Received: 25 October 2022, Accepted: 20 June 2022, Published: 30 June 2022
http://dx.doi.org/10.17576/ajtlhe.1401.2022.05

ASSESSING STUDENTS' FOCUS AND INTEREST USING YOUTUBE ANALYTICS

Zamira Hasanah Zamzuri[1*]

[1]Department of Mathematical Sciences, Faculty of Science and Technology
Universiti Kebangsaan Malaysia, 43600 UKM Bangi, Selangor, Malaysia
***(Corresponding author: zamira@ukm.edu.my)**

Abstract

Knowing the fact whether your students are paying attention in the class is a valuable information that helps educators to reassess their teaching quality. With the recent global situation, opting to the recorded or online streaming session has become more common and needed. Various platforms are available in order to share these sessions and one of the most commonly used is YouTube. One attractive property offered by YouTube is the analytics are readily available for each video with various important and interesting facts. From these facts, the educators are able to assess the students' focus and further strategizing for the future lessons. Based on the data of 22 videos uploaded on October 2020 to January 2021, the average view of duration for each video is around 10 to 12 minutes, which is approximately one third from the total duration of the video. The students were giving more attention to the demonstration sessions, when white board being used and visualizations are shown in the videos. Less interest is observed from the students when the video contents are chapters' summary or the formulas.

Keywords: data analytics; lesson videos; students' focus; YouTube

Abstrak

Mengetahui samada pelajar anda menumpukan perhatian di dalam kelas merupakan maklumat berharga yang dapat membantu pengajar untuk menilai semula kualiti pengajaran. Dengan situasi global semasa, memilih untuk merakam atau penstriman dalam talian sudah menjadi kebiasaan malahan keperluan. Pelbagai platform tersedia untuk tujuan ini yang mana YouTube merupakan salah satu pilihan yang kerap digunakan. Satu kelebihan menarik yang ditawarkan oleh YouTube adalah analitik yang tersedia bagi setiap video yang membekalkan pelbagai maklumat penting dan menarik. Berdasarkan maklumat ini, pengajar mampu untuk menilai fokus pelajar dan seterusnya membina strategi untuk sesi pengajaran dan pembelajaran di masa akan datang. Berdasarkan 22 video yang dimuat naik dari Oktober

2020 sehingga Januari 2021, didapati purata masa menonton bagi setiap video adalah sekitar 10 ke 12 minit, yang mana hanya sepertiga dari jumlah tempoh masa video. Pelajar memberi lebih perhatian kepada bahagian demonstrasi, apabila papan putih digunakan serta paparan visualisasi di dalam video. Pelajar menunjukkan kurang minat kepada kandungan video seperti formula atau ringkasan dari bab.
Kata kunci: analitik data; fokus pelajar; video pelajaran; YouTube

1.0 INTRODUCTION

The ultimate aim in any teaching and learning activities is to ensure that the students able to achieve the learning outcome set for the session, for example comprehending the concept or able to replicate the examples given; and the students' focus is one of the main ingredient to this aspect. During the classes, it is difficult to judge students' focus and attention solely based on the interactions in the class or the students' reaction and body language. With the advent of technology and adapting to the current situation with the existence of the pandemic, recorded lectures or lesson videos has become a popular alternative in the education world. Along with the videos properties, the students' focus and attention can be measured from their interactions navigating the videos which give the educator more information on the engagement and retention of the lesson session. Using lesson video analytic is not relatively new, as can be found in Draus et al. (2014), Sinha et al. (2014) and Mirriahi and Vigentini (2017), and in more recent papers using the data from YouTube Analytics (Farrell, 2021; Walsh et al., 2019; Emerson et al., 2019).

Due to the lockdown restriction, imposed for the first time in Malaysia on March 2020, courses taught at Universiti Kebangsaan Malaysia have been conducted online for almost three semesters. One of the approaches employed in this remote learning experience is to record the lecture videos and shared in video sharing platform, such as YouTube.

