

Active Viewing: A Study of Video Highlighting in the Classroom

Samuel Dodson
University of British Columbia
dodsons@mail.ubc.ca

Ido Roll
University of British Columbia
ido.roll@ubc.ca

Matthew Fong
University of British Columbia
mfong@ece.ubc.ca

Dongwook Yoon
University of British Columbia
yoon@cs.ubc.ca

Negar M. Harandi
University of British Columbia
negarm@ece.ubc.ca

Sidney Fels
University of British Columbia
ssfels@ece.ubc.ca

ABSTRACT

Video is an increasingly popular medium for education. Motivated by the problem of video as a one-way medium, this paper investigates the ways in which learners' active interaction with video materials contributes to active learning. In this study, we examine *active viewing* behaviors, specifically seeking and highlighting within videos, which may suggest greater levels of participation and learning. We deployed a system designed for active viewing to an undergraduate class for a semester. The analysis of online activity traces and interview data provided novel findings on video highlighting behavior in educational contexts.

CCS CONCEPTS

• **Information systems** → **Video search**; • **Applied computing** → **Interactive learning environments**; **Annotation**;

KEYWORDS

active viewing, video search, learning, highlighting

ACM Reference Format:

Samuel Dodson, Ido Roll, Matthew Fong, Dongwook Yoon, Negar M. Harandi, and Sidney Fels. 2018. Active Viewing: A Study of Video Highlighting in the Classroom. In *CHIIR '18: 2018 Conference on Human Information Interaction & Retrieval, March 11–15, 2018, New Brunswick, NJ, USA*. ACM, New York, NY, USA, 4 pages. <https://doi.org/10.1145/3176349.3176889>

1 INTRODUCTION

Video is an increasingly popular medium for education. Video offers many advantages, such as the opportunity for self-paced learning. Given the benefits of video based learning, many blended, flipped, or hybrid courses employ video lectures. Video is not perfect, however. Watching video often creates a passive learning experience for students. Motivated by the problem of video as a one-way medium, this paper investigates the ways in which learner's interaction with video materials contribute to active learning.

Theories of active learning [e.g., 4] argue that effective learning takes place while students engage and interact with the educational

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than the author(s) must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

CHIIR '18, March 11–15, 2018, New Brunswick, NJ, USA

© 2018 Copyright held by the owner/author(s). Publication rights licensed to Association for Computing Machinery.

ACM ISBN 978-1-4503-4925-3/18/03...\$15.00

<https://doi.org/10.1145/3176349.3176889>

materials, such as textbooks and video lectures. When studying reading, nonlinear navigating and annotating can be used as indicators of active reading [1]. For evaluating video viewing, we examine *active viewing* behaviors, specifically seeking and highlighting within videos, which may suggest greater levels of participation and learning. By “seeking” we mean moving to a new part of the video. Previous studies have shown that video navigation behaviors are associated with active information seeking and engaged learning [5, 9]. Nonetheless, it is largely unknown how video highlighting constitutes an active viewing experience.

Recently, Fong et al. [6] presented ViDeX, a system that supports video highlighting. Although the design of the system was assessed, the ways in which active viewing features are used was not examined. Consequently, we are left with the following research questions: (i) How are video seeking activities associated with highlighting behaviors? (ii) How do active viewing behaviors contribute to learning? (iii) What are the appropriate affordances of video interfaces for engagement?

To answer these questions, we deployed ViDeX to an undergraduate class for a semester. The analysis of online activity traces and interview data provided novel findings on video highlighting behavior in educational contexts. For instance, the type of content (i.e., textual or visual) was an important determinant of highlighting behavior. The users of the highlighting features relied more often on semantic cues of the Transcript than visual cues of the Filmstrip.

2 RELEVANT WORK

For online learning to be effective, students must have access to tools that facilitate active learning. For example, in the video medium, this could be done by supporting aggregation and visualization of viewer interactions to facilitate navigation [7, 9]. In other mediums, a common way that learners interact with information is by commenting, highlighting, tagging, or otherwise annotating. Annotation is a type of information interaction that has been found to have a positive relationship with learning outcomes, such as recall, comprehension, and engagement [2, 3, 15]. Consequently, many digital reading environments have been designed to support active reading by providing annotation [e.g., 11, 17]. Tashman and Edwards [16] define active reading as “reading activities that involve more interaction with the reading media than simple sequential advancement through the text.”

