077.

Scheiter, K. & Gerjets, P. (2007). Learner Control in Hypermedia Environments. *Educational Psychology Review, 19*, 285-307.

Seo, K., Dodson, S., Harandi, N. M., Roberson, N., Fels, S., & Roll, I. (2021). Active learning with online video: The impact of learning context on engagement. *Computers & Education, 165*, 104132.

Shah, P., Mayer, R., & Hegarty, M. (1999). Graphs as aids to knowledge construction. *Journal of Educational Psychology, 91*, 690-702.

Shao, Z. & Chen, K. (2021). Understanding individuals' engagement and continuance intention of MOOCs: the effect of interactivity and the role of gender. *Internet Research, 31*(4), 1262-1289.

Shoufan, A., & Mohamed, F. (2022). YouTube and education: A scoping review. *IEEE Access*, 10, 125576 - 125599.

Skinner, E. A., & Pitzer, J. R. (2012). Developmental dynamics of student engagement, coping, and everyday resilience. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of Research on Student Engagement* (pp. 21-44). New York, NY: Springer.

Sweller, J. (1988). Cognitive load during problem-solving: Effects on learning. *Cognitive Science, 12*, 257-285.

Sweller, J., Van Merrienboer, J., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*(3), 251-296.

Tabbers, H., Martens, R., & van Merriënboer, J. (2004). Multimedia instructions and cognitive load theory: effects of modality and cueing. *British Journal of Educational Psychology, 74*, 71-81.

Tafesse, W. (2020). YouTube marketing: how marketers' video optimization practices influence video views. *Internet Research, 30* (6), 1689-1707.

Tan, E. (2013). Informal learning on YouTube: Exploring digital literacy in independent online learning. *Learning, Media and Technology, 38*(4), 463-477.

ten Hove, P., & van der Meij, H. (2015). Like it or not: What characterizes YouTube's more popular instructional videos? *Technical Communication, 62* (1), 48-62.

Tversky, B., Zacks, J., Lee, P. U., et al. (2000). Lines, blobs, crosses, and arrows: Diagrammatic communication with schematic figures. In M Anderson, P. Cheng & V. Haarslev (Eds.), *Theory and Application of Diagrams*, (pp. 221-230). Berlin: Springer.

Tversky, B., Morrison, J., & Bétrancourt, M. (2002). Animation: Can it facilitate? *International Journal of Human-Computer Studies, 57*, 247-262.

van Wermeskerken, M., Ravensbergen, S., & van Gog, T. (2018). Effects of instructor presence in video modeling examples on attention and learning. *Computers in Human Behavior, 89*, 430-438.

Vizcaíno-Verdú, A., Contreras-Pulido, P., & Guzmán-Franco, M. D. (2019). Reading and informal learning trends on YouTube: The booktuber. *Comunicar - Media Education Research Journal*, 27(1), 93-101.

Voelkl, K. E. (1997). Identification with school. *American Journal of Education, 105*, 204319.

Wang, H., & Chen, C. W. (2020). Learning English from YouTubers: English L2 learners' self-regulated language learning on YouTube. *Innovation in Language Learning and Teaching, 14*(4), 333-346.

Wang, X., Lin, L., Han, M., & Spector, J. M. (2020). Impacts of cues on learning: Using eye-tracking technologies to examine the functions and designs of added cues in short instructional videos. *Computers in Human Behavior, 107*, 106279.

Wouters, P., Tabbers, H. K., & Paas, F. (2007). Interactivity in video-based models. *Educational Psychology Review, 19*(3), 327e342.

Xu, B., Chen, N., & Chen, G. (2020). Effects of teacher role on student engagement in WeChat-Based online discussion learning. *Computers & Education, 157*, 103956.

Xu, J., Fedorowicz, J., & Williams, C. B. (2019). Effects of symbol sets and needs gratifications on audience engagement: Contextualizing police social media communication. *Journal of the Association for Information Systems, 20*(5), 536-569.

Zhou, Q., Lee, C. S., Sin, S. C. J., Lin, S., Hu, H., & Ismail, M. F. F. B. (2020). Understanding the use of YouTube as a learning resource: a social cognitive perspective. *Aslib Journal of Information Management,* (3), 339-359.

Zhu, E. (2006). Interaction and cognitive engagement: An analysis of four asynchronous online discussions. *Instructional Science, 34*, 451-480.

About the Authors

Zixing Shen received her Ph.D. in Management from Case Western Reserve University, USA. She is currently an Associate Professor of Information Systems at the College of Business, New Mexico State University, USA. Her research has been published in journals such as the *European Journal of Information Systems*, the *Journal of the Association for Information Science and Technology*, and *Communications of the Association for Information Systems*, among others. Her current research focuses on engagement and sense-making on social media and AI in healthcare.