Based on Bonafini et al. (2017), patterns and behavior of the students while navigating the videos can be defined as their engagement on the learning process. Kadoic and Oreski (2021) measures the students' engagement on learning videos through a set of variables extracted from YouTube Analytics including the number of spikes and dips. Meanwhile, Guo et al. (2014) use how long the students watch the video as a measurement of the students' engagement. Interactions during the video's navigation may provide certain insight on the viewers' focus and interest towards the video. Actions such as forwarding, rewinding or repeating the videos cannot be executed during a live session. Hence, through the viewers' or the video's navigators action towards the video, their focus can be measured. This is the main objective of this paper in which we want to measure and gain further insight on the students' focus, attention and interest to the lecture videos shared using YouTube platform.

2.0 MATERIALS AND METHODS

2.1 Data Source

The data used in this paper is based on 22 lecture videos for the course STQS4113: Applied Multivariate uploaded in YouTube. This course was taught in the first semester of the 2020/2021 academic session, in which the videos were uploaded from October 2020 to January 2021. There were 38 fourth year students who taken this course under Bachelor of Science (Statistics) program. There are two videos uploaded each week during the scheduled day of class for the students to watch, followed by an interaction session for any query.

2.2 YouTube Analytics

YouTube is one of the most popular platform for sharing online content, specifically recorded videos. The content varies from entertainment, social awareness, kids and also education. Before a video is uploaded to the platform, there are three types of privacy setting for the video's visibility; which are public, unlisted and private. Each video uploaded to the YouTube will be equipped with updated information related to the video visibility and engagement through the YouTube Analytics Tab. Information provided from YouTube Analytics are then partitioned into three categories with variables as shown in Table 1.

Table 1: *The categories and variables provided in YouTube Analytics*

Categories	Variables
Engagement	Number of views, Average view duration, Average percentage viewed, Number of spikes, Number of dips, Number of likes, Number of dislikes
Reach	Traffic source types, Impressions and how they led to watch time
Audience	Number of unique viewers, average views per viewer and number of subscribers

In this paper we extract the variables from the overview section that combines these three categories. The variables extracted are average view duration, average percentage viewed, number of spikes and number of dips. The number of spikes refers to the frequency of parts in which more viewers watched from the previous parts whereas the number of dips Is the opposite of it. The videos considered for this study are the unlisted videos, which means the video cannot be accessed publicly, hence the other three variables under the engagement category; the number of view, the number of likes and the number of dislikes are not being considered.

3.0 RESULTS AND DISCUSSION

We present the statistical descriptive analysis for the data extracted from YouTube analytics for the 22 lecture videos. For each chapter, the concept and mathematical formulation are presented followed by the examples of application. After that, a session on performing the analysis in R is explained in order for the students be able to replicate the example and know how to perform the analysis numerically. Hence the videos format can be categorized into three: concept and theory only, R demonstration session only or combination of both. For each video, we extract the variables as explained in the previous section. Then, we look at two parts which are the view duration and number of spikes and dips, and discuss the findings in relation to the students' focus, attention and interest towards the videos.

3.1 The video and view duration

Most of the lecture videos for this course have duration around 30 – 40 minutes. As shown in Figure 1, the mode for the video duration is a little over 35 minutes with the maximum is around 50 minutes. However, when we look at Figure 2, the view duration tops at around 12 minutes, which is about a third of the video duration. This shows that the students' attention towards the video can be measured at around 34%, in which the content that delivered to the students from the video is not at the full capacity, not even half. We further look at the scatter plot between the video duration and average view duration, in which a positive linear relationship observed. The correlation computed for these two variables is 0.78, quite a strong relationship. We infer that given a longer video duration, a longer average view duration is observed; however, this does not help in strategizing in terms of improving students' focus on the videos. As we have pointed out before, the average view duration is only a third of the whole video duration. And it is important to highlight that the video with longest duration (53 minutes) only have the view duration about 12 minutes. This is the downfall on using videos to teach as the viewers or the students can choose what they want to hear and view and what they don't. It is also support the findings by Guo et al. (2014), Zainuddin and Attaran (2016) and Eick and King Jr (2012), that lesson videos should be concise and short, in order to help with the students' focus and attention.