While annotation is a well-established information practice in text environments, it is less common when viewing video. This may be because there are few video viewing systems that enable this functionality, although there are exceptions [e.g., 10]. The lack

of systems that allow for annotation may be due to the difficulty designing systems that enable video annotation in a user-friendly way. Hayles [8] discusses the challenge of translating information practices for learning activities from one medium to another, arguing that some functionality of the particular practice is both lost and gained in the process. As a result, the design of video based learning systems must not simply borrow practices from other media without considering the differences between media.

While active reading is specific to text, we can learn from the developments of digital reading environments to develop new approaches to encourage active viewing. Highlighting is one of the most popular types of annotation in text environments [13]. This is, perhaps, because of the low cognitive load required to emphasize content, which does not distract from the individual's primary task. The popularity of highlighting may also be due to learners' assumption that highlighting has positive effects on their reading processes and outcomes. By supporting highlighting in video, we can attempt to translate a common text based practice and test its effectiveness in video based learning environments.

3 ViDeX

ViDeX is a system designed for active viewing through content interaction and personalization. Encouraging information interactions similar to the more familiar practice of annotating text, ViDeX aims to make interacting with video intuitive. ViDeX currently supports the ability to highlight intervals of video and text. We created ViDeX to explore the effectiveness of highlighting video. We recognize that annotation is an important way in which many individuals learn and interact with information in other media, and suspect that viewers may benefit from similar tools. While commenting exists in many video viewing systems, highlighting is a unique feature.



Figure 1: A screenshot of ViDeX, displaying the Transcript (far left) and Filmstrip (bottom).

The ViDeX interface is comprised of three elements: the Player, the Filmstrip, and the Transcript. The Filmstrip provides users with a visual summary of the video, allowing users to skim the video content. The Filmstrip itself is a sequence of thumbnails, where each thumbnail represents a segment of the video. By placing the cursor across the width of each thumbnail, the picture in the thumbnail changes according to the linear mapping between the width of the thumbnail and the part of the video it represents. This allows the user to preview the visual content that the video represents.

To highlight, users must either drag and select across the Filmstrip or Transcript to first select a span of content, and subsequently select a highlight color using the toolbar above the Player. Highlighting creates a strip of color along the top edge of the Filmstrip in addition to the corresponding text in the Transcript.

The Transcript brings the user a textual representation of the video. Each word in the Transcript is associated with a time in the video. Consequently, selections and highlights in either the Filmstrip or the Transcript are synchronized. Any actions made in one widget transfers to the other.

4 METHODS

We evaluated engagement with ViDeX using a mixed-method approach, employing log data analysis and interviews, over the course of a semester in an undergraduate chemistry class. Quantitative log data were used to identify patterns of information interaction. Qualitative interview data were used to better understand and explain these patterns. Students enrolled in the class were provided with videos related to lectures and labs. The videos ranged in length from 3.93 to 12.92 minutes ($M=8.73$ minutes, $SD=2.94$ minutes). One of the authors visited the class at the beginning of the semester to introduce students to ViDeX and invite them to participate in the study by using the system to watch videos assigned by their instructor. Using ViDeX was not a class requirement nor was watching the videos. The videos were also available on YouTube, so students could watch the content without participating in the study. Participants were not compensated for using ViDeX.

All relevant system events were logged, including video events (such as loading, playing, and seeking) and highlighting events (such as color, source, and span of each highlight). All data were anonymously logged. To allow students time to familiarize themselves with the system and its novel features without affecting our analysis of typical usage, we analyzed data starting a month after the start of the semester. By the end of the semester, 6,995 system events from 28 participants were logged. A limitation of the study was the relatively small sample size. The log data were analyzed, using R, with an exploratory data analysis approach. We treated system events as independent in our data analysis.

In addition to logging system events, we also collected qualitative data through interviews with students from the class at the end of the semester. One of the authors recruited six participants for one-on-one and small group semi-structured interviews of about ten minutes each. The interviews explored participants' typical use of ViDeX and their learning practices with and without ViDeX. Interviewees were paid a \$15.00 honorarium for their time.

5 RESULTS & DISCUSSION

In this section, we explore students' viewing and highlighting behaviors. We supplement the quantitative analysis with feedback we received from interviewees about their typical use of ViDeX throughout the semester.

5.1 Highlighting

Of the 28 participants whose interactions with the system and content were logged, five (18%) highlighted at least once. In total, these

five participants created 51 highlights. All but one of the participants highlighted in more than one login session. The highlights varied in color, length, and location within the videos. The median length of highlight spans was 6.23 seconds ($M=11.55$, $SD=13.15$), and ranged from 0.16 to 54.69 seconds. Most highlights ($n=29$, 56.86%) were less than ten seconds and only 6 (11.76%) were longer than thirty seconds. Unlike previous studies of text highlighting [14], we did not find that some participants engaged in heavy highlighting of the content (so called “happy highlighters”).

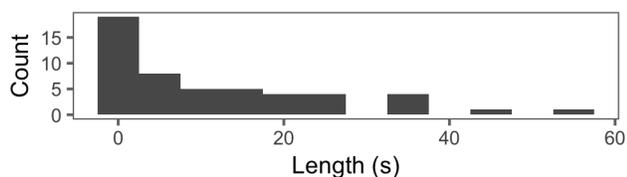


Figure 2: Highlight length in seconds showing that 56.86% of highlights were less than ten seconds and only 11.76% were longer than thirty seconds.

To highlight, the user selects content by clicking and dragging content via either the Filmstrip or the Transcript then chooses a highlight color using the toolbar above the Player. Few selections became highlights. While students made 942 selections, only 51 highlights were created. One reason for this may be the interface itself. The select-and-highlight sequence may not have been very intuitive or visible to most users. At the end of the semester, four of the six interviewees explained they did not know they could highlight using ViDeX, which is surprising given that ViDeX was demoed at the beginning of the semester. Highlighting may have been used more if students had known about it, especially considering all interviewees reported taking notes. The four interviewees that did not highlight explained that they either took notes using a word processor, with a paper notebook, or by taking screenshots.

Of the highlights that were created, many appeared to be used as bookmarks. 56.86% ($n=29$) of the highlights were less than ten seconds. This suggests that students were placing bookmarks in the video more than selecting spans of content, such as highlighting full sentences or paragraphs in the Transcript. One of the interviewees explained that he used highlights as bookmarks for spans of video to watch later. He would replay the bookmarked video to learn more about a keyword. Using highlights exclusively as bookmarks is somewhat surprising, since highlights have been found to have additional functions in text [e.g., 12].

Given that students mainly highlighted content using the Transcript, why did they use the highlights as bookmarks and not for other types of functions that are common when annotating text? This tendency may suggest that students were indeed strategic about their behavior. Even when highlighting text, students realized that this activity is done in the context of video. A bookmark highlight is sufficient to locate the area of interest. By clicking and jumping to a highlight, they were able to bring the playhead to the area of interests, and play the relevant segment of video. The majority ($n=26$, 50.98% of highlights were also jumped to by clicking within the highlighted span of content. The actual number of

highlights that were jumped to is likely higher, because students may have clicked just before the highlighted span, especially for short highlights that are difficult to click. Indeed, 86.27% highlighted spans were viewed again.

Notably, while most selections were made using the Filmstrip ($n=764$, 81.00%), all but two highlights were created with the Transcript. Students’ preference to highlight using the Transcript may be explained in a similar way: highlighting was used mainly to bookmark keywords, rather than emphasize the full extent of visual content of interest. As highlighting was based on textual information, the Transcript was the natural place for highlighting.

Another important aspect of highlights is their location relative to the playhead position, which represents the current playing frame of video content. Approximately two-fifths of the highlights were created within 10 seconds of the playhead (see Figure 3). Highlights were created behind ($n=8$, 15.69%) and ahead ($n=23$, 45.10%) of the playhead. The behavior of highlighting close to and far from the playhead was consistent across participants. Highlighting ahead of the playhead suggests that students may have been skimming the text to identify important content to highlight and guide their future information interaction. This pattern, again, is consistent with the role of highlights as bookmarks described above.