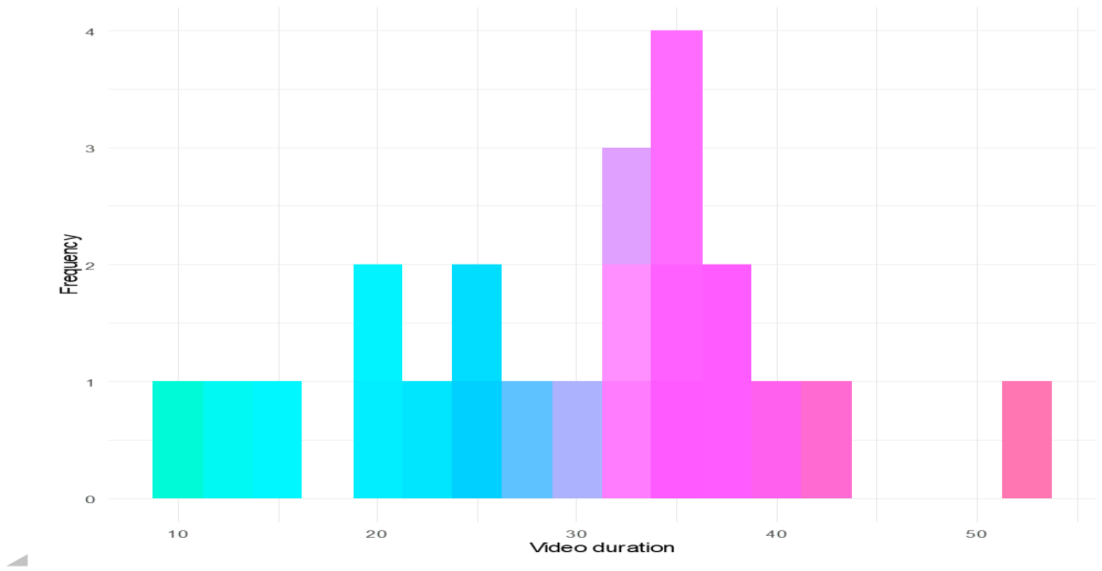

Figure 1: The histogram of the videos duration

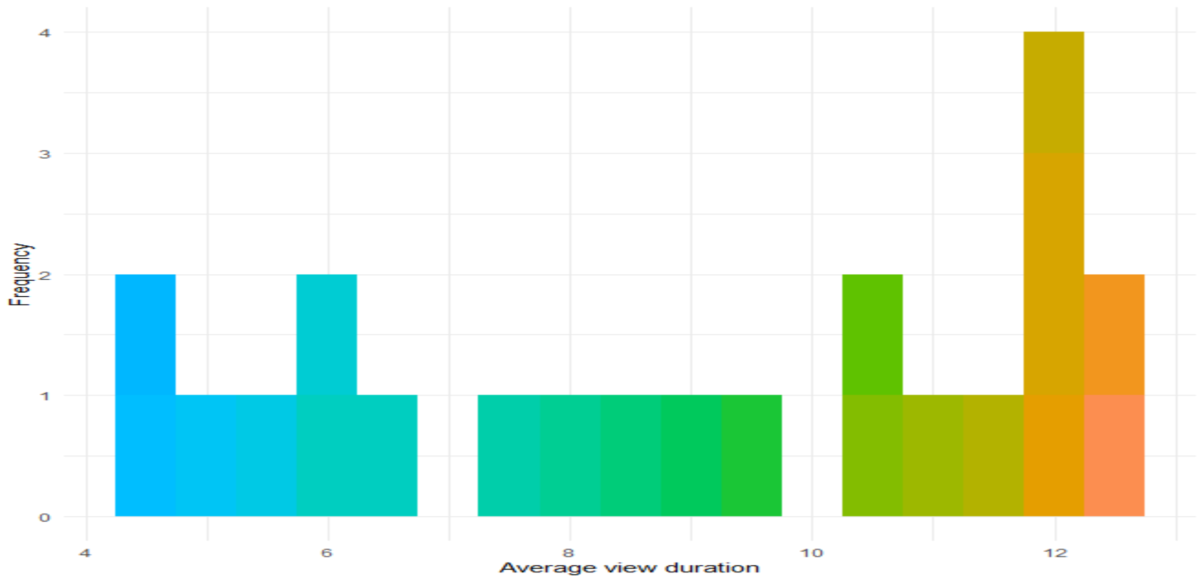

Figure 2: The histogram of the average view duration

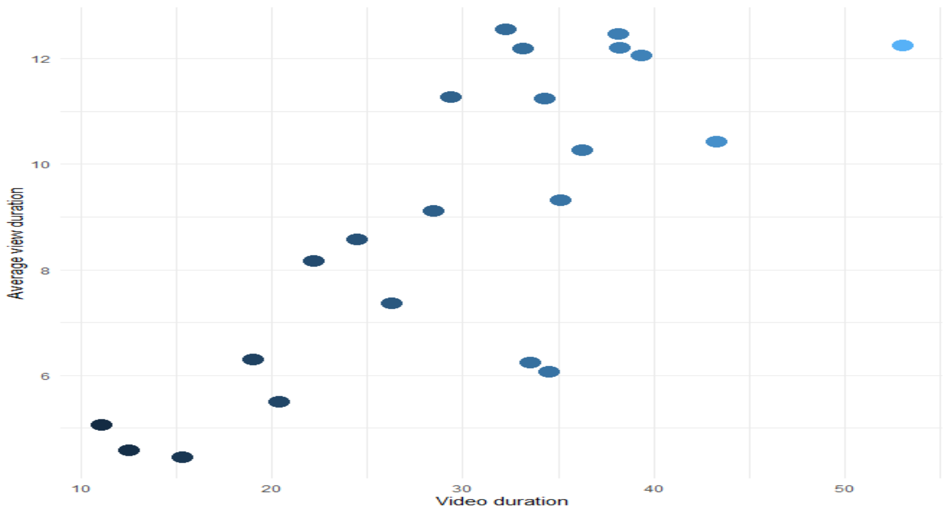

Figure 3: The scatter plot of the video duration vs the average view duration

3.1 The number of spikes and dips

The number of spikes and dips provided in YouTube Analytics are relevant info in conjunction with the behavior of video navigation, in which the viewer can choose which part the want to watch or to skip. When more viewers are watching a certain part of the video, then a spike will appear while dips happening when viewers choose to abandon or skip that specific part of the video. In this part of analysis, for each video, we extract the number of spikes and dips. Higher number of spikes than dips infer that the video has more attention – appealing parts. Using number of spikes and dips in measuring videos audience engagement and retention can be found in a number of studies such as educational (Kadoic & Oreski, 2021; Kim et al. 2014), fitness (Sui et al., 2022) and marketing (Richardson & Vallone, 2014; Grundy & Grundy, 2018). More similar research can be found in the literature as ones conducted by Li et al. (2015) in which they evaluate how click actions (pauses, seeking, skipping, replaying) as a pattern of interactions and Sinha et al. (2014) use click sequences to predict in-video dropouts.

Adopting the similar approach, we then look further what is the content when most of the viewers skipped or watch and categorize the content whether it is a concept explanation, examples, R demonstration session or other types of content. Based on Figure 4, the maximum number of spikes in a video is 9 whereas for dips, is 7. The mode for the number of dips is 1, whereas for the number of spikes are 2 and 4. This indicates that the videos analyzed in this paper are considered interesting to the viewers, since the number of spikes are greater than the number of dips, on average per video. The videos' viewers chose to focusing on the parts of the videos that are of their interest, hence the higher number of spikes compared to dips.