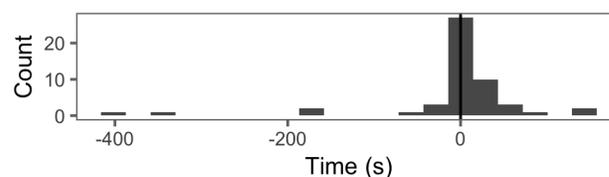


Figure 3: Highlight relative to playhead position. 39.22% of highlights were created within 10 seconds of the playhead.

How information is communicated through video may affect highlighting and seeking. An interviewee explained that he would have highlighted more frequently if the videos had more “show how” rather than “tell how” information. When communicating embodied knowledge, the Transcript would likely contain less textual and more visual information, requiring the student to watch the video to fully understand the context of what is being communicated.

5.2 Seeking

Students used the Filmstrip and Transcript to seek to information differently. The Filmstrip was used for navigation of visual content, such as seeking through the video to find a specific PowerPoint slide, whereas the Transcript was used for scanning for keywords and highlighting. Given that the content was video lectures, it is not surprising that 76.6% of all seeks were made using the Filmstrip. Interviewees explained that the Filmstrip was useful for searching for visual content. The interviewees, for example, explained they would seek through the video for specific PowerPoint slides and transitions to new slides. Once they found the slide they sought, the interviewees explained they would copy the information using a word processor, a paper notebook, or by taking screenshots

It seems that the two user interface elements were used for different types of information seeking, suggesting that seeking video

and text support different information needs. The median seek distance using the Filmstrip was 18.27 seconds, whereas the median seek using the Transcript was 108.03 seconds. The Filmstrip seeks may have been used to fine-tune the current playhead location: for example, seeking backwards to repeat the last few spoken words or jumping forwards to search for a transition to a new PowerPoint slide. In contrast, Transcript seeks were often larger and occasionally made across the entire length of the video. This may have been the result of finding important textual information while ignoring the current position of the playhead. Indeed, the Transcript allowed students to navigate and highlight the content without watching it. An interviewee explained he would scan the Transcript from top to bottom for keywords to identify which segments of video to watch. Unlike the Filmstrip, the Transcript was used for both navigation and highlighting. Thus, it comes as little surprise that when jumping to highlights, the Transcript was used more often (79.7%) than the Filmstrip (20.3%).

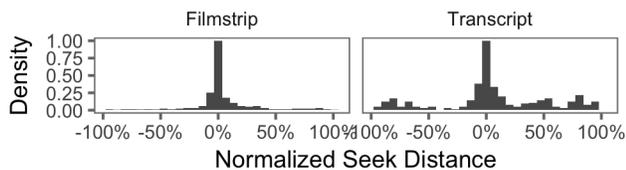


Figure 4: Jumping direction and distance by user interface element. The Filmstrip was used for shorter jumps than the Transcript.

6 CONCLUSION

The results suggest that students made strategic use of the textual and visual user interface elements. We see two main differences with regard to students' interaction with textual information (using the Transcript) and visual information (using the Filmstrip). Most seeks were made using the Filmstrip. This is likely because of the rich cues aided seeking to visual changes in the video, such as PowerPoint slide changes, and since students are used to seeking video using the horizontal bar beneath it in other video viewing environments, such as YouTube. Highlighting, on the other hand, was conducted primarily via the Transcript. Similarly, when jumping to highlights, the Transcript was used most often.

A more careful analysis suggests a specific use for highlights — that of bookmarks for future watching, based on textual keywords. This is suggested by seeing that students (i) highlighted very short spans, (ii) did so using the Transcript, and (iii) highlighted far away from the current location of the playhead. This interpretation is supported by students' interviews. One explanation for these patterns is that students brought in their work habits, using the tools in familiar ways and not making use of the linked nature of these two elements in ViDeX. It is possible that better integration of the two elements, such as a vertical Filmstrip adjacent to the Transcript, would help learners see the connection between the two. However, another possibility is that this behavior demonstrates a desired and adaptive behavior. Students knew that they could use both elements for seeking, and they were strategic about which element to use.