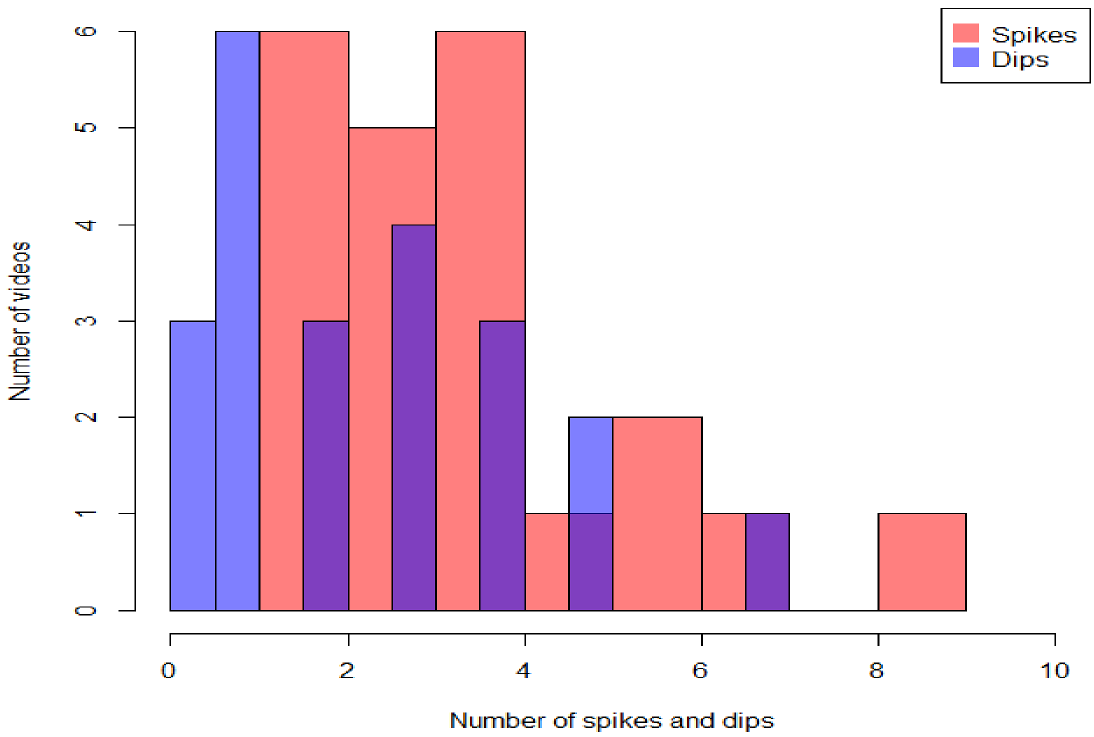

Figure 4: Overlapped histograms for the number of spikes and dip

We dig further by itemizing the content of the video for each spikes and dips. The total number of spikes is 83 whereas for dips, it is 53 for all 22 videos, in which it is about 4 spikes and 2 dips per video. The results are then summarized in Table 2.

Table 2: *The numbers of spikes and dips for each content type*

	Spikes			Rank		
Content	Frequency	%	Rank	Frequency	%	Rank
Instructions	1	1.205	7	0	0	7
Summary/ Intro	0	0	8	6	11.321	3
White Board	6	7.229	6	0	0	7
R Session	29	34.940	1	6	11.321	3
Visualization	7	8.434	5	0	0	7
Examples	10	12.048	3	8	15.094	2
Concept	20	24.096	2	18	33.962	1
Interpretation	9	10.483	4	4	7.547	5
Formula	1	1.205	7	5	9.434	4
Further explanation	0	0	8	1	1.887	6
Zoom in	0	0	8	5	9.434	4

Through inspecting Table 2, the content with the highest number of spikes is the R sessions, with 29 spikes. This signify that the students are looking forward to the demonstration sessions using R to employ the analysis. There is also dips for the R sessions which is 6, perhaps for the session in which the steps are trivial and already known by the students. As using R specifically to perform the mathematical computation in the subject taught is considered new and more complex to the students, then it raises more interest hence the increment in the number of spikes. Shirey and Reynolds (1988) fundamentally found that the complexity level and interest are correlated, in which a more complex content can raise more interest on the students. Contents such as concepts explanation, interpretation and examples have the almost equal number of spikes and dips, as these types of content are the ones that are commonly in the videos, hence the results. It shows that the students know the importance of concept, interpretation and explanation since they are paying attention to them but being selective as shown by the existence of the number of dips. This is aligning with findings in Kim et al. (2014), in which peaks are observed when students want to replaying a brief segment, repeating a non-visual explanation and returning to missed content. Another important point inferred from this table is the students show interest when the white board is being used (6 spikes, 0 dips) and when any visualizations is being shown (7 spikes, 0 dips). It is apparent to us; these two types of content can grab the students' attention to focus on the video. Formulas, chapters intro and summary and video editing error (zoom in situations) are the ones that being skipped and ignored by the students as shown by the higher number of dips compared to spikes for these categories. Consistently, also pointed out by Kim et al. (2014), video length, abrupt visual transitions, and interface characteristics as reasons for the dips or in-video dropouts.

4.0 CONCLUSION

In this paper, we utilize the information provided through YouTube Analytics. Narrowing down the scope to assess the students' focus and interest, we further analyze the data based on the view duration and numbers of spikes and dips. We found that the students only watch a third of the videos, and this supports the known fact of recording long lesson videos are not effective. Shorter lesson videos are preferable and can deliver the information in a more effective manner. Although only a third of the videos being watched, more interest was observed from the students as measured through the number of spikes and dips. Since the number of spikes are higher on average, this indicates that the students are more interested to the videos.

We also examine the types of content in which a spike or a dip occurred. Through this analysis, we found that the demonstration session, the usage of white board and visualizations are able to capture the students' interest and attention. Contents such as formulas and chapters summaries are the ones that are often ignored by the students. This does not mean that a good video should only contain the white boards, demonstration and visualizations; but it may help in strategizing in developing a more effective learning videos. For example, a brief peek to the formulas during the demonstration session may help in garnering the students' attention to the formulas. As concluded by Van der Sluis et al. (2016), it is still challenging to understand why certain characteristics of videos affect student behavior but what is more important is how we can turn this understanding into appropriate online interventions in support a student's learning process. Integrating the findings from this paper with other similar studies may contribute to a proposed guideline on creating "good" online learning videos. As mentioned by Guo et al. (2014), 'less than 6 minutes' guide is a popular recommendation for online learning videos, as found through analyzing dwelling times. As concluded in Bruff et al. (2013), Kim et al. (2014) and Li et al. (2015), students tend to become selective on certain parts of the educational video and re-watching the video. This justify the use of peaks and dips as the measurement since the action of selecting and skipping parts of the video indicate the pattern of students' behavior towards the video. It then further impact on educator's strategies to manipulate these features in order to produce an effective and highly functional learning session.

Findings from this paper can shed some lights to the educators in reassessing their teaching technique in lesson videos and further strategizing for improved teaching and learning experience in the future. It is also important to highlight that the analysis conducted in this study is at the descriptive level and can be considered preliminary, hence the findings only can be inferred to the context of the study, final year students who taking a mathematical subject. Further research is needed to take into account other factors so that the findings can be applied to a broader context. The inclusion of other contributing factors should also be considered in the future. For example, perceived difficulty on the subject affecting the students' engagement in a negative way as found in Li et al. (2015).

5.0 ACKNOWLEDGEMENTS

We would like to thank Faculty of Science and Technology, Universiti Kebangsaan Malaysia for sponsoring the financial means to participate in presenting the work in this paper.

6.0 REFERENCES

Bonafini, F., Chae, C., Park, E., & Jablokow, K. (2017). How Much Does Student Engagement with Videos and Forums in a MOOC Affect Their Achievement?. *Online Learning Journal*, *21*(4).