The different patterns may make sense, given the different goals of seeking. It is important to remember that the videos used in this study were highly verbal, which invites this kind of behavior. In that case, the fact that ViDeX supported different uses of its elements to achieve different goals may suggest a useful implementation.

To summarize, we identified strategic use of textual and visual user interface elements for different types of information interactions (short and long seeking, jumping to highlights, and highlighting). Specifically, we found that highlighting of video is very different from highlighting text. For one, it does not seem to be very intuitive. Also, these highlights tend to be shorter, while typical text highlights often select phrases, sentences, or paragraphs. Video highlights also appeared to serve only one function, unlike text highlights. Further research and support for learners are needed to better understand how to make video highlighting a more effective tool for interaction with video content.

REFERENCES

- [1] Mortimer J Adler and Charles Van Doren. 1972. *How to read a book*. Simon & Schuster, New York, NY.
- [2] Thomas H Anderson and Bonnie B Armbruster. 1984. Studying. In *Handbook of Reading Research*. Pearson, New York, NY, 657–679.
- [3] David C Caverly and Vincent P Orlando. 1991. Textbook study strategies. In *Research and Advanced Technology for Digital Libraries*. International Reading Association, Newark, DE, 86–165.
- [4] Michelene T H Chi and Ruth Wylie. 2014. The ICAP framework: linking cognitive engagement to active learning outcomes. *Educational Psychologist* 49, 4 (2014), 219–243.
- [5] Suzanne L. Dazo, Nicholas R. Stepanek, Robert Fulkerson, and Brian Dorn. 2016. An empirical analysis of video viewing behaviors in flipped CS1 courses. *ACM Inroads* 7, 4 (Nov. 2016), 99–105.
- [6] Matthew Fong, Gregor Miller, Xueqin Zhang, Ido Roll, Christina Hendricks, and Sidney Fels. 2016. An investigation of textbook-style highlighting for video. In *Graphics Interface*. 201–208.
- [7] Philip J. Guo, Juho Kim, and Rob Rubin. 2014. How video production affects student engagement: an empirical study of MOOC videos. In *Proceedings of the First ACM Conference on Learning @ Scale Conference (L@S '14)*. ACM, New York, NY, 41–50.
- [8] N Katherine Hayles. 2005. *My mother was a computer: digital subjects and literary texts*. University of Chicago Press, Chicago, IL.
- [9] Juho Kim, Philip J. Guo, Carrie J. Cai, Shang-Wen (Daniel) Li, Krzysztof Z. Gajos, and Robert C. Miller. 2014. Data-driven interaction techniques for improving navigation of educational videos. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology (UIST '14)*. ACM, New York, NY, 563–572.
- [10] Scott LeeTiernan and Jonathan Grudin. 2001. Fostering engagement in asynchronous learning through collaborative multimedia annotation. In *INTERACT*. 472–479.
- [11] Chunyuan Liao, François Guimbretière, and Ken Hinckley. 2005. PapierCraft: a command system for interactive paper. In *Proceedings of the 18th Annual ACM Symposium on User Interface Software and Technology*. ACM, New York, NY, 241–244.
- [12] Catherine C Marshall. 2009. *Reading and writing the electronic book*. Morgan & Claypool Publishers, San Rafael, CA.
- [13] Iliia A Ovsianikov, Michael A Arbib, and Thomas H McNeill. 1999. Annotation technology. *International Journal of Human-Computer Studies* 50, 4 (1999), 329–362.
- [14] Frank Shipman, Morgan Price, Catherine C Marshall, and Gene Golovchinsky. 2003. Identifying useful passages in documents based on annotation patterns. In *Research and Advanced Technology for Digital Libraries*. Springer, Berlin, Germany, 101–112.
- [15] Michele L Simpson and Sherrie L Nist. 1990. Textbook annotation: an effective and efficient study strategy for college students. *Journal of Reading* 34, 2 (1990), 122–129.
- [16] Craig Tashman and W Keith Edwards. 2011. Active reading and its discontents: the situations, problems and ideas of readers. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 2927–2936.
- [17] Craig Tashman and W Keith Edwards. 2011. LiquidText: a flexible, multitouch environment to support active reading. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, New York, NY, 3285–3294.