Bruff, D. O., Fisher, D. H., McEwen, K.E. & Smith, B. E. (2013). Wrapping a MOOC: Student perceptions of an experiment in blended learning. *MERLOT Journal of Online Learning and Teaching 9, 2*(2013), 187–199.

Draus, P., Curran, M., & Trempus, M. (2014). The influence of instructor-generated video content on students' satisfaction with and engagement in asynchronous online classes. *Journal of Online Learning & Teaching, 10*(2), 240-254.

Eick, C. J & King Jr., D. T. (2012). Nonscience Majors' Perception on the Use of YouTube Video to Support Learning in an Integrated Science Lecture. *Journal of College Science Teaching*, *42*(1), 26-30.

Emerson, E., Tansley, G. & Javanbakht, Z. (2019). Blended learning in Engineering Design: Using YouTube analytics to track student engagement with online content. *The 30th Annual Conference for the Australasian Association for Engineering Education (AAEE 2019).*

Farrell, M. (2021). Using YouTube Analytics to Measure the Effectiveness of Instructor-Generated Video in Online Courses. *The 11th International Conference on Learning Analytics and Knowledge.*

Grundy, P. & Grundy, T. (2018). Videos: Documenting the Production of Best Practice Australian Cotton. *CRDC Final Report.*

Guo, P. J., Kim, J. & Rubin, R. (2014). How video production affect student engagement: an empirical study of MOOC videos. *Proceedings of the first ACM conference on Learning*, 41-50.

Kadoic, N. & Oreski, D. (2021). Learning Analytics of YouTube Videos Linked to LMS Moodle. 2021 44th International Convention on Information, Communication and Electronic Technology (MIPRO).

Kim, J., Guo, P. J., Seaton, D. T., Mitros, P., Gajos, K. Z., & Miller, R. C. (2014). Understanding In-video Dropouts and Interaction Peaks in Online Lecture Videos. *In Proceedings of the First ACM Conference on Learning @ Scale Conference (L@S '14).* ACM, New York, NY, USA, 31–40.

Li, N., Kidzinski, L., Jermann, P. & Dillenbour, P. (2015). MOOC Video Interaction Patterns: What Do They Tell Us? In Design for Teaching and Learning in a Networked World, Gránne.

Conole, Tomaž, K., Christoph, R., Johannes, K., & ˇÉlise, L., (Eds.). *Lecture Notes in Computer Science*.

Mirriahi, N. & Vigentini, L. (2017). Analytics of Learner Video Use in C. Lang. Handbook of Learning Analytic. Siemens, A. Wise & D. Gasevic (Eds).

Richardson, A. & Vallone D. M. (2014). YouTube: a promotional vehicle for little cigars and cigarillos?. Tobacco Control 23: 21-26.

Shirey, L. L & Ralph, E. R. (1988). Effect of interest on attention and learning. *Journal of Educational Psychology, 80*(2), 159.

Sinha, T., Jermann, P., Li, N. & Dillenbourg, P. (2014). Your click decides your fate: Inferring Information Processing and Attrition Behavior from MOOC Video Clickstream Interactions. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 3–14.

Van der Sluis, F., Ginn, J., & Van der Zee, T. (2016, April). Explaining student behavior at scale: the influence of video complexity on student dwelling time. In *Proceedings of the third (2016) acm conference on learning@ scale* (pp. 51-60).

Sui, W., Rush, J., & Rhodes, R. E. (2022). Engagement With Web-Based Fitness Videos on YouTube and Instagram During the COVID-19 Pandemic: Longitudinal Study. *JMIR formative research*, *6*(3), e25055.

Walsh, J. N., O'Brien, M. P. & Slattery, D. M. (2019). Video Viewing Patterns Using Different Teaching Treatments: A Case Study Using YouTube Analytics. Research in Education and Learning Innovation Archives, 22.

Zainuddin, Z. & Attaran, M. (2016). Malaysian students' perception of flipped classroom: a case study. *Innovations in Education and Teaching International*, *53*(6), 660-